INSIGHT INTO THE INTEGRASE-DNA RECOGNITION MECHANISM:
A SPECIFIC DNA-BINDING MODE REVEALED BY AN
ENZYMATICALLY LABELED INTEGRASE*

Olivier Delelis†, Kevin Carayon†, Elvire Guiot, Hervé Leh, Patrick Tauc, Jean-Claude Brochon,
Jean-François Moussadet and Eric Deprez
LBPA, CNRS, Ecole Normale Superieure Cachan, Institut d’Alembert, 61 av. du Président Wilson, 94235
Cachan, France.

Running head: Heterogeneity and catalytic properties of IN-DNA complexes
†O.D. and K.C. contributed equally to this work
address correspondence to: E. Deprez, LBPA, ENS-Cachan. Tel. 33-147-40-23-94; Fax. 33-147-40-76-84;
E-mail: deprez@lbpa.ens-cachan.fr

Integration catalyzed by integrase (IN) is a key process in the retrovirus life cycle. Many biochemical or structural HIV-1 IN studies have been severely impeded by its propensity to aggregate. We characterized a retroviral IN (PFV-1, Primate Foamy Virus) that displays a different solubility profile to HIV-1 IN. Using various techniques, including fluorescence correlation spectroscopy, time-resolved fluorescence anisotropy, size-exclusion chromatography, we identified a monomer-dimer equilibrium for the protein alone, with a half-transition concentration of 20-30 µM. We performed specific enzymatic labeling of PFV-1 IN and measured the resonance energy transfer (FRET) between TAMRA-labeled IN and fluorescein-labeled DNA substrates. FRET and fluorescence anisotropy highlight the preferential binding of PFV-1 IN to the 3’-end processing site. Sequence-specific DNA-binding was not observed with HIV-1 IN, suggesting that the intrinsic ability of retroviral INs to bind preferentially to the processing site is highly underestimated in the presence of aggregates. IN is in a dimeric state for 3’-processing on short DNA substrates, whereas IN polymerization, mediated by nonspecific contacts at internal DNA positions, occurs on longer DNAs. Additionally, aggregation, mediated by nonspecific IN-IN interactions, occurs preferentially with short DNAs at high IN:DNA ratios. The presence of either higher-order complex is detrimental for specific activity. Ionic-strength favors catalytically competent over higher-order complexes by selectively disrupting nonspecific IN-IN interactions. This counteracting effect was not observed with polymerization. The synergic effect on the selection of specific/competent complexes, obtained by using short DNA substrates under high-salt conditions, may have important implications for further structural studies in IN-DNA complexes.

Retroviral integrases (IN1) catalyze the integration of viral DNA into the host genome ensuring its perpetuation in the host cell. The integration process requires two catalytic steps. During the first step, titled 3’-processing (3’-P), IN specifically removes two nucleotides from each viral DNA end. IN then transfers both ends into target DNA by a one-step transesterification reaction, resulting in full-site integration. IN alone is competent for the insertion process (1). Retroviral INs consist of three functional domains. The core domain contains the DDE catalytic triad, and is flanked by the N-terminal domain (involved in multimerization) and the C-terminal DNA-binding domain. The integrity of the DDE triad and a metallic cofactor are strictly required for enzymatic activity.

Biochemical and structural studies of HIV-1 IN have been severely impeded because of solubility. No structural data are available for the full-length protein (free or bound to DNA) to date, although X-ray structures of DNA-free single- or double-domains have been successfully solved (2-5). Hence, the sequence-specificity of the 3’-P reaction remains poorly understood and little is known about the mechanism of IN/DNA substrate recognition. Point mutations or the use of detergent may improve the solubility but only to a limited extent (6;7), and may also cause changes to IN properties. For instance, detergent is
detrimental to 3'-P activity in the presence of the physiological cofactor Mg\(^{2+}\) (8), and one soluble mutant was found to be resistant to diketo-acid compounds (9).

Endonucleolytic cleavage depends primarily on the presence of the canonical CA sequence preceding the removed dinucleotide. Other positions are crucial for 3'-P activity under Mg\(^{2+}\) conditions, despite the absence, in vitro, of IN preference/specificity for the cognate U5 or U3 terminal sequence at the DNA-binding level (10-12). Thus, it is generally hypothesized that IN specificity is fully explained at the catalytic level. However, due to poor solubility, it is a difficult task to determine which enzymatic properties of HIV-1 IN actually correspond to intrinsic properties of retroviral INs, and which properties are related to aggregation. During the course of our study on HIV-1 IN, we have observed significant differences relating to solubility between HIV-1 IN and another retroviral IN, Primate Foamy Virus-1 (PFV-1) IN. PFV-1 is the prototype of Foamy viruses belonging to the retrovirus family and differ from lentiviruses in some aspects. For instance, PFV-1 reverse transcription occurs at late stages of the replication cycle (13;14). Interaction with host cellular cofactors is also distinct, as LEDGF/p75 only interacts with lentiviral INs (15). Furthermore, only the U5 end is processed due to the asymmetric nature of the integration process in the case of PFV-1 (16). However, integration is an obligatory step for replicating PFV-1, a hallmark of retrovirus life cycle (16). PFV-1 IN shares common structural features with other retroviral INs, such as three-domain organization as well as the three major catalytic activities, 3'-P, strand transfer and disintegration (17-19).

