the FEBS Journal

Integrase of Mason–Pfizer monkey virus

Jan Snášel[1,2], Zdeněk Krejčík[1,2], Věra Jenčová[1,2], Ivan Rosenberg[1], Tomáš Ruml[1], Jerry Alexandratos[3], Alla Gustchina[3] and Iva Pichová[1]

1 Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague, Czech Republic
2 Institute of Molecular Genetics and Center for Integrated Genomics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
3 Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD, USA

Keywords
integrase; Mason–Pfizer monkey virus; HIV-1; specificity; structure

Correspondence
I. Pichová, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo n. 2, 166 10 Prague 6, Czech Republic
Fax: +42 02 20183556
Tel: +42 02 20183251
E-mail: iva.pichova@uochb.cas.cz

(Received 21 July 2004, revised 21 September 2004, accepted 22 September 2004)

doi:10.1111/j.1432-1033.2004.04386.x

The gene encoding an integrase of Mason–Pfizer monkey virus (M-PMV) is located at the 3′-end of the pol open reading frame. The M-PMV integrase has not been previously isolated and characterized. We have now cloned, expressed, isolated, and characterized M-PMV integrase and compared its activities and primary structure with those of HIV-1 and other retroviral integrases. M-PMV integrase prefers untranslated 3′-region-derived long-terminal repeat sequences in both the 3′-processing and the strand transfer activity assays. While the 3′-processing reaction catalyzed by M-PMV integrase was significantly increased in the presence of Mn^{2+} and Co^{2+} and was readily detectable in the presence of Mg^{2+} and Ni^{2+} cations, the strand transfer activity was strictly dependent only on Mn^{2+}. M-PMV integrase displays more relaxed substrate specificity than HIV-1 integrase, catalyzing the cleavage and the strand transfer of M-PMV and HIV-1 long-terminal repeat-derived substrates with similar efficiency. The structure-based sequence alignment of M-PMV, HIV-1, SIV, and ASV integrases predicted critical amino acids and motifs of M-PMV integrase for metal binding, interaction with nucleic acids, dimerization, protein structure maintenance and function, as well as for binding of human immunodeficiency virus type 1 and Rous avian sarcoma virus integrase inhibitors 5-CI-TEP, DHPTPB and Y-3.

Mason–Pfizer monkey virus (M-PMV) was originally isolated from a spontaneous mammary carcinoma in a rhesus monkey [1]. While this exogenous virus has not been demonstrated to be oncogenic [2], it has been associated with an acquired immunodeficiency syndrome in macaques [3,4]. M-PMV, together with mouse mammary tumor virus, simian retrovirus, squirrel monkey retrovirus, and Jaagsiekte sheep retrovirus represent genus *Betaretrovirus*, and exhibit a D-type morphology, i.e. form immature capsids within the host cells. The process of integration and characterization of the integrase have not been elucidated in these types of retroviruses. The genome of M-PMV consists of four genes: 5′-gag-pro-pol-env 3′. The gene encoding integrase is located at the 3′-end of the pol and thus two ribosomal frameshifts within the overlap of the gag-pro and pro-pol are necessary to yield the Gag-Pro-Pol polyprotein [5]. The M-PMV protease specifically cleaves this precursor to yield integrase, reverse transcriptase, and a few structural proteins. Several integrases of other retroviruses have been isolated and their activities characterized, i.e. integrase of AMV [6], HIV-1 [7],

Abbreviations
AMV, avian myeloblastosis virus; ASV, Rous avian sarcoma virus; CA, capsid protein; CAEV, caprine arthritis-encephalitis virus; EIAV, equine infectious anemia virus; FIV, feline immunodeficiency virus; HFV, human foamy virus; HSRV, human spumaretrovirus; HTLV I and II, human T-cell leukemia virus type I and II; LTR, long-terminal repeat; MA, matrix protein; MLV, murine leukemia virus; Mo-MuLV, Moloney murine leukemia virus; M-PMV, Mason–Pfizer monkey virus; U3, untranslated 3′-region; U5, untranslated 5′ region; wt, wild type.

and HIV-2 [8], CAEV and MVV [9], HTLV I [10] and HTLV II [11], FIV [12,13], ASV [14], and HSRV [15].

Integration of retroviral cDNA into the host cell chromosome catalyzed by integrase is crucial for virus replication. Therefore, the enzyme has become an attractive novel target for antiviral drug design [16]. The integration proceeds *in vivo* and *in vitro* in three steps. In the 3′-processing reaction, two nucleotides are removed from each cDNA 3′-end and the newly generated 3′-hydroxyl groups provide the sites for joining with the 5′-ends of the target host DNA in the strand transfer reaction. The product of integration is a gapped intermediate in which the nonjoined 5′-viral DNA ends are flanked by short single-stranded gaps in the host DNA. Removal of mispaired nucleotides and gap repair are carried out by cellular enzymes [17].

Retroviral integrases contain two known metal-binding domains. The N-terminal domain includes a zinc-finger motif and the central catalytic core domain contains a triad of acidic amino acids that bind Mn^{2+} or Mg^{2+}, the metal cofactors necessary for enzymatic activity. Binding of zinc to the N-terminal part enhances multimerization of the native enzyme and increases its enzymatic activity [18]. Crystal structures of the catalytic cores or two-domain derivatives of several integrases have been determined in the absence and presence of bound inhibitors and/or metal ions [19–24]. The three-dimensional structures of the individual N- and C-terminal domains were determined by NMR spectroscopy [25–27].

Here we show that M-PMV integrase with and without a His-tag at the C-terminus display the identical 3′-processing, strand transfer, and disintegration activities preferentially with U3-derived sequences, leading to the conclusion that the His-tag does not influence enzymatic activities of M-PMV integrase. The protein catalyzes the cleavage of M-PMV and HIV-1 long-terminal repeat (LTR)-derived substrates with very similar efficiency.

Results

Determination of the N-terminal sequence of M-PMV integrase

DNA sequences encoding M-PMV integrase are located at the 3′-end of the pol reading frame. An alignment of the amino acid sequences of ASV, SIV, and HIV integrases predicted the N-terminal sequence of M-PMV integrase as Ile-Asn-Thr-Asn. To determine the precise N-terminus of M-PMV integrase, we have used the property of the retroviral proteases to cleave the polyprotein precursors into functional proteins and enzymes. The DNA encoding the predicted integrase and a substantial part of the 3′-end of the gene encoding the reverse transcriptase were cloned into a bacterial expression vector. The precursor was isolated from inclusion bodies and conditions for cleavage with M-PMV protease were optimized. The cleavage was performed at pH 6 in the presence of 0.3 M NaCl. Biochemical characterization of M-PMV protease showed that this protease preserves 80% of the proteolytic activity under these conditions [28]. Edman degradation of the cleavage product with mobility of about 33 kDa revealed the N-terminal sequence Ser-Asn-Ile-Asn-Thr-Asn-Leu-Glu.

Cloning, expression and isolation of M-PMV IN

To simplify the purification, we cloned and expressed integrase with a His_6-tag attached to the C-terminus of the enzyme. To evaluate any influence of the His-tag on the activities of integrase, we also prepared integrase lacking the His anchor. When a standard protocol for bacterial pET expression of proteins at 37 °C was used, the yield of both integrases {[+]His-tag (integrase His-tag) and [–]His-tag (integrase)} was low and the purified proteins were insoluble in common buffers in the absence of urea. The expression of M-PMV integrase His-tag was confirmed by immunoblot analysis with anti His-tag antibodies (data not shown). The solubility of bacterially expressed M-PMV integrases was improved by the decrease of cultivation temperature of transformed bacterial cells to 18 °C. The integrase His-tag, eluted from the Ni-nitrilotriacetic acid column by a gradient of 20–600 mM imidazole in TNM buffer, and concentrated either by ultrafiltration (Amicon membrane; cut off 10 000) or on Centricon filters (cut off 10 000), was soluble only up to 0.1 $mg \cdot mL^{-1}$. Interestingly, the highest concentration of integrase (0.5 $mg \cdot mL^{-1}$) was achieved when integrase His-tag was eluted from the Ni-nitrilotriacetic acid column with 600 mM imidazole (Fig. 1A).

Wild type (wt) integrase was purified using extraction from the homogenized bacterial pellet into the HED buffer with 1 M NaCl, followed by ammonium sulfate precipitation (30% saturation) and chromatography on butyl-Sepharose and heparin-Sepharose columns. Nucleic acids were removed by a phosphocellulose chromatography step (A_{260}/A_{280} ratio was decreased from 1.2 to 0.6) (Fig. 1B). The overall yield of the native enzyme, with or without the C-terminal His-tag, was 3–5 mg.

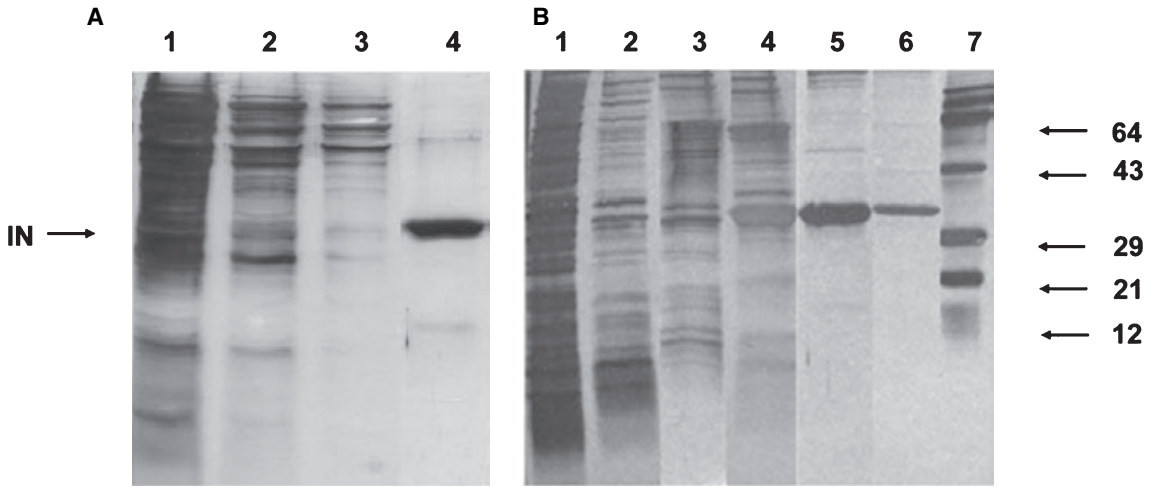

Fig. 1. Purification of M-PMV integrase. (A) Samples from purification of M-PMV integrase-His-tag. Lane 1, total protein from induced cells; lane 2, pellet extracted into TNM buffer with 2 M NaCl; lane 3, flow-through fractions from Ni-nitrilotriacetic acid column; lane 4, protein eluted from Ni-nitrilotriacetic acid column with 600 mM imidazole. (B) Samples from purification of M-PMV integrase. Lane 1, total protein from induced cells; lane 2, pellet extracted with HED buffer containing 1 M NaCl; lane 3, the ammonium sulfate precipitate; lane 4, protein after chromatography on butyl-Sepharose; lane 5, protein after chromatography on Heparin-Sepharose; lane 6, integrase eluted from phosphocellulose; lane 7, molecular mass standards.

Enzymatic activities of M-PMV IN

Both M-PMV integrase and M-PMV integrase His-tag were assayed for 3′-processing and strand transfer activities with M-PMV U5 or U3 LTR-derived substrates labeled at the 5′-end with ^{32}P. The results showed that the 3′-processing reaction catalyzed by M-PMV integrases occurs with both substrates (Fig. 2A,B). However, an analysis of kinetic data showed that the U3 LTR oligonucleotide is a slightly better substrate (with an apparent $K_m' = 58$ nM, $V_{max} = 13$ fmol·min^{-1}) than U5 LTR oligonucleotide (app. $K_m' = 78$ nM, $V_{max} = 10$ fmol·min^{-1}). The concentration of integrase in the assays was determined by the Bradford method [29] and thus represents the total concentration of integrase without discrimination between monomeric or multimeric forms of the enzyme. The experiments also confirmed that the presence of the C-terminally attached His-tag has no influence on the 3′-processing activity of M-PMV integrase.

The 3′-processing activity was stimulated by increasing the temperature. An almost twofold concentration

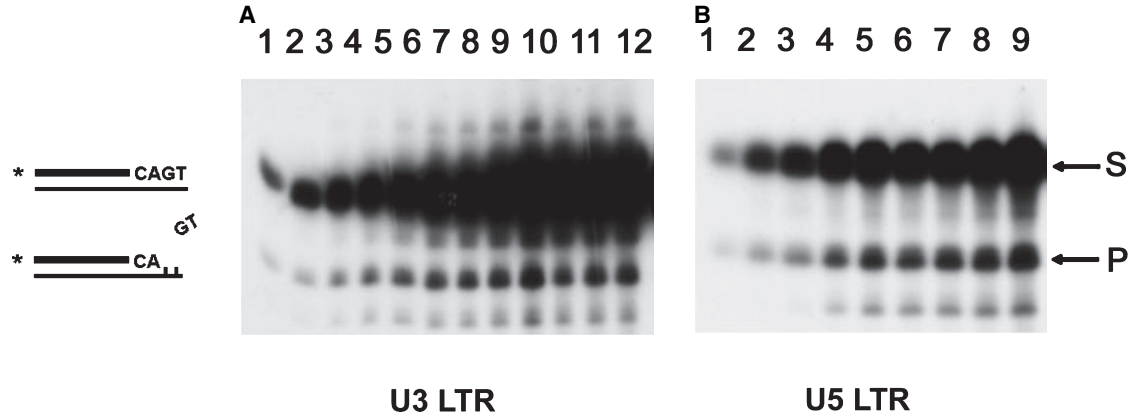

Fig. 2. The 3′-processing activity of M-PMV integrase shown as a function of substrate concentration. (A) Lanes 1–12: the 5′-end ^{32}P-labeled U3 LTR substrate (S) of concentration 5, 10, 20, 30, 40, 60, 80, 120, 140, 160, 180 and 200 nM; (B) Lanes 1–9: the 5′-end ^{32}P-labeled U5 LTR substrate (S) of concentration 5, 10, 15, 20, 30, 40, 50, 70, 90 nM were incubated with 150 nM integrase for 20 min at 37 °C. P, products of the cleavage reactions catalyzed with the integrase.

of product was generated after 50 min of incubation at 37 °C compared to 30 °C. Increasing temperature to 44 °C did not change the reaction rate (data not shown). The cleavage of 30 nM U3 with 150 nM M-PMV integrase was evident after 1 min of incubation and was linear for 15 min at 37 °C.

The analysis of M-PMV integrase integration activity confirmed that joining of substrates catalyzed by M-PMV integrase is much less efficient than that catalyzed by HIV-1 integrase. The products of the integration reaction were visible on gels only after a long exposure time. To enhance the detection of this reaction, we used a 'precleaved' 19-mer U3 and U5 substrates with sequences 5′-ACTGTCCCGACCCGC GGGA-3′ and 5′-GATCCCGCGGGTCGGGACA-3′, respectively. These single stranded 19-mer oligonucleotides were annealed to the complementary 21-mer oligonucleotides. The results showed that the yield of integration reaction catalyzed by integrase was also more efficient with U3 LTR derived substrate and a maximum of products was obtained after 30 min of incubation (Fig. 3). Identical results were obtained for integrase (His-tag), confirming that the His-tag has no influence on the integration activity of M-PMV integrase.

The disintegration reaction representing the reverse reaction of the strand transfer occurs *in vitro* with high efficiency [30]. The significance of disintegration *in vivo* is unclear, but *in vitro* it is the most robust reaction and is performed by many mutated or truncated integrase proteins that display only low or undetectable levels of processing and strand transfer [7]. The sealing of the nick in the target DNA with the substrates catalyzed by M-PMV integrase (see Materials and methods) resulted in the formation of a 30-nt labeled product.

The maxima of 3′-processing and strand transfer activities catalyzed by integrase were detected in the presence of 10 mM Mn^{2+} and 8–15 mM Mn^{2+}, respectively. The 3′-processing activity in the presence of Mg^{2+} was about tenfold lower than that in the presence of Mn^{2+}. Surprisingly, M-PMV integrase exhibits a readily detectable cleavage activity even in the presence of Co^{2+} and Ni^{2+}. Higher conversions of the substrate were achieved in the presence of Co^{2+} compared to Mn^{2+} (Fig. 4). The strand transfer activity is strictly dependent on Mn^{2+}, i.e. only trace levels of autointegration products were obtained in the presence of Mg^{2+} and no products were detected in the presence of Co^{2+} and Ni^{2+} (data not shown). The disintegration activity of M-PMV integrase was reproducible at a manganese ion concentration ranging from 0.2 mM to 35 mM. No activity was detected in the presence of 1–50 mM magnesium (Fig. 5).