To get deeper insight into the IN/DNA recognition mechanism, we studied and compared solubility, oligomeric status and catalytic efficiency of PFV-1 and HIV-1 IN. We used various methods, such as time-resolved fluorescence anisotropy (TFA), fluorescence correlation spectroscopy (FCS) and size-exclusion chromatography, and found that free PFV-1 IN was significantly more soluble than HIV-1 IN. This enabled us to successfully apply transglutaminase (TGase)-mediated TAMRA labeling to a soluble IN at a single predefined position. Fluorescence resonance energy transfer (FRET) experiments and DNA-binding anisotropy-based assay were conducted to assess the specificity of IN/DNA interactions. A specific mode of interaction was clearly shown for PFV-1 IN, but not for HIV-1 IN. DNA promotes the catalytically competent dimeric state. The dimerization process is in competition with other processes (which were found to be detrimental to activity) such as IN polymerization and aggregation, mediated by nonspecific IN-DNA and IN-IN interactions, respectively. A selection of specific/competent complexes may be achieved using short DNA substrates under conditions of high ionic strength.

**EXPERIMENTAL PROCEDURES**

**Purification and IN labeling by guinea pig (gp) TGase-** For DNA-binding and 3'-P assays, unlabeled wild-type HIV-1 and PFV-1 IN were purified as previously described (8). The plasmid encoding PFV-1 IN with the specific C-term PKPQQFM tag for TGase-mediated labeling was constructed by PCR amplification of pET15b-PFV-IN (20) with P1 and P2 primers (P1: 5'-ACA TAT GTG TAA TAC CAA AAA ACC AAA CCT GG-3'. P2: 5'-AGG ATC CTA CAT AAA CTG CTG AGG TTT TGG CTC GAG TTC ATT TTC -3'). The corresponding plasmid for tagged HIV-1 IN was constructed by amplification of pET15b-HIV-IN (8) with P3 and P4 primers (P3: 5'-CCA TAT GTT TTT AGA TGG AAT AGA TAA-3'. P4: 5'-CGG ATC CTA CAT AAA CTG CTG AGG TTT TGG GTC CTC ATC CTG TCT ACT-3'). The underlined nucleotides correspond to the TGase tag – PKPQQFM – located at the C-term of IN. The resulting DNAs encoding IN-PKPQQFM were inserted into pET15b.

For labeling, IN attached to Ni\(^{2+}\)-beads was incubated overnight at 4°C with 0.5 mM TAMRA cadaverine (Molecular Probes) and 1 unit of gp TGase (Sigma) in the labeling buffer (20 mM Hepes pH 7.5, 10 mM CaCl\(_2\)). Beads were extensively washed with Tris buffer (50 mM pH 8, 1 M NaCl) to remove free fluorophores. Elution was done with 1 M imidazole. TAMRA-labeled or unlabeled PFV-1 IN was further purified by cation-exchange chromatography on a Mono S column (BioRad) previously equilibrated with 20 mM Hepes pH 7.5, 0.1 M NaCl. IN was diluted...
before loading (1 mL/min) and a linear salt gradient (0.1-1 M NaCl) was applied at 1 mL/min. IN was eluted at 450-550 mM NaCl (cation-exchange chromatography was unsuccessful with HIV-1 IN for solubility reasons). It was possible to further concentrate PFV-1 IN using Amicon Ultra centrifugal filters (Millipore). Concentrations up to 250 µM in 20 mM Hepes pH 7.5, 75 mM NaCl were obtained. Ionic strength was adjusted by buffer exchange during the dialysis step. Dialysis was not used as protein aggregation was reproducibly observed during the dialysis step of concentrated samples (significant precipitation occurred at PFV-1 IN concentrations of 12 mg/mL, whichever protocol was used).

Size-exclusion chromatography- 100 µL of varying concentrations of purified PFV-1 IN (from 2 to 60 µM) were applied at room temperature onto a Superdex 200 prep grade column (Hiload 16/60, Amersham Pharmacia biotech) using a AKTA FPLC system (Amersham Bioscience) previously equilibrated with the protein buffer (20 mM Tris-HCl pH 8, 0.05 or 0.5 M NaCl, 10% glycerol, 50 µM ZnSO₄ and 4 mM β-mercaptoethanol). The column was calibrated with the following globular proteins: sweet potato β-amylase (200-kDa), yeast alcohol dehydrogenase (150-kDa), bovine serum albumin (66-kDa), ovalbumin (45-kDa), carbonic anhydrase (29-kDa) and horse heart cytochrome c (12.5-kDa) (MW-GF-200 Kit, Sigma). Fractions were collected at a flow rate of 0.3 ml/min, under a pressure of 0.5 mPa. Absorbance was measured at 280 nm. The exclusion volume of the column V₀ was measured by calibration with blue dextran (2000-kDa).

Characterization of TAMRA-labeled IN (INᵀ)- We evaluated the number of TAMRA per labeled protein by mass spectrometry analysis: INᵀ and the unlabeled tagged IN were subjected to SDS-PAGE electrophoresis. After gel staining, the gel slides containing the corresponding proteins were destained in 25 mM NH₄CO₃, 50% acetonitrile. A proteolysis using sequence grade trypsin (Promega) was then performed overnight at 37°C. The resulting peptides were recovered by addition of 50% acetonitrile, 5% trifluoroacetic acid and dried in a speed-vac. They were reconstituted in 5% acetonitrile, 0.1% trifluoroacetic acid and applied on a SEND-ID protein array according to the Ciphergen’s protocol. The mass spectra were obtained by reading the protein chip using a SELDI mass spectrometer. The resulting spectra were compared with the in silico profile obtained with GPMAW software (Lighthouse Data).