The ionic strength significantly influences the activity of M-PMV integrase. The enzyme precipitated in buffers with a concentration of NaCl lower than 25 mM. The highest levels of integrase activity were detected in the presence of 25 mM NaCl. Higher concentrations of salt decreased the activity of M-PMV integrase (Fig. 6) and concentrations above 170 mM NaCl abolished both the 3′-processing and joining reactions. Similar results were reported for M-MuLV and visna virus integrases which were inhibited by 25–100 mM NaCl

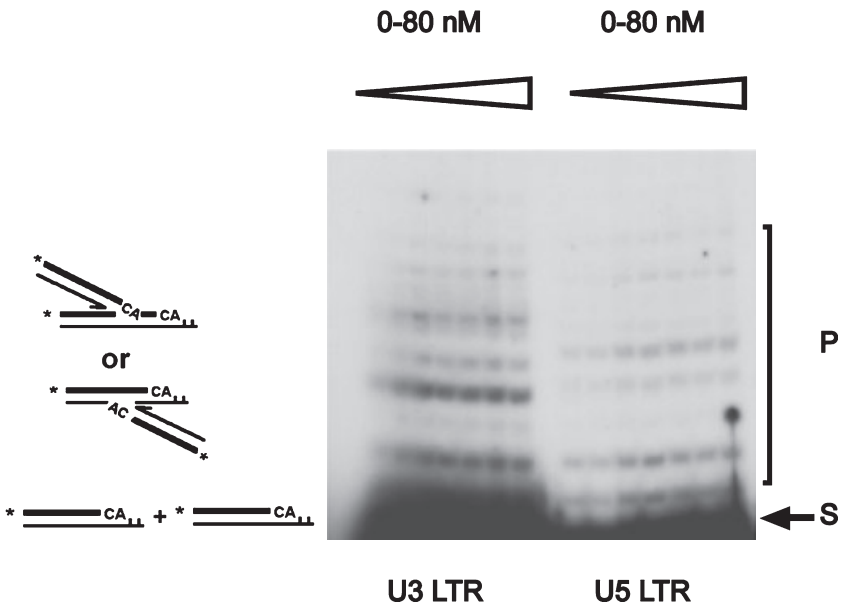

Fig. 3. The strand transfer activity of M-PMV integrase shown as a function of substrate concentration. Integrase at a concentration of 150 nM was incubated with preprocessed U3 or U5 M-PMV LTR substrates (S) at concentration ranging from 0 to 80 nM at 37 °C for 10 min. P, products of the strand transfer reactions catalyzed with the integrase.

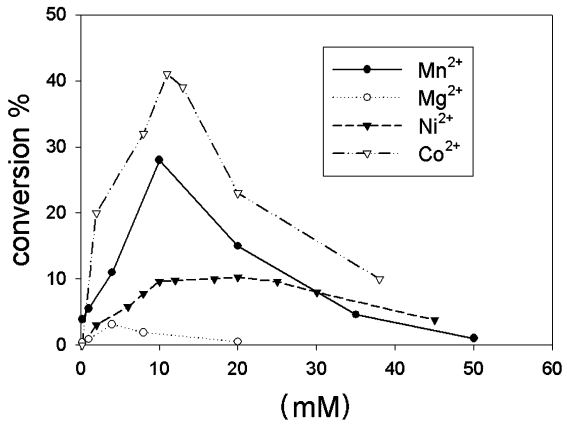

Fig. 4. The effect of Mn^{2+}, Mg^{2+}, Co^{2+} and Ni^{2+} on the M-PMV integrase 3′-processing activity. M-PMV integrase (150 nM) was incubated with 30 nM U3 LTR substrate in the presence of increasing concentrations of different cations for 30 min at 37 °C.

and by 50–150 mM NaCl, respectively [9,31]. AMV integrase displays a higher salt requirement; the maximal activity was detected at 145 mM NaCl [6].

We have confirmed that M-PMV integrase catalyzed reactions are dependent on pH. The 3′-processing activity was readily detected at pH values ranging from pH 7.0–9.0. A dramatic decrease of the 3′-processing activity was observed at pH higher than 9.5 and lower than 7. The optimal strand transfer activity was observed at pH ranging from 6 to 7, basal levels of the activity were noted at pH 8–9, and no strand transfer activity was observed at pH below 5 and pH higher than 9.5 (data not shown). We can conclude that whereas the maximal strand transfer activity was detectable under acidic conditions, the optimal processing activity of M-PMV integrase proceeded at neutral pH. The best conditions for both reactions were found to be at pH 7.4 and 25 mM NaCl. The pH profile of M-PMV integrase catalyzed reactions is similar to those of HIV-1, HTLV I and II, Mo-MLV, and ASV integrases [7,10,11,31,32].

Substrate specificity of integrase-catalyzed reactions

To compare the substrate specificity of M-PMV integrase with that of HIV-1 integrase, we used the integrase's own LTR substrate and an LTR substrate of the opposite virus. Moreover, single-stranded (ss) vs. double-stranded (ds) oligonucleotide substrates were tested.

HIV-1 and M-PMV integrases most efficiently catalyzed the 3′-processing of their own LTR substrates (Fig. 7A). The efficiency of the cleavage of two conserved nucleotides from the single-stranded HIV-1 U5 LTR substrate by HIV-1 integrase was 50% lower than that from the double-stranded substrate. HIV-1 integrase did not process the ds M-PMV U3 LTR but surprisingly generated −1, −2, and −3 products from

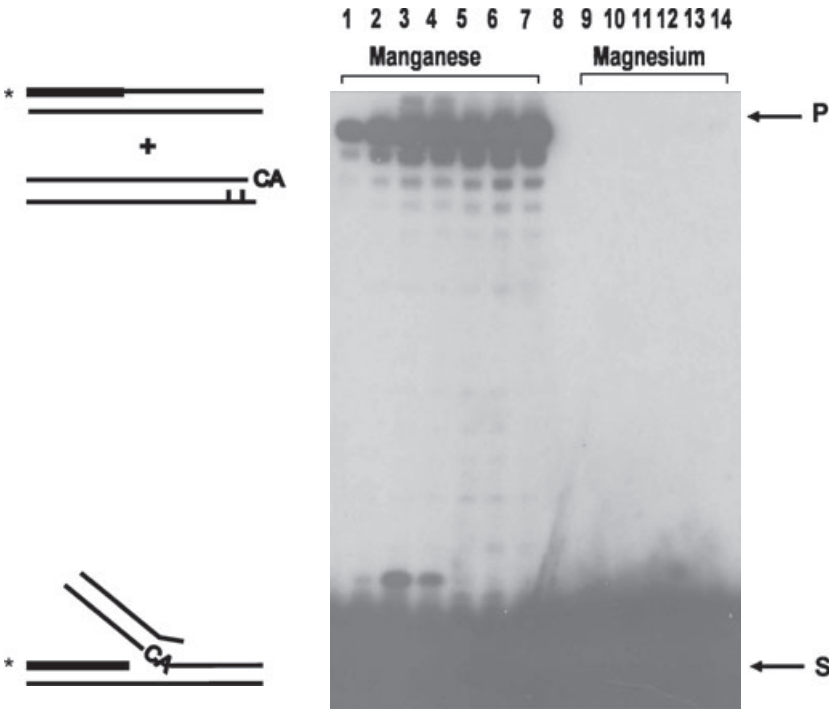

Fig. 5. Disintegration activities of M-PMV integrase as a function of metal ion concentration. Integrase (150 nM) was incubated with 50 nM Y-disintegration substrate (prepared as described in Materials and methods) at 37 °C for 40 min in the presence of: Lanes 1–7: 200 μM, 1 mM, 4 mM, 8 mM, 20 mM, 35 mM and 50 mM Mn^{2+}; lane 8: without metal ions; lanes 9–14: 20 μM, 1 mM, 4 mM, 8 mM, 20 mM and 50 mM Mg^{2+}. In the schematic diagram, oligonucleotide substrates are represented by lines, and the labeled oligonucleotide is in bold.

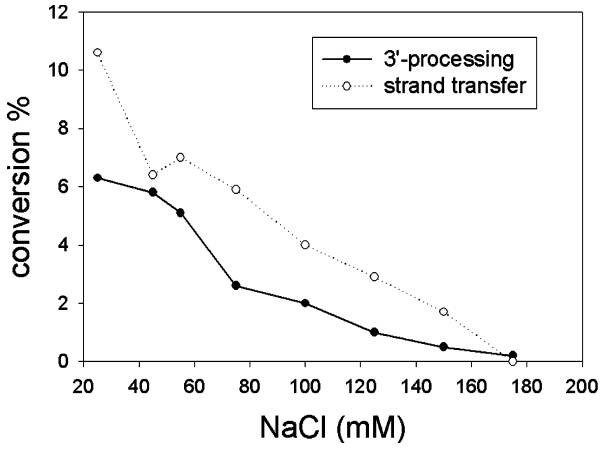

Fig. 6. The influence of ionic strength on 3′-processing and strand transfer activities of M-PMV integrase. M-PMV integrase at 150 nM was incubated for 30 min at 37 °C with 30 nM M-PMV U3 LTR substrate or preprocessed U3 LTR substrate in 20 mM Mops, pH 7.2, containing 50 μM EDTA, 10 mM 2-mercaptoethanol, 10% glycerol (w/v), 7.5 mM $MnCl_2$, 0.1 mg·mL^{-1} BSA, and desired concentration of NaCl.

the ss M-PMV LTR. On the other hand, M-PMV integrase efficiently cleaved both ds U3 M-PMV and U5 HIV LTR substrates (Fig. 7A) and generated the −1 cleavage product from ss HIV-1 LTR. However the cleavage of the ss M-PMV U3 LTR substrate with M-PMV integrase was not detected.

Whereas HIV-1 integrase catalyzed only the covalent joining of its ds blunt-ended LTR substrate, M-PMV integrase integrated both ds M-PMV U3 LTR substrate and weakly ds HIV-1 U5 LTR; however, the integration patterns were slightly different (Fig. 7B). Identical results were obtained when LTR 'preprocessed substrates' were used for an analysis of the strand transfer reaction (data not shown). Similarly to HIV and other retroviral integrases, M-PMV integrase can cleave but not integrate the ss M-PMV U5 and U3 oligomers.

The processing of viral DNA catalyzed by the integrase can be considered a site-specific alcoholysis reaction. HIV-1 integrase was shown to exhibit also a nonspecific alcoholysis, during which the enzyme attacks multiple sites in a target DNA of random sequence (nonviral ds oligonucleotides) and generates product bands other than −2 [33,34]. To prove that M-PMV integrase could catalyze the nonspecific alcoholysis, we used a ds 24-mer oligonucleotide of a random sequence and a (homo)oligonucleotide dT_{10} as substrates. We found that M-PMV integrase, when incubated with ds 24-mer oligonucleotide, generated preferentially three oligonucleotides corresponding to −18, −17, and −16 mers. However, HIV-1 integrase cleaved the same substrate only at the −1 position (not

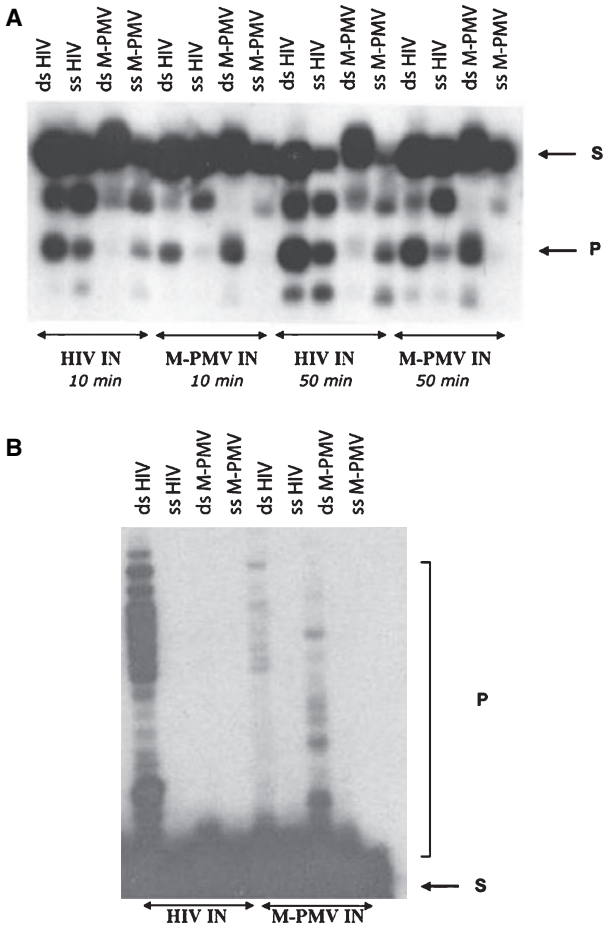

Fig. 7. Substrate specificity of HIV and M-PMV integrases. Enzymes at concentration 150 nM were incubated with 30 nM double and single-stranded LTR derived substrates (S) at 37 °C. (A) The 3′-processing reaction catalyzed with integrases for 10 and 50 min; (B) the strand transfer activity detected after 50 min of incubation. P, products of the cleavage and strand transfer reactions catalyzed with the integrase.

shown). Both integrases cleaved oligonucleotide dT_{10} only in the presence of metal ions and generated 9, 8, and 7-mer oligonucleotides. However, these products were not covalently joined in the reaction catalyzed by integrases, confirming that only the viral DNA ends or oligonucleotides with sequences close to genuine viral DNA ends can be joined by the retroviral enzymes. The exact physiological role of nonspecific nuclease activity of retroviral integrases is not known.

Analysis of the sequence of M-PMV integrase and a comparison with other integrases

The structures of retroviral integrases that have been solved to date show considerable sequence similarity within a common set of three domains with conserved

three-dimensional folds. All known retroviral integrases comprise a zinc-binding N-terminal domain, a catalytic core domain, and a ds DNA-binding C-terminal domain. Although no structure of a full-length retroviral integrase has been published to date, the structures of isolated domains have been solved by X-ray crystallography or by nuclear magnetic resonance. In addition, crystal structures of constructs containing two out of three domains together are also available.

A structurally based sequence alignment of three retroviral integrases was used as a template for subsequent alignment of the sequence of M-PMV integrase that matched the most important structural and functional characteristics of these enzymes. The initial sequence alignment for HIV, SIV, ASV and M-PMV integrases was obtained using the program CLUSTALW [35]. Because the fragments of the compared protein sequences, which contain insertions and deletions, are not usually superimposed accurately by using an automatic mode of alignment, manual corrections were introduced in those parts based on the comparison with the superimposed crystal structures of HIV, SIV and ASV integrases, using the program INSIGHTII 2000 from Accelrys. The three-dimensional structures that were used in structure-based sequence alignment include the single- and two-domain constructs of HIV integrase [24–26,36], and two-domain structures of SIV integrase [37] and ASV integrase [23]. Cα coordinates of the corresponding domains were superimposed using the program ALIGN [38]. The resulting sequence alignment of four integrases allowed us to infer the most important regions of M-PMV integrase and postulate the course of future experiments. The M-PMV integrase numbering scheme used below corresponds to Fig. 8, with HIV integrase or ASV integrase numbering in parentheses, when appropriate.

M-PMV integrase exhibits 13% identity and 31% identity and similarity across this set of four proteins (Table 1). M-PMV integrase shows greater sequence homology with individual integrases, as expected in this group of evolutionarily diverse retroviruses. When examined separately, the individual domains show only slightly different homology characteristics compared to full-length enzymes. The catalytic core domains show slightly higher identity levels than the full sequences, while the C-terminal domains show greater homology than average. This may reflect the higher level of requirement for the conservation of the core residues, which are involved in the catalytic mechanism and the binding of the metal cofactors, as compared with less specific interactions with DNA.

Several metal ions such as Zn^{2+}, Mg^{2+} and Mn^{2+} have been shown to regulate the activity of integrase and affect the stability of the tertiary structure. M-PMV integrase retains the critical amino acid residues for binding metal ions; in the N-terminal zinc-binding domain the HHCC motif is conserved (H14, H18, C42, and C45, corresponding to HIV H12, H16, C40, C43) [26]. Binding of a zinc cation in this domain has been shown to alter and stabilize the overall protein structure, thereby accentuating catalytic activity [39]. The core domain retains the essential DD(35)E motif common to all integrase endonuclease catalytic active sites (D70, D127, E163 in M-PMV and D64, D116, E152 in HIV integrases). The corresponding residues in ASV integrase have been shown to bind catalytic metal cations (Mg^{2+}, Mn^{2+}, or Zn^{2+} among others) [22,40,41]. A noncatalytic residue H103 in ASV integrase, which binds Zn^{2+} and thus stabilizes the local fold [41], is conserved in M-PMV integrase (H110), therefore a similar function can be implied to this residue in the latter.

Another important residue in the active site area is Q148 in HIV integrase. This residue is shown to interact with nucleic acid in this enzyme [42] and is conserved in all integrases that were included in the structure based sequence alignment (Fig. 8). In ASV integrase, the corresponding residue is Q153, while in M-PMV the equivalent residue is Q159. In ASV integrase, Q153 stabilizes the conformation of the active site region by forming a hydrogen bond with a main chain of the catalytic residues.

Other amino acids important for maintaining protein structure and function are also conserved in M-PMV integrase. The N-terminal domain includes a number of key residues implicated in structure stabilization via dimeric contacts, such as I3, N6, L7, E33, R36, Q37, K40, V46, and T47 (HIV F1, L2, I5, V31, K34, E35, A38, Q44, L45) [26]. A comparison of HIV-1 and HIV-2 integrases indicates that the latter part of the secondary structure in this region is significantly less well conserved, based upon variability in the primary structure, but we have noted all residues that have been shown to form N-terminal dimeric contacts in any HIV integrase. A highly conserved serine residue which facilitates a structurally important tight β turn in ASV integrase core (S85) corresponds to S91 in M-PMV integrase, implying a similar conservation of the protein fold in this region [21]. The active site present in the core domain has a highly conserved flexible loop implicated in binding DNA with a 'hinge' formed by two immutable glycines. Both features, the conserved DNA binding residues and hinge glycines G151 and G160 (HIV G140 and G149) are also present in M-PMV integrase [43].