The labeling yield of purified INᵀ was estimated by both absorbance spectroscopy and FCS. Given the molar extinction coefficient of TAMRA-cadaverine at 556 nm (62,000 M⁻¹.cm⁻¹) and the results from Bradford protein assay, the labeling yield was estimated to be 45% by absorbance spectroscopy. The FCS autocorrelation curve (Fig. 1E) was also used to estimate the labeling yield, as it is possible to deduce from the inverse autocorrelation amplitude the mean number of labeled molecules in the detection volume. Here, g(0)⁻¹ was equal to 0.025 and the detection volume was 1.27 µm³ (the lateral ω₀ and axial z₀ dimensions are equal to 0.45 and 1.5 µm, respectively) giving an estimated concentration of INᵀ (52 nM). Considering the concentration of total IN (100 nM), the deduced labeling yield (≈ 50%) is in good agreement with that determined by absorbance spectroscopy. This calculation is evidently correct only if a single fluorophore is attached to the protein (SELDI analysis of INᵀ confirmed that a single TAMRA moiety was covalently and specifically attached to the protein C-term end).

INᵀ, as obtained after the labeling procedure, was also tested for 3'-P at the U5 end (21-mer oligonucleotide) using standard assay (8) and was compared with both the untagged and the unlabeled tagged proteins. 3'-P and half transfer products were about 23-27% and 3.5-4.5%, respectively, for three hours incubation (data not shown) and no significant difference was observed between the three samples showing that neither the PKPQQFM tag nor the attached fluorophore influences the catalytic activity.

PFV-1 ODNs for DNA-binding/3'-P anisotropy-based assays and FRET experiments- All ODNs (≤ 65-mer) were purchased from Eurogentec (Liège, Belgium) and further purified by gel-electrophoresis. ODNs of various lengths (15→65-mer) were tested for 3'-P activity, all mimicking the specific (cognate) terminal U5 sequence: PFV-CAAT (sequence of the 65-mer: 5'-GAA CTA CAC TTA TCT TAA ATG ATG TAA CTC CTT AGG ATA ATC CAA AAT TCC ATG
ACA AT-3’. The 3’-dinucleotide cleaved by IN is underlined. The processed strand and its complementary nonprocessed strand were denoted A and B, respectively. Fluorescein (Fl) was attached at the 5’- or 3’-end of strand A for DNA-binding experiments. Fl was attached at the 3’-end of strand A for the activity assay. The PFV-GTAT sequence was identical, except that the four 3’-term bases of strand A were GTAT instead of the canonical sequence CAAT. The specific HIV-1 ODN (U5 sequence) is provided in reference (21).

The three nonspecific (random) sequences used for competition experiments (Fig. 4) were: 5’-ACA TAA TCT AAA ATA ATT GCC-3’ (21-mer), 5’-ACC TAT GCG CCG CTA GAT TCC-3’ (21-mer) and 5’-TCA AGC TAG AAG ATT ATC TCA AGT ACA TAA TCT AAA ATA ATT GCC-3’ (45-mer). Annealing of oligonucleotides was obtained by mixing equimolar amounts of complementary strands in 20 mM Hepes pH 7.5, 100 mM NaCl, heating to 85°C for 5 min, and slow cooling to 25°C.

For FRET experiments, ODNs of various lengths (21 → 300-mer) were studied, all mimicking the PFV-1 U5 end. The double-stranded (ds) ODN was Fl-labeled at the 5’-end (A or B strand) (Fig. 3A). Fl-labeled 100- and 300-mer ds ODNs were obtained by PCR. All PCRs were performed using the pHSRV13 plasmid encoding the proviral PFV-1 genome. The 100-mer BFl5 and 300-mer BFl5 were obtained using, as primers, the single-stranded 21-mer BFl5 ODN and a second ODN hybridizing to an internal sequence located, respectively, 100- or 300-bp from the U5 extremity. The 100-mer AF15 and 300-mer AF15 were obtained using, as primers, the non-fluorescent single-stranded 21-mer B ODN and a 5’-Fl-labeled ODN hybridizing to an internal sequence located, respectively, 100- or 300-bp from the U5 extremity. After amplification, PCR products were purified on agarose gel using the Qiagen gel extraction kit.

Steady-state fluorescence anisotropy-based assay- Anisotropy-based assay is a quantitative method to determine both DNA-binding and catalytic parameters of IN using the same sample (21). Briefly, IN binding to Fl-labeled DNA increases the steady-state anisotropy value (\(r\)) (measured on a Beacon Instrument (Panvera, Madison, WI)), allowing the calculation of the fraction of DNA sites bound to IN. The DNA-binding step can be recorded at 25°C using Fl-labeled ODNs at the 5’ or 3’-end of either the A or B strand. By shifting the temperature to 37°C, the activity-dependent decrease in the \(r\) value allows quantification of the 3’-P reaction. This occurs only if Fl is initially linked to the AT dinucleotide (the released PFV-1 dinucleotide). The fraction of released dinucleotides (\(F_{\text{dinu}}\)) can be calculated using the relative decrease of \(r\) compared with the initial value obtained at the end of the DNA-binding step (\(r_{\text{t=0}}\)) (real-time assay) (21). No decrease in the \(r\) value was observed with the PFV-GTAT control sequence, in contrast to the PFV-CAAT canonical sequence (data not shown). The activity can be also calculated in fixed-time experiments (3’-P is stopped by SDS):

\[F_{\text{dinu}} = \frac{r_{\text{NP}} - r}{r_{\text{NP}} - r_{\text{dinu}}}\]

where \(r_{\text{NP}}\) and \(r_{\text{dinu}}\) are the anisotropies for pure solutions of nonprocessed ds ODN and dinucleotide, respectively (\(r_{\text{dinu}}\) was measured using the 5’-AT-3’Fl dinucleotide).