Although no three-dimensional structures of integrases with bound nucleic acids are presently available,

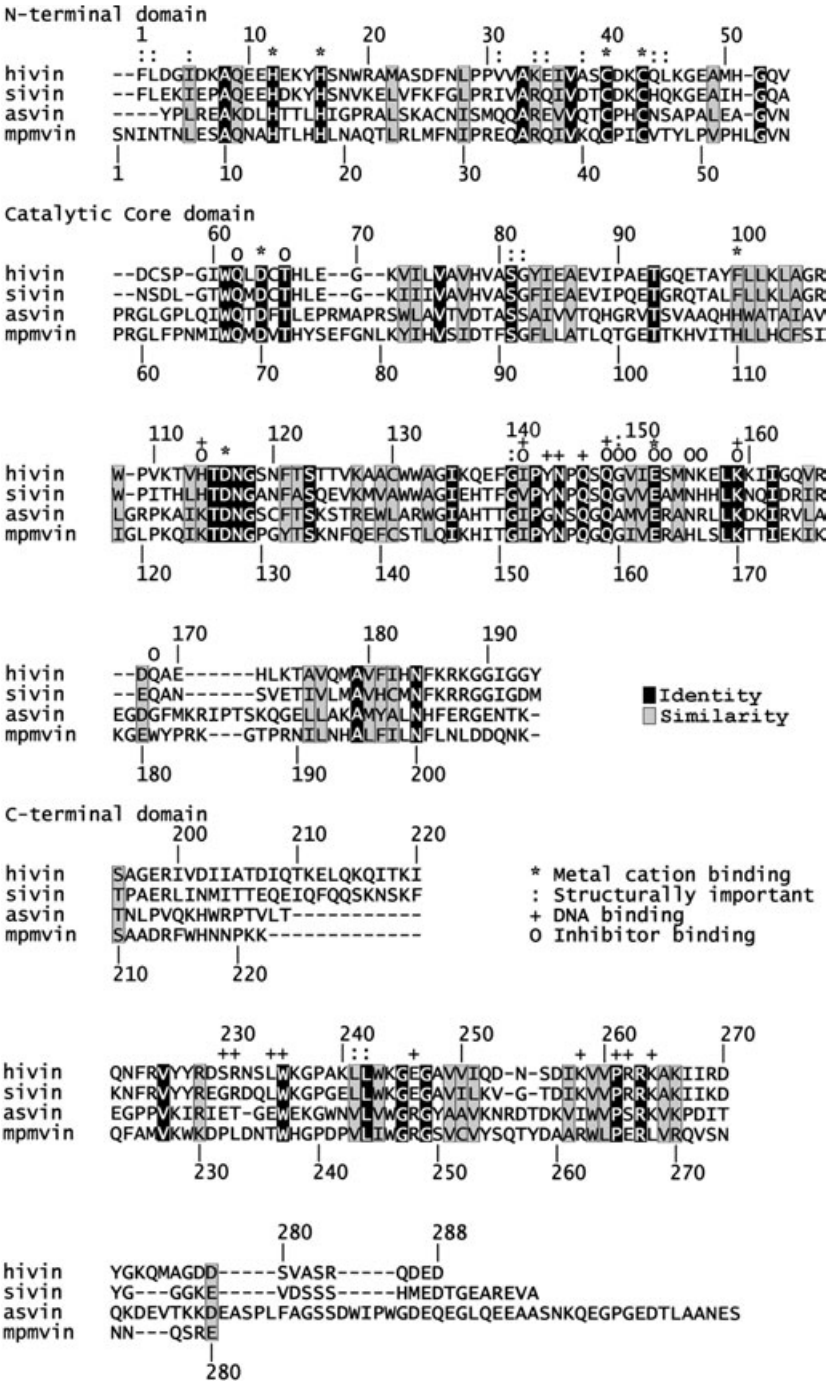

Fig. 8. Structure based alignment of HIV, ASV, and SIV integrases, with M-PMV integrase aligned based upon its primary structure. Identical amino acid residues conserved across all four proteins are marked in black, while similar residues are marked in grey. *, residues which bind metal cations; :, residues found to be important in maintaining protein three dimensional structure and stability; +, residues which may bind DNA; O, residues which bind inhibitors.

DNA crosslinking studies have implicated certain positively charged or hydrophobic residues to be involved in nucleic acid binding. The residues that might play this role in M-PMV integrase are K125, Y154, and K170 (HIV H114, Y143, and K159) [44]. The DNA-binding C-terminal domain contains less well conserved residues, R248, R262, P265, E266, L268, and perhaps P232, L233 (HIV E246, K258, P261, R262, K264, perhaps S230, R231) [45]. This lower degree of identity may reflect a difference in specificity, a lower stringency in the residue identities needed to hold DNA in this region, or simply a difficulty in aligning the sequences in this region.

Finally, several integrase structures have been solved with bound inhibitors. Leaving aside the purely computationally derived models, which have not

Table 1. Comparison of the homology of M-PMV integrase and other selected integrases (HIV, SIV, and ASV). Percentages of identity and similarity are based on structure-based alignment as shown in Fig. 8.

	Full	N-term	Core	C-term
All				
Identity	13	12	15	12
Similarity	18	12	19	21
Ident. + Sim.	31	25	33	33
HIV				
Identity	27	26	30	21
Similarity	19	14	19	24
Ident. + Sim.	46	40	49	45
SIV				
Identity	27	32	28	21
Similarity	21	16	21	29
Ident. + Sim.	48	47	49	50
ASV				
Identity	30	26	33	26
Similarity	21	19	23	19
Ident. + Sim	52	46	56	45

necessarily agreed with the solved structures, we analyzed integrase proteins with bound inhibitors 5-Cl-TEP (with HIV) [46], DHPTPB 3,4-dihydroxyphenyltriphenylphosphonium bromide [47], and Y-3, an anti-HIV integrase inhibitor which also inhibits ASV integrase and was only solved bound to ASV integrase [48]. Most, but not all, HIV integrase amino acid contacts for 5-Cl-TEP were retained in M-PMV: T72, Q149, E162, H166, L167, and K170 (HIV T66, Q148, E152, N155, K156, and K159). DHPTPB appeared to inhibit integrase activity by binding to the dimer interface at HIV integrase Q168, corresponding to W180 in M-PMV integrase, which leads us to predict that this may not be a cross-species specific inhibitor like Y-3. The Y-3 inhibitor contacts are conserved slightly better in M-PMV, including residues Q68, K125, I152, G160, I161, R164 (vs. ASV Q62, K119, I146, A154, M155, R158) than in HIV (Q62, H114, I141, G149, V150, S153). A study of M-PMV integrase activity inhibition (or three-dimensional structure solution) using 5-Cl-TEP, DHPTPB, or Y-3 might indicate which of these residues are critical for inhibitor binding, aiding future antiviral drug development and design.

Discussion

Knowledge of the formation of the preintegration complex is generally limited, and for simple retroviruses such as M-PMV, this process is almost unknown. There is also a gap in structural characterization of integrases of these viruses. Here we present characterization of the integrase of M-PMV and its functional and structural comparison with integrases of other retroviruses.

The M-PMV integrase is cleaved out from the Gag-Pro-Pol polyprotein precursor during maturation through the action of the viral protease [5]. The *in vitro* cleavage of the N-terminally extended precursor of integrase with M-PMV protease allowed us to determine the N terminus of integrase. The substrate specificity mapping of M-PMV protease, which was performed with substrates derived from the cleavage sites within the Gag-polyproteins and peptidomimetic inhibitors designed originally for proteases of HIV and ASV, showed that the specificity of M-PMV protease is similar to that of ASV [28]. The cleavage site Tyr-Lys-Ile-Val-Ala*Ser-Asn-Ile-Asn-Tyr between M-PMV reverse transcriptase and integrase, that we determined here, fulfils well the substrate requirements of M-PMV protease. The amino acid alignment of N-termini of four integrases (Fig. 8) indicates the sequence identity of His residues important for metal binding across the compared proteins. In common with other retroviral integrases, M-PMV integrase is a protein with limited solubility in aqueous solutions. Several integrases (HIV, ASV, HTLV-II, CAEV, and MLV) can be purified using the extraction from insoluble inclusion bodies into a buffer with high ionic strength (1 M NaCl). However, this protocol failed for M-PMV integrase when the protein was expressed in transformed *Escherichia coli* cells cultivated at 37 °C. Protein with better solubility was obtained when *E. coli* cells were cultivated for prolonged time at 15 °C. Similar conditions improved the solubility of HTLV I integrase as described by Müller and Kräusslich [10]. We also show that a His-tag attached to the C-terminus does not influence the solubility of M-PMV integrase or its reactions (i.e. 3′-processing, strand transfer and disintegration). Interestingly, Shibagaki *et al.* [13] reported that the presence of an N-terminal His-tag decreased the 3′-end joining activity of FIV integrase and significantly modified the selection of integration sites. They also hypothesized that the His-tag could alter the binding affinity of the protein to DNA. We examined the DNA-binding affinity of both M-PMV integrase and M-PMV integrase (His-tag) by short wavelength UV cross-linking at 254 nm using ds and ss 21-U3 LTR substrates (data not shown). Both integrases exhibited identical binding affinity to ds DNA substrate and did not bind to ss LTR substrate, confirming that ds DNA is a better substrate for the integrase and that the His-tag at the C-terminus of M-PMV integrase does not influence the binding of the enzyme to ds DNA.

M-PMV integrase prefers the substrate derived from M-PMV U3 LTR in both the 3′-processing and the strand transfer reactions. However, the strand transfer activity of M-PMV integrase was very weak using blunt-ended substrates, with the integration products visible only after overexposing the gels during autoradiography. Better covalent joining of the substrates was achieved using precleaved oligonucleotide substrates. Both substrates (U5 and U3) showed distinct integration patterns, confirming that the integration biases are not completely random but rather dependent on the nucleotide sequence and/or the secondary structure of the DNA. Similar results were reported for the visna virus integrase, which cleaved comparably the U5 and U3 substrates but the integration of the U3 substrate yielded a higher number of products [49]. CAEV and MVV integrases demonstrated comparable cleavage activities with both U5 and U3 substrates [50] and HTLV I integrase displayed a significant preference for the U5 LTR substrate in both 3′-processing and strand transfer reactions [10]. Preferential cleavage of the U5 substrate was also reported for HFV [15], FIV [12] and HIV-1 integrases [50].

A number of divalent cations were shown to bind and modulate the activities of retroviral integrases [10,11,18,21,51,52]. The 3′-processing and strand transfer reactions catalyzed by M-PMV integrase are supported by Zn^{2+}, Mn^{2+}, or Mg^{2+} ions. While the extent of both reactions *in vitro* is noticeably higher in the presence of Mn^{2+}, under physiological conditions the integration process is efficient in the presence of magnesium ions. Interestingly, M-PMV integrase efficiently catalyzed the 3′-processing reaction also in the presence of Co^{2+} or Ni^{2+} ions. These divalent cations also supported the cleavage of (homo)oligonucleotide dT_{10} by M-PMV integrase in the reaction considered as a nonspecific alcoholysis. However, strand transfer as well as disintegration activities were detected only in the presence of Mn^{2+}. The presence of divalent metal cations also facilitates formation of multimeric forms of M-PMV and HIV integrases. The chemical crosslinking experiments in the presence of 1-ethyl-3-(3-dimethylamino-propyl)carbodiimide showed that when divalent ions were removed by dialysis of the integrase samples against the buffer containing 10 mM EDTA, both integrases were present in solution only as monomers. However, these enzymes form dimers and higher multimers in the presence of 10 mM Mg^{2+}, 10 mM Mn^{2+}, or 10 mM Zn^{2+} ions (data not shown). Our results are consistent with metal cation induced multimerization of HIV-1 integrase at submicromolar concentrations [53].

M-PMV integrase displayed more relaxed sequence requirements for site-specific cleavage and strand transfer compared to HIV integrase, which efficiently catalyzed the reactions with substrates derived only from HIV LTRs. FIV [12], HTLV II [11], CAEV and MVV integrases [9] also display similar substrate flexibility, recognizing both their cognate and HIV-1 LTR substrates. The sequence requirements for disintegration catalyzed by HIV-1 integrase, as well as with M-PMV integrase, are less stringent, because both integrases retain similar disintegration activity with both the HIV-1 and M-PMV LTR derived substrates.

Comparative analysis of the primary structure of M-PMV integrase involving the other integrases for which the three-dimensional structures are available provides guidance for future experiments aimed at the explanation of functional and structural properties of this enzyme. These data can be used for designing mutagenesis experiments. However, complete understanding of the specificity of this enzyme may not be possible without additional experiments aimed at determination of the crystal structure of at least the isolated domains of M-PMV integrase, and possibly of the complete protein.

Materials and methods

Construction of expression vectors

A plasmid pSARM 15 containing the full-length coding sequences of M-PMV was used for cloning of all expression vectors. The coding regions for predicted M-PMV integrase and 15 adjacent amino acids at the N-terminus were amplified by polymerase chain reaction (PCR) using *Pfu* polymerase (New England Biolabs) and primers 5′-CGGAATTCATATGATTGGACATGTCAGGG-3′, complementary to the N-terminal sequences of the precursor, and 5′-CCCTCGAGTCACTCCCTGGATTGG-3′, complementary to sequences preceding the stop codon of the *pol* reading frame, respectively. The primers introduced *Nde*I (*Eco*RI) and *Xho*I restriction sites (underlined) at the 5′- and 3′- ends of the encoded DNA sequence.

Primer sequences used for the amplification of DNA encoding wt M-PMV integrase were as follows: 5′-CGGAATTCATATGAGTAACATAAACACA-3′ and 5′-CCCTCGAGTCACTCCCTGGATTGG-3′. The oligonucleotides used for the amplification of M-PMV integrase coding region with a C-terminal His-tag were 5′-CGGAATTCATATGAGTAACATAAACACA-3′ and 5′-CCCTCGAGTCACTCCCTGGATTGG-3′. The PCR-amplified DNAs, digested with *Nde*I and *Xho*I, were isolated from the gel and ligated into pET22b (Novagen) to generate expression vectors pET22bprecursor, pET22bM-PMVin and pET22bM-PMVinhistag. Cloning procedures were

performed using established techniques [54]. The clones were characterized by restriction analysis and verified by DNA sequencing.

Expression of M-PMV integrase in *E. coli* and protein purification

Purification of M-PMV integrase containing C-terminal His-tag (integrase His-tag)

E. coli BL21(DE3) cells were transformed with the pET22M-PMVinhistag plasmid. A single colony was used to inoculate 10 mL of Luria–Bertani (LB) medium containing ampicillin (final concentration of 100 $\mu g \cdot mL^{-1}$) and grown for 10 h at 37 °C. The culture was then diluted 1 : 400 with fresh LB medium with ampicillin, incubated at 18 °C to a D_{600} of 0.5, and the expression was induced by addition of isopropyl thio-β-D-galactoside to a final concentration of 0.4 mM. The cells were harvested after 20 h of further cultivation at 18 °C by centrifugation at 5000 ***g*** for 20 min in a Beckman JA-14 rotor. The cells were solubilized in 20 mM Tris, pH 8.0, 0.1 mM EDTA, 2 mM 2-mercaptoethanol, 500 mM NaCl (buffer A), lysed by lysozyme ($0.2 mg \cdot mL^{-1}$) for 30 min, sonicated for 4× 25 s on ice, and centrifuged at 10 000 ***g*** for 20 min in a Beckman JA-18 rotor. The pellet was resuspended in 20 mM Tris, pH 8.0, 2 M NaCl, 2 mM 2-mercaptoethanol (TNM buffer), stirred for 24 h at 4 °C, and centrifuged again at 10 000 ***g*** for 45 min in the Beckman JA-18 rotor. The supernatant was filtered through a 0.45 μm filter and loaded onto a Ni-nitrilotriacetic acid column (Qiagen). The impurities were removed from the column by extensive washing with 50 mM imidazole in TNM buffer. The integrase protein was eluted with a gradient of 50–500 mM imidazole in TNM buffer. The peak fractions were collected and dialyzed against 20 mM Hepes, pH 7.5, 1 mM EDTA, 1 mM dithiothreitol, 1 M NaCl, and 20% glycerol. The concentration of protein was determined by Bradford method [29].

Purification of M-PMV integrase

Cells BL21(DE3) carrying plasmid pET22BM-PMVin were grown, induced, and harvested under conditions similar to those used for the production of integrase with His-tag. The pellet from the lysed cells was resuspended in 100 mL of 20 mM Hepes, pH 7.5, 1 mM EDTA, 1 mM dithiothreitol (HED buffer) containing 100 mM NaCl. Following the second centrifugation step (10 000 ***g***, 45 min), the integrase was repeatedly extracted from the pellet with 100 mL of HED buffer with 1 M NaCl and stirred for several hours on ice. The protein was precipitated from the collected supernatants by the addition of ammonium sulfate to 30% saturation. The protein was resuspended in HED buffer containing 1 M NaCl by stirring for 30 min at 4 °C and dialyzed against HED buffer containing 800 mM ammonium sulfate and 200 mM NaCl (buffer B). The suspension was clarified by centrifugation (10 000 ***g***, 30 min, Beckman JA-18 rotor) and loaded onto a butyl-Sepharose 4B column equilibrated with buffer B. Integrase was eluted from the column with a linear gradient of buffer B and HED buffer containing 300 mM ammonium sulfate, 80 mM NaCl and 10% glycerol. The purification of integrase was monitored by SDS/PAGE. Fractions containing integrase were diluted with two volumes of HED containing 10% glycerol and were loaded onto a heparin-Sepharose column equilibrated with HED with 200 mM NaCl and 10% glycerol (buffer C). Integrase was eluted from the column by a linear gradient of NaCl (from 200 mM to 1 M) in buffer C. Fractions containing integrase were pooled and dialyzed against 20 mM Tris, pH 7.5, 2 mM dithiothreitol, 1 mM EDTA, 10% glycerol (TDEG buffer) with 200 mM NaCl and then loaded onto a phosphocellulose column equilibrated with the same buffer. The integrase was eluted from the column by 1 M NaCl in TDEG buffer. The peak fractions were collected, dialyzed against HED with 1 M NaCl and 20% glycerol and were aliquoted and stored at −70 °C.