The formation of IN/DNA complexes and the subsequent 3’-P catalytic reaction were performed by incubating Fl-labeled ODNs with unlabeled IN in buffer A (20 mM Hepes pH 7.5, 1 mM DTT, 10 mM MgCl2). IN, DNA and NaCl concentrations are explicitly mentioned in the figure legends.

FRET experiments- FRET between Fl-labeled DNA (donor) and TAMRA-labeled IN (acceptor) was monitored with an Eclipse (Varian) spectrofluorimeter. DNA concentration was constant (20 nM) in buffer A+50 mM NaCl. INT was progressively added and the decrease in the steady-state emission intensity of the donor was recorded at 25°C. Excitation wavelength was 500 nm. The donor quenching (\(q_D\)) was measured at 520 nm. \(q_D\) was estimated using:

\[q_D = 1 - \left( \frac{I_D,0}{I_D,[\text{INT}]} \right)\]

where \(I_D,0\) and \(I_D,[\text{INT}]\) represent the donor emission intensity at 520 nm in absence (= DNA alone) and in presence of acceptor (= along the titration), respectively. Excitation and emission slit widths were 10 and 5 nm, respectively. For each DNA/IN\(^\top\) mixture, a corresponding control experiment was done using unlabeled IN to check the signal stability.

Fluorescence Correlation Spectroscopy- FCS measurements were performed under two-photon excitation (840 nm) on a home-built system using
a 80 MHz mode-locked Tsunami Ti:Sapphire laser (pulse 100 fs) pumped by a Millenia solid-state laser (Spectra Physics) and a Nikon TE2000 inverted microscope. Before entering through the epifluorescence port of the microscope, the laser beam was expanded with a two lenses afocal system to fill the back aperture of the objective (Nikon, Plan Apo, 100x, numerical aperture 1.4, oil immersion). The setup was optimized to obtain a diffraction-limited focal spot. Measurements were typically carried out in a 50 µl solution dropped on the coverslip (buffer A+50 mM NaCl). Fluorescence was collected by the same objective, separated from the excitation by a dichroic mirror (Chroma 700DCSPXR) and focused onto an avalanche photodiode (SPCM-AQR-14, Perkin Elmer). An additional filter was used to reject the residual excitation light (Chroma 2P-Emitter E700SP). The detector was connected to a digital correlator (ALV 6000) that calculates the normalized correlation function $g(\tau)$ of fluorescence fluctuations according to:

$$g(\tau) = \frac{\langle I(t) \cdot I(t+\tau) \rangle}{\langle I(t) \rangle^2}$$

where $I(t)$ is the number of detected fluorescence photons per time unit. Assuming a 3-D Gaussian distribution of the excitation intensity, the FCS function for a free Brownian diffusion process is given by:

$$g(\tau) = \frac{1}{N} \cdot \frac{1}{(1 + \frac{\tau}{\tau_D}) \cdot \sqrt{1 + \frac{\omega_0^2}{\omega_0^2} \cdot \frac{\tau}{\tau_D}}}$$

where $N$ is the mean number of fluorophores in the excitation volume and $\tau_D$ is the translational diffusion time. $\omega_0$ and $z_0$ are the lateral and axial dimensions of the excitation volume, respectively. The mean translational diffusion time was determined by fitting autocorrelation curves using a Levenberg–Marquardt nonlinear least-squares fitting algorithm according to the analytical model (Eq. 2). For two-photon excitation, the diffusion coefficient $D$ is then calculated according to: $D = \frac{\omega^2}{8 \tau_D}$. The calibration of the excitation volume was performed using a 10 nM water solution of TAMRA ($D = 2.8 \times 10^{-10}$ m$^2$ s$^{-1}$). The lateral $\omega_0$ and axial $z_0$ dimensions of the excitation volume were 0.45 and 1.5 µm, respectively. The excitation power was adjusted using a neutral density filter. For our setup, we determined that an excitation power of 10 mW was suitable for two-photon excitation of both TAMRA alone and IN$^T$ (fluorescence intensity exhibits a quadratic dependence and $\tau_D$ value is constant – absence of photobleaching – as a function of the incident power below 16 mW). Recording times were typically of 5 min (average of 10 cycles of 30 s).

**Time-resolved Fluorescence Anisotropy**

Fluorescence of Trp or Fl was used for hydrodynamic studies of free or DNA-bound PFV-1 IN, respectively (in buffer A+50 mM NaCl). IN-DNA complexes were obtained using unlabeled PFV-1 IN and Fl-labeled 15- and 21-mer U5 PFV-1 ODN (Fl at the 5’-end of strand A). Correlation times $\theta$ were calculated from the two polarized fluorescence decays $I_\perp(t)$ and $I_\| (t)$, using time-correlated single photon counting (21;22). Briefly, the time scaling was 19.5 ps per channel and 4096 channels were used. The excitation light pulse source was a Ti:sapphire laser (Millennia-pumped Tsunami femtosecond laser, Spectra Physics) (repetition rate, 8 MHz) associated with a 2$^{nd}$ or a 3$^{rd}$ harmonic generator tuned to 490 or 296 nm for Fl or Trp, respectively. The emission monochromator (ARC SpectraPro-150) was set to 530 or 340 nm ($\Delta\lambda=15$ nm) for Fl or Trp, respectively. The two polarized components were collected alternately over a period of 30 s (total count of $I_\perp(t)$: 15 millions, which is a condition compatible with the recovery of correlation times up to $\approx 80$ ns using Trp fluorescence (7)). $I_\perp(t)$ and $I_\| (t)$ were analyzed by the maximum entropy method (23;24) to determine the distributions of both lifetimes $\tau$ and rotational correlation times $\theta$, according to Eq. 3 and 4:

$$I_\perp(t) = \frac{1}{3} \sum_{j=1}^{m} \alpha_i e^{-1/\tau} \left( 1 + 2 \sum_{j=1}^{m} \rho_j e^{-1/\theta} \right)$$

$$I_\| (t) = \frac{1}{3} \sum_{j=1}^{m} \alpha_i e^{-1/\tau} \left( 1 - \sum_{j=1}^{m} \rho_j e^{-1/\theta} \right)$$

with $\sum_{i=1}^{n} \alpha_i = 1$ and $\sum_{j=1}^{m} \rho_j = r_0$

where $\alpha_i$ represents the relative population characterized by the lifetime $\tau_i$ and $\rho_j$ represents the initial anisotropy related to a motion characterized by the rotational correlation time $\theta_j$. 

...
r’0 is the apparent fundamental anisotropy value (typically 0.24-0.26 in the present study, which is below the fundamental anisotropy value – 0.3 – for Trp at λex = 296 nm, suggesting that a fast component in the PFV-1 IN anisotropy decay cannot be resolved by the instrument as previously found for many proteins (25)). The decay of the total fluorescence intensity, I T(t), and the fluorescence anisotropy decay, r(t) are then calculated from both polarized components according to Eq. 5 and 6, respectively:

\[ I_r(t) = I_\parallel(t) + 2G \times I_\perp(t) \] (5)

\[ r(t) = \frac{I_\parallel(t) - G \times I_\perp(t)}{I_r(t)} = \sum_{j=1}^{\infty} e^{-t/\theta_j} \] (6)

where G represents the correction factor for the difference in the monochromator transmission between parallel and perpendicular polarized components.

RESULTS

TFA and FCS reveal distinct solubility properties between HIV-1 and PFV-1 INs. TFA and FCS are complementary methods used to study the hydrodynamic properties of proteins, measuring the rotational and translational diffusion of molecules, respectively (25). Overall rotational diffusion or flexibility are major causes of light depolarization, and anisotropy decay analysis allows the determination of rotational correlation time (θ) distribution. Long correlation time (θlong) represents the overall tumbling motion related to the protein hydrodynamic volume. In FCS, the temporal behavior of fluorescence fluctuations within a small excitation volume, described by the autocorrelation function, allows the determination of translational diffusion times.

PFV-1 IN was first studied by TFA using Trp fluorescence. Distributions of θ are shown for two IN concentrations (Fig. 1A). θlong were approximately 25 and 53 ns for IN concentrations of 2 and 125 µM, respectively. Given a molecular weight of 46.4 kDa for monomeric PFV-1 IN, these values are consistent with a monomeric (Mo) and a dimeric (Di) form, respectively (Table I). θlong increased continuously from 25 to 53 ns as a function of protein concentration with the half-transition state occurring at 20 µM (Fig. 1B). The intermediary θlong values represent average correlation times during the transition and the observed transition reasonably corresponds to a Mo-Di equilibrium. This result was qualitatively confirmed by size-exclusion chromatography (Fig. 1C).

These results significantly differ to those of HIV-1 IN (Table I). Under similar conditions, i.e. absence of detergent, HIV-1 IN was mainly tetrameric below 200 nM, and strongly aggregated above 200 nM (8;26). This critical concentration was only ten-fold higher with detergent (7). Monomeric HIV-1 IN (32 kDa) (characterized by θlong = 19 ns (7)) was obtained only at low submicromolar concentrations under detergent conditions.

The TGase-mediated TAMRA labeling of HIV-1 and PFV-1 INs (INT) was next performed to study their translational diffusion properties by FCS. This enzymatic labeling (27) occurs specifically on a Gln residue of the PKPQQFM peptide substrate (fused to the C-terminal IN end) and leads to IN being labeled at a single position (confirmed by mass spectrometry analysis; Fig. 1D). HIV-1 INT did not display autocorrelation curves that were satisfactory, due to the presence of bright spikes that originate in the presence of aggregates, regardless of the concentration used (data not shown). The number/frequency of bright spikes was slightly lower but remains significant upon addition of detergent, indicating that irreversible aggregates were formed during the purification/labeling procedure. Thus, HIV-1 INT was not further used in the present study. By contrast, PFV-1 INT resulted in satisfactory autocorrelation curves (Fig. 1E) (no bright spike was detected during the acquisition). The fit yielded a diffusion time (τD) of 565 µs at submicromolar concentrations. The FCS acquisition was repeated using a constant concentration of INT and varying the concentration of unlabeled IN up to 250 µM to study τD and its concentration dependence (Fig. 1E, inset). τD significantly increased from 565 to 680 µs in this concentration range (again, no bright spike was detected in the different mixtures of labeled/unlabeled PFV-1 INs). The τD ratio was 1.2, close to the theoretical ratio (1.3) for a Mo-Di equilibrium.
transition. The FCS half-transition concentration (30 µM) was consistent with that observed in TFA.

We have previously shown that the HIV-1 IN DNA-binding step is very slow (kon=0.23 min⁻¹) (28). We suggested that these slow kinetics originate in pre-binding transition from higher order multimeric states (tetramers and aggregates) to smaller molecular species (Mo) whereas the DNA-bound dimeric form correlates with optimal 3’-P activity (21;26;28). Accordingly, Faure et al. found that the Mo•DNA precedes the Di•DNA complexes by analyzing the time-dependent formation of LTR cross-linked species (29). Interestingly, fluorescence anisotropy shows that a rapid equilibrium is reached after mixing PFV-1 IN and DNA, as compared to the much slower formation kinetics of HIV-1 IN-DNA complexes (Fig. 2), confirming TFA and FCS data (lack of detection of large PFV-1 IN molecular species before DNA binding). Therefore, the conversion of higher order multimers to Mo, responsible for the slow DNA-binding step, is a characteristic of HIV-1 IN. PFV-1, which is mainly monomeric at low-micromolar concentrations, is characterized by a rapid DNA-binding step.