HIV-1 integrase expression and purification

HIV-1 integrase was expressed and purified from bacterial cells as described previously [55].

Purification and cleavage of M-PMV integrase precursor

BL21(DE3) cells transformed with the plasmid pET22B carrying DNA sequences corresponding to predicted M-PMV integrase extended at the N-terminus with sequences encoding 15 amino acids from the reverse transcriptase gene were grown, induced, and harvested under conditions similar to those used for the production of integrase with His-tag.

The precursor was expressed into inclusion bodies. For purification, the inclusion bodies were washed several times with 20 mM Tris, pH 8.0, 0.1 mM EDTA, 2 mM 2-mercaptoethanol, 500 mM NaCl, and 2 M urea. The integrase precursor was solubilized in 20 mM Tris, pH 8.0, 1 mM EDTA, 2 mM dithiothreitol buffer with 8 M urea and then step-wise dialyzed against 20 mM Tris, pH 6 with 0.3 M NaCl, 60 mM 2-mercaptoethanol (buffer B). The precursor was cleaved with the 13 kDa form of M-PMV protease (protease) [28]. The standard reaction mixture contained integrase precursor at a final concentration 50 μM in a total volume 50 μL and protease at a final concentration of 5 μM. The digestion mixture was incubated at 37 °C overnight. The cleavage products were separated on SDS/PAGE, transferred onto PVDF membrane, and the N-terminal amino acid sequence was determined for a protein using automated Edman degradation on an Applied Biosystems Procise sequencer.

Oligonucleotide substrates

M-PMV integrase oligonucleotide substrates

The oligonucleotides derived from the M-PMV LTR with the following sequences were used for activity tests. M-PMV U5 LTR (+) strand: 5′-GATCCCGCGGGTCGGG ACA(GT)-3′, (–) strand: 5′-ACTGTCCCGACCCGCGGG ATC-3′; M-PMV U3 LTR (+) strand: 5′-GGCAGCACG GCTCCGGACA(TG)-3′, (–) strand: 5′-CATGTCCGGAG CCGTGCTGCC-3′. The disintegration substrate is composed of four oligonucleotides (ON1–4): ON1: 5′-GA AAGCGACCGCGCC-3′; ON2: 5′-GGACGCCATAGCCC CGGCGCGCGGTCGCTTTC-3′; ON3: 5′-CATGTCCGG AGCCGTGCTGCC-3′; ON4: 5′-GGCAGCACGGCTCCG GACAGGGGCTATGGCGTCC-3′.

HIV-1 integrase oligonucleotide substrates

The following HIV-1 substrates were used. HIV U5 LTR (+) strand: 5′-ATGTGGAAAATCTCTAGCA(GT)-3′, (–) strand 5′-ACTGCTAGAGATTTTCCACAT-3′.

Non-viral oligonucleotide substrates for nonspecific activity of integrases

The following substrates were used: 5′-GTCGTCACTGG GAAAACCCTGGCG-3′, 5′-CAGCAGTGACCCTTTTGC GACCGC-3′. Synthetic oligonucleotides were purified on 15% denaturing polyacrylamide gel by electrophoresis. The separated bands were detected by UV shadowing. Oligonucleotides were extracted from the gel, loaded onto a DEAE-Sephacel column, eluted by 1 M LiCl, and concentrated with a Speedvac (Savant) centrifuge. Oligonucleotides were desalted by passing through a gel filtration NAP-10 column containing Sephadex G-15 (Sigma) and concentrated again in the Speedvac. Radiolabeling of the oligonucleotides (14 pmol) was performed using 30 pmol of [^{32}P]ATP[γP] (6000 Ci·mmol^{-1}, ICN) and 10 U of T4 polynucleotide kinase (PNK, New England Biolabs). The mixture of a total volume of 16 μL was incubated for 45 min at 37 °C and the reaction was stopped by adding 1 μL of 0.25 M EDTA, pH 8.0 and heating at 85 °C for 15 min. Non-incorporated radioactive nucleotides were removed from the labeled oligonucleotides by centrifugation through a spin column containing Sephadex G-15 (Sigma). The labeled oligonucleotide was annealed with its complement by heating at 90 °C followed by slow cooling to room temperature.

To detect the disintegration activity of M-PMV integrase, the ON1 was labeled with PNK and [^{32}P]ATP[γP]. After inactivation of PNK, the equimolar amounts of ON2, ON3 and ON4 were added to the mixture containing 80 mM NaCl. The annealing of the complementary strands was performed by heating the mixture to 85 °C followed by slow cooling to room temperature. The resulting labeled substrate was separated from the unincorporated [^{32}P]ATP[γP] by passing through a spin-column (Sigma).

M-PMV integrase assays

For activity measurements, purified M-PMV integrase at a concentration of 100–220 nM was incubated with the appropriate 5′-end ^{32}P-labeled linear oligonucleotide substrate at concentrations ranging from 3 to 200 nM in a reaction buffer (20 mM Mops, pH 7.2, 50 mM NaCl, 50 μM EDTA, 10 mM 2-mercaptoethanol, 10% glycerol (w/v), 7.5 mM $MnCl_2$, 0.1 mg·mL^{-1} (BSA) at 37 °C for 1–60 min. The final reaction volume was 20 μL. The reaction was stopped by addition of an equal volume of Maxam–Gilbert loading buffer [98% (v/v) deionized formamide, 10 mM EDTA, 0.025% (w/v) xylene cyanol, and 0.025% (w/v) bromophenol blue). Samples were heated at 100 °C for 3 min and the aliquots (5 μL) were resolved by electrophoresis on a denaturing 15% polyacrylamide gel (7 M urea, 0.09 M Tris borate, pH 8.3, 2 mM EDTA, and 15% acrylamide). Gels were dried and subjected to autoradiography or analyzed using a Molecular Dynamics Phosphor Imager.

HIV-1 integrase assays

For activity measurements, purified HIV-1 integrase at a final concentration of 220 nM, determined by the Bradford method, was incubated with the 5′-end ^{32}P-labeled linear oligonucleotide substrate at concentrations ranging from 20 to 80 nM in a reaction buffer at 37 °C for 15 min. Separation of reaction products and visualization were performed similarly to M-PMV integrase activity assays.

Acknowledgements

This work was supported in part by the Grant Agency of the Academy of Sciences of the Czech Republic under Contract No. IAA 40 55 304 and Research project Z4055905.

References

1 Chopra HC & Mason MM (1970) A new virus in a spontaneous mammary tumor of a rhesus monkey. *Cancer Res* **30**, 2081–2086.
2 Fine D, Landon JC, Pienta RJ, Kubicek MT, Valerio MJ, Loeb WF & Chopra HC (1975) Responses of infant rhesus monkeys to inoculation with Mason–Pfizer monkey virus materials. *J Natl Cancer Inst* **54**, 651–658.
3 Daniel MD, King NW, Letwin NL, Hunt RD, Sehgal PK & Desroisiers RC (1984) A new type D retrovirus isolated from macaques with an immunodeficiency syndrome. *Science* **223**, 602–605.