Study of PFV-1 IN/DNA interaction by FRET reveals a preferential DNA-binding for the 3’-end processing site- FRET experiments were performed to get deeper insight into the positioning of IN on the DNA substrate. DNA of various sizes (increasing from 21- to 300-mer) were Fl-labeled at the 5’-end, either at the processing end or at the opposite end (Fig. 3A). Fig. 3B shows one example for the quenching donor (qD) plotted against PFV-1 IN³ concentration. Only the quenching donor was measured (at 520 nm), not the acceptor sensitization (at 580 nm), since the acceptor concentration was a varying parameter along the titration. qD increased as [IN³] increased and reached a plateau corresponding to a maximum quenching efficiency qD,max of 78% (the donor emission did not change using equivalent amounts of unlabeled IN). qD,max = 78% is consistent with the theoretical value (at least 75%), assuming a labeling yield of 50% (see Experimental procedures) and a minimal IN/DNA stoichiometry of 2:1 (see below). The qD,max value decreased continuously by decreasing the IN³/unlabeled IN ratio (Fig. 3B), confirming that the decrease in the donor intensity was actually due to resonance energy transfer.

We found that qD was dependent on the DNA length (Fig. 3C): for a given IN concentration, qD was systematically higher for short DNAs. This result was obtained even though the apparent affinity decreases as the DNA size decreases (Table II), demonstrating that the IN positioning on DNA was the main factor influencing the DNA size-dependence of qD. Only positions of IN³ in close proximity to donor (< 60-80 Å for the Fl-TAMRA pair) can significantly influence the energy transfer. The continuous decrease in the qD value obtained with DNA molecules of greater sizes demonstrates that IN may occupy many internal positions in DNA leading to a “pearl chain” structure. Such a polymerization on long DNAs is consistent with the observation of light scattering that occurs for only 100- and 300-mers at IN concentrations above 600 nM (data not shown).

Interestingly, for 45-, 100- and 300-mer DNA, qD was systematically higher - for a given DNA size - when the Fl donor was attached on the terminal processing site - up to two-fold higher for 300-mer (FRET results are summarized in Fig. 3D). No such difference was observed with the 21-mer. Our results indicate a significant preference for IN binding at the processing end. This is consistent with the absence of bias for the short 21-mer substrates where a difference in the quenching donor between the two ends is not expected as the dimension of DNA is comparable with both the overall dimension of IN and the Förster distance.

Anisotropy-based assay confirms a specific DNA-binding mode for PFV-1, but not for HIV-1 IN- We next assessed the specific DNA-binding of PFV-1 IN (suggested by FRET) by competition experiments. Unlabeled PFV-1 IN was incubated with Fl-labeled DNA, 45-mer PFV-1 U5 (Fig. 4A) or 45-mer random sequence (Fig. 4B), in the presence of increasing concentrations of unlabeled DNA (45-mer PFV-1 U5 or random sequence). The decrease in the anisotropy value was then recorded. Using the cognate Fl-labeled DNA, unlabeled cognate and random sequences competed with different IC50 values (respectively 1.4 and >2.5 µM) (Fig. 4A). Consistently, the
competition was easier when the Fl-labeled DNA was random, with corresponding values of 0.5 and 1.6 µM, respectively (Fig. 4B). These results indicate that PFV-1 IN has a higher apparent affinity for the cognate sequence. The IC50 value found for the PFV-1 IN/45-mer U5 complex (1.4 µM at 50 mM NaCl) was used to estimate the apparent Kd value (Kd,app) according to (30;31) (Cheng-Prusoff relationship was not used here because [IN] > [DNA substrate]): The calculated Kd,app from competition assay (0.13 µM) was then compatible with the corresponding Kd,app obtained by direct titration measurement (0.1 µM in Table II), indicating no or only a slight influence of the Fl moiety on the DNA-binding properties of IN.

Preferential/specific DNA-binding of HIV-1 IN is generally difficult to assess in vitro (10;12;32) despite (i) specific and nonspecific complexes are distinct in term of radius of gyration (33), (ii) several residues (Q148, Y143, K156, K159, K14) appear to be involved in such specific contacts before or/and after the 3'-P reaction (11;32;34-36). In particular, it was recently shown that Q148 is also involved in maintaining the complex stability for the subsequent strand transfer reaction by specifically contacting the 5’-LTR end of the donor (viral) DNA (34;35). Competition experiments were repeated with HIV-1 IN and confirm that no significant difference can be observed between the cognate and a random 45-mer DNA sequence, irrespective of the Fl-labeled DNA sequence (Fig. 4C & 4D). Moreover, HIV-1 and PFV-1 INs display different apparent affinities for their respective DNA substrate (compare Fig. 4C & 4D). For 45-mer DNA, the specificity and the absence of specificity for PFV-1 and HIV-1, respectively, were obtained regardless of the ionic strength (50 or 150 mM NaCl). For 21-mer DNA, PFV-1 IN displayed differential DNA-binding between cognate and random sequences at 150 mM NaCl (two different random sequences were tested and gave similar results), whereas no specific DNA-binding was observed at 50 mM NaCl (data not shown). Again, no specificity was observed with HIV-1 IN using 21-mer DNA at either salt concentration (data not shown).

Comparative study of catalytic properties of HIV-1 and PFV-1 INs- Taking into account that solubility and DNA-binding properties of HIV-1 and PFV-1 INs were significantly different, we next compared catalytic features of both enzymes using a fluorescence 3’-P assay (21). Briefly, this assay is based on steady-state anisotropy (r) and allows the separation of DNA-binding and catalytic parameters of 3’-P reaction as well as the study of real-time kinetics. A simultaneous quantitative analysis of DNA-binding and dinucleotide release is possible as both steps strongly influence the molecular size of the fluorescent moiety when Fl is linked to the released dinucleotide.

For HIV-1, we have previously found that the r value, as obtained after DNA-binding and before catalysis (r0), was fully predictive of the subsequent activity at low IN:DNA ratios, according to the fractional saturation function (21). By contrast, for higher ratios, r continued to increase but the 3’-P activity significantly decreased. The variable r is dependent on both the fractional DNA saturation and the molecular size of the complex. Therefore, higher-order IN multimers/aggregates are detrimental to 3’-P activity and account for the characteristic bell-shaped curves when plotting activity versus IN concentration or the r0 parameter. We obtained a similar behavior with PFV-1 IN using the cognate 21-mer PFV-CAAT substrate when the 3’-P activity was plotted against IN concentration (Fig. 5A) or r0 (Fig. 5B). This suggests that, although HIV-1 and PFV-1 INs, free in solution, are characterized by distinct self-association states, they display comparable enzymatic features for the 3’-P catalytic reaction. HIV-1 IN was previously shown to work very slowly as a single-turnover enzyme for 3’-P (rate constant=0.004 min⁻¹) (28). Here, PFV-1 IN possesses similar slow single-turnover kinetics (Fig. 5D), suggesting an intrinsic single-turnover property of retroviral INs which is clearly independent of the initial aggregation state. Interestingly, the 3’-P activity at the top of the bell-shaped curve (Actmax) was significantly better (from 50 to 75%) with greater ionic strength (from 50 to 200 mM NaCl) (Fig. 5A), although more IN was required to obtain optimal activity, because the overall IN/DNA affinity is weakened by ionic strength (Table II). This result suggests that large complexes, responsible for the activity drop off, were more sensitive to ionic strength than catalytically competent complexes.
Furthermore, the 3'-P activity is better at higher ionic strength for a given $r_{t=0}$ value in the decreasing phase (Fig. 5B) confirming that $r_{t=0}$ accounts for both the number of 3'-P-competent complexes and the less active large complexes. The enhanced activity in high-salt concentration at the optimal $r_{t=0}$ value confirms that ionic strength differentially affects the two types of complexes and delays the formation of large complexes.

We obtained bell-shaped curves for all DNA lengths tested (Fig. 5C). However, we found that the optimal activity ($Act_{max}$) was strongly length-dependent and better 3'-P activity was obtained using shorter ODNs (3-fold higher for 15-mer than that of 65-mer) (confirmed by complete kinetics study in Fig. 5D). Again, $Act_{max}$ did not occur at similar IN concentrations. The optimal concentration depends on DNA size, and this is consistent with apparent $K_d$ values showing a better IN affinity for long DNAs (Table II). Interestingly, the activities in the decreasing phase reached a plateau ($Act_{plat}$) (Fig. 5C). The decreasing phase was more pronounced for longer DNA substrates, as evidenced when plotting the $Act_{max}/Act_{plat}$ ratio versus DNA size (Fig. 5E). These results indicate that shortening the DNA substrate improves the number of catalytically competent complexes over the number of less active large complexes. The stimulating effect of ionic strength on $Act_{max}$ (as described in Fig. 5A for the 21-mer) was reproduced with the 15-mer DNA substrate ($Act_{max} = 90\%$ at 200 mM NaCl; data not shown). Surprisingly, no such effect was observed for the longer 45- and 65-mer DNAs (data not shown). This result will be further discussed in the next section.

Fluorescence of Fl was next used for the TFA study of the hydrodynamics of PFV-1 IN-DNA complexes using 15- and 21-mer Fl-labeled ODNs (Table III). At the optimal IN concentration for activity ($Act_{max}$), $\theta_{long}$ values were consistent with dimeric forms of PFV-1 IN bound to short DNAs (53-63 ns). Above this concentration, we detected large complexes ($\theta_{long} > 100$ ns), consistent with the bell-shaped curves shown in Fig. 5A. Therefore, the dimeric form appears to be the most catalytically competent form for 3'-P and, higher order multimers or aggregates, obtained for high IN:DNA ratio, lead to suboptimal activity. These results strongly parallel those previously obtained with HIV-1 (21).

**DISCUSSION**

We used various biophysical techniques to address the problem of IN solubility and found that HIV-1 and PFV-1 INs are characterized by different solubility properties. PFV-1 IN, free in solution, was found to be more soluble than HIV-1 IN. Moreover, fluorescence anisotropy and FRET experiments highlight sequence-specific IN/DNA recognition that is measurable only with PFV-1 IN. However, most of the enzymatic features are similar for both types of INs (slow single turnover kinetics, dimeric form responsible for optimal 3'-P). In addition, although more soluble, PFV-1 IN can lead to higher order multimers or aggregates when bound to its DNA substrate under conditions of high IN:DNA ratio (leading to suboptimal activity), as previously found for HIV-1 IN (21). Finally, DNA size and ionic strength were found to be important parameters that modulate the number of higher-order complexes. In addition, we have demonstrated the feasibility of the enzymatic labeling of a retroviral integrase with a low-molecular weight fluorophore. In contrast to conventional chemical labeling methods, such a strategy leads to a site-specific labeling of the protein of interest, avoiding multiply labeled proteins and heterogeneous samples, and then is particularly well suited for FRET and FCS studies. Regrettably, bright spikes were detected by FCS with HIV-1 INT, suggesting the presence of aggregates, while PFV-1 IN INT resulted in satisfactory autocorrelation curves with no bright spike detected, reinforcing the idea that HIV-1 IN has a higher propensity for aggregation than PFV-1 IN.