4 Marx PA, Maul DH, Osborn KG, Lerche NW, Moody P, Lovenstine LJ, Henrickson RV, Arthur LO, Gilden RV, Gravell M, et al. (1984) Simian AIDS: isolation of a type D retrovirus and transmission of the disease. *Science* **223**, 1083–1086.
5 Sonigo P, Barke RC, Hunter E & Wain-Hobson S (1986) Nucleotide sequence of Mason-Pfizer monkey virus: An immunosuppressive D-type retrovirus. *Cell* **45**, 375–385.
6 Vora AC, Fitzgerald ML & Grandgenett DP (1990) Removal of 3′-OH-terminal nucleotides from blunt-ended long terminal repeat termini by the avian retrovirus integration protein. *J Virol* **64**, 5656–5659.
7 Vincent KA, Ellison V, Chow SA & Brown PO (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.
8 van Gent DC, Elgersma Y, Bolk MWJ, Vink C & Plasterk RH (1991) DNA binding properties of the integrase proteins of human immunodeficiency viruses types 1 and 2. *Nucleic Acids Res* **19**, 3821–3827.
9 Störmann KD, Schlecht MC & Pfaff E (1995) Comparative studies of bacterially expressed integrase proteins of caprine arthritis-encephalitis virus, maedi-visna virus and human immunodeficiency virus type 1. *J General Virol* **76**, 1651–1663.
10 Müller B & Kräusslich H-G (1999) Characterization of human T-cell leukemia virus type I integrase expressed in *Escherichia coli*. *Eur J Biochem* **259**, 79–87.
11 Balakrishnan M, Zastrow D & Jonsson CB (1996) Catalytic activities of the human T_cell leukemia virus type II integrase. *Virology* **219**, 77–86.
12 Vink C, van der Linden KH & Plasterk RHA (1994) Activities of the feline immunodeficiency virus integrase protein produced in *Escherichia coli*. *J Virol* **68**, 1468–1474.
13 Shibagaki Y, Holmes ML, Appa RS & Chow SA (1997) Characterization of feline immunodeficiency virus integrase and analysis of functional domains. *Virology* **230**, 1–10.
14 Andrake MD & Skalka AM (1996) Retroviral integrase, putting the pieces together. *J Biol Chem* **271**, 19633–19636.
15 Pahl A & Flügel RM (1995) Characterization of the human spuma retrovirus integrase by site-directed mutagenesis, by complementation analysis, and by swapping the zinc finger domain of HIV-1. *J Biol Chem* **270**, 2957–2966.
16 Anthony NJ (2004) HIV-1 integrase: a target for new AIDS chemotherapeutics. *Curr Top Med Chem* **4**, 979–990.
17 Engelman A (2003) The roles of cellular factors in retroviral integration. *Curr Top Microbiol Immunol* **281**, 209–238.
18 Wlodawer A (1999) Crystal structures of catalytic core domains of retroviral integrases and role of divalent cations in enzymatic activity. *Adv Virus Res* **52**, 335–350.
19 Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie R & Davies DR (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **266**, 1981–1986.
20 Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R & Davies DR (1998) Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. *Proc Natl Acad Sci USA* **95**, 9150–9154.
21 Bujacz G, Jaskolski M, Alexandratos J, Wlodawer A, Merkel G, Katz RA & Skalka AM (1995) High-resolution structure of the catalytic domain of avian sarcoma virus integrase. *J Mol Biol* **253**, 333–346.
22 Maignan S, Guilloteau JP, Zhou-Liu Q, Clement-Mella C & Mikol V (1998) Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases. *J Mol Biol* **282**, 359–368.
23 Yang Z-N, Mueser TC, Bushman FD & Hyde CC (2000) Crystal structure of an active two-domain derivative of Rous sarcoma virus integrase. *J Mol Biol* **296**, 535–548.
24 Wang JY, Ling H, Yang W & Craigie R (2001) Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J* **20**, 7333–7343.
25 Cai M, Zheng R, Caffrey M, Craigie R, Clore GM & Gronenborn AM (1997) Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *Nature Struct Biol* **4**, 567–577.
26 Eijkelenboom AP, van den Ent FM, Wechselberger R, Plasterk RH, Kaptein R & Boelens R (2000) Refined solution structure of the dimeric N-terminal HHCC domain of HIV-2 integrase. *J Biomol NMR* **18**, 119–128.
27 Eijkelenboom AP, Lutzke RA, Boelens R, Plasterk RH, Kaptein R & Hard K (1995) The DNA-binding domain of HIV-1 integrase has an SH3-like fold. *Nature Struct Biol* **2**, 807–810.
28 Zábranský A, Andreánsky M, Heidengsfeldová-Hrušková O, Havlíček V, Hunter E, Ruml T & Pichová I (1998) Three active forms of aspartic proteinase from Mason-Pfizer monkey virus. *Virology* **245**, 250–256.
29 Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.
30 Chow SA, Vincent KA, Ellison V & Brown PO (1992) Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science* **255**, 723–726.
31 Jonsson CB, Donzella GA & Roth MJ (1993) Characterization of the forward and reverse integration reactions of the Moloney murine leukemia virus integrase

protein purified from *Escherichia coli*. *J Biol Chem* **268**, 1462–1469.
32 Lubkowski J, Yang F, Alexandratos J, Wlodawer A, Zhao H, Burke TR Jr, Neamati N, Pommier Y, Merkel G & Skalka AM (1998) Structure of the catalytic domain of avian sarcoma virus integrase with a bound HIV-1 integrase-targeted inhibitor. *Proc Natl Acad Sci USA* **95**, 4831–4836.
33 Katzman M & Sudol M (1996) Nonspecific alcoholysis, a novel endonuclease activity of human immunodeficiency virus type 1 and other retroviral integrases. *J Virol* **70**, 2598–2604.
34 Katzman M, Sudol M, Pufnock JS, Zeto S & Skinner LM (2000) Mapping target site selection for the nonspecific nuclease activities of retroviral integrase. *Virus Res* **66**, 87–100.
35 Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.
36 Chen JC, Krucinski J, Miercke LJW, Finer-Moore JS, Tang AH, Leavitt AD & Stroud RM (2000) Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc Natl Acad Sci USA* **97**, 8233–8238.
37 Chen Z, Yan Y, Munshi S, Li Y, Zruygay-Murphy J, Xu BM, Witmer Felock P, Wolfe A, Sardana V, Emini EA *et al.* (2000) X-Ray structure of simian immunodeficiency virus integrase containing the core and C-terminal domain (Residues 50–293) – an initial glance of the viral DNA binding platform. *J Mol Biol* **296**, 521–533.
38 Cohen GE (1997) ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J Appl Crystallogr* **30**, 1160–1161.
39 Asante-Appiah E & Skalka AM (1999) HIV-1 integrase: structural organization, conformational changes, and catalysis. *Adv Virus Res* **52**, 351–369.
40 Bujacz G, Jaskólski M, Alexandratos J, Wlodawer A, Merkel G, Katz RA & Skalka AM (1996) The catalytic domain of avian sarcoma virus integrase: conformation of the active-site residues in the presence of divalent cations. *Structure* **4**, 89–96.
41 Bujacz G, Alexandratos J, Wlodawer A, Merkel G, Andrake M, Katz RA & Skalka AM (1997) Binding of different divalent cations to the active site of avian sarcoma virus integrase and their effects on enzymatic activity. *J Biol Chem* **272**, 18161–18168.
42 Jenkins TM, Esposito D, Engelman A & Craigie. R (1997) Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking. *EMBO J* **16**, 6849–6859.
43 Greenwald J, Le V, Butler SL, Bushman FD & Choe S (1999) The mobility of an HIV-1 integrase active site loop is correlated with catalytic activity. *Biochemistry* **38**, 8892–8898.
44 Heuer TS & Brown PO (1998) Photo-cross-linking studies suggest a model for the architecture of an active human immunodeficiency virus type 1 integrase–DNA complex. *Biochemistry* **37**, 6667–6678.
45 Gao K, Butler SL & Bushman F (2001) Human immunodeficiency virus type 1 integrase: arrangement of protein domains in active cDNA complexes. *EMBO J* **20**, 3565–3576.
46 Goldgur Y, Craigie R, Cohen GH, Fujiwara T, Yoshinaga T, Fujishita T, Sugimoto H, Endo T, Murai H & Davies DR (1999) Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design. *Proc Natl Acad Sci USA* **96**, 13040–13043.
47 Molteni V, Greenwald J, Rhodes D, Hwang Y, Kwiatkowski W, Bushman FD, Siegel JS & Choe S (2001) Identification of a small-molecule binding site at the dimer interface of the HIV integrase catalytic domain. *Acta Crystallogr Sect D* **57**, 536–544.
48 Lubkowski J, Yang F, Alexandratos J, Merkel G, Katz RA, Gravuer K, Skalka AM & Wlodawer A (1998) Structural basis for inactivating mutations and pH-dependent activity of avian sarcoma virus integrase. *J Biol Chem* **273**, 32685–32689.
49 Katzman M & Sudol M (1994) *In vitro* activities of purified visna virus integrase. *J Virol* **68**, 3558–3569.
50 Bushman FD & Craigie R (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.
51 Zheng R, Jenkins TM & Craigie R (1996) Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc Natl Acad Sci USA* **93**, 13659–13664.
52 Hindmarsh P & Leis J (1999) Retroviral DNA integration. *Microbiol Mol Biol Rev* **63**, 836–843.
53 Deprez E, Tauc P, Leh H, Mouscadet JF, Auclair C & Brochon JC (2000) Oligomeric states of the HIV-1 integrase as measured by time-resolved fluorescence anisotropy. *Biochemistry* **39**, 9275–9284.
54 Sambrook J, Fritsch E, F & Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring. Harbor Laboratory Press, Cold Spring Harbor, New York.
55 Snášel J, Rejman D, Liboska R, Točík Z, Ruml T, Rosenberg I & Pichová I (2001) Inhibition of HIV-1 integrase by modified oligonucleotides derived from U5′ LTR. *Eur J Biochem* **268**, 980–986.