In previous studies, HIV-1 IN was found mainly tetrameric, below 200 nM in absence of detergent, and significant protein aggregation was observed above 200 nM (8,26). In the presence of detergent, this critical concentration was found in the low-micromolar range. Here, PFV-1 IN was found mainly monomeric in the low-micromolar range, in the absence of any detergent and DNA. Using PFV-1 IN, either unlabeled (in size-exclusion chromatography and TFA experiments) or TGase-mediated TAMRA-labeled (in FCS...
experiments), we have clearly identified a Mo⇌Di equilibrium for the protein, free in solution. TFA and FCS gave approximately the same half-transition concentration, 20-30 µM. PFV-1 IN is therefore characterized by a Mo⇌Di equilibrium and aggregation does not significantly occur in absence of DNA during the transition. The reason for better solubility than HIV-1 IN remains unclear, as no evident differences appear when comparing the global hydrophobic profiles of both INs (not shown). Mutating only a few surface residues significantly improves HIV-1 IN solubility. For instance, the single mutation F185K improves the solubility of the entire protein (6). It is then possible that only a small number of key residues may account for enhanced PFV-1 IN solubility. Due to the poor sequence similarity between both INs (≈15%), a greater understanding of solubility determinants requires further structural studies.

Interestingly, using the TAMRA-labeled PFV-1 IN and fluorescein-labeled DNA substrates, we found that FRET was systematically more efficient when fluorescein was attached at the processing as compared to the non-processing end. This difference was observed only with long DNA substrates (from 45 to 300-mer) and not with a short 21-mer DNA substrate, with size comparable to the Förster distance. This result highlights a preference of DNA-binding for the processing end and was confirmed by competitive DNA-binding experiments. Altogether, these results suggest that, using PFV-1 IN, it is possible to distinguish between specific and nonspecific complexes, not only at the catalytic level (as typically found for HIV-1 (11;21)) but also at the DNA-binding level. In our opinion, these results do not reveal a differential DNA-binding property of both INs but, most likely, that the aggregative properties and high propensity of HIV-1 IN to establish nonspecific contacts mask in vitro an intrinsic ability of the protein to bind preferentially to its cognate sequence. However, FRET and competition experiments suggest that the specificity, although significant, is modest for PFV-1 IN. This could be due to the inherent nonspecific DNA-binding mode (responsible for the binding to the target DNA) that can be mainly ascribed to the C-term domain, as for the HIV-1 protein (12) although we cannot exclude the possibility that this property is also mediated by the N-term region in the case of PFV-1 IN (18). Furthermore, the catalytic domain which ensures 3’-P and the subsequent joining reaction into a large variety of target DNA sequences should be also in part responsible for nonspecific DNA-binding.

FRET analysis, as a function of DNA size, also suggests a nonspecific DNA binding mode at internal positions onto long DNA substrates. In addition, we found that the 3’-P catalytic process was stimulated by (i) shortening the DNA size and (ii) increasing the ionic strength. Nevertheless, the stimulating effect of ionic strength was not observed for long DNA substrates such as 45- and 65-mer. This result, together with the FRET data, suggests the existence of two types of higher-order complexes. IN polymerization, mediated by nonspecific internal IN-DNA interactions, is the main phenomenon occurring on long DNAs and is strongly detrimental to 3’-P activity. Consistently, the Act_{max}/Act_{plat} ratio (as defined in Fig. 5) is lower for 15 and 21-mer DNAs as no polymerization is expected to occur on short DNAs for evident steric reasons. However, their slight decreasing phases may be assigned to the presence of aggregates, mediated by nonspecific IN-IN interactions at high IN:DNA ratios. Only a differential effect of high salt concentration on nonspecific IN-DNA and IN-IN interactions may explain why ionic strength only stimulates Act_{max} of short DNAs. Ionic strength disrupts more efficiently nonspecific IN-IN interactions than nonspecific IN-DNA interactions. This also explains why the specific interaction between PFV-1 IN and the short 21-mer DNA was only evidenced under high-salt conditions (see above) as aggregates on DNA underestimate detection of specific complexes at low ionic strength. Altogether, our results show that two types of large complexes are present at high IN:DNA ratios and both are detrimental for IN activity. The polymerization process is mainly influenced by the DNA length while aggregation is more sensitive to the ionic strength. Consequently, an additive/synergic positive effect on the IN activity is observed by combining both parameters, i.e. shortening the DNA length and increasing the ionic strength.
FRET analysis indicates that IN on long DNAs can be positioned at internal DNA sites due to its nonspecific DNA-binding mode, decreasing the 3’-P efficiency. The DNA-length dependence of 3’-P activity raises the question of the particularly low 3’-P efficiency which can be expected for long DNAs such as physiological viral DNAs (about 9 and 13 kbp-long for HIV-1 and PFV-1, respectively). To date, the mechanism by which IN is specifically positioned at the processing site in vivo is still unclear. Although IN prefers to bind to its cognate DNA sequence above other sequences, the relative specificity of IN described in this study may not be sufficient to ensure its selective in vivo positioning at the processing ends. It has been previously proposed for HIV-1 that the nucleocapsid protein binds tightly to viral DNA, except at the DNA ends (37). Additionally, this protein stimulates IN activity in vitro (38), and may indirectly assist the correct positioning of IN onto the processing ends. It is important to note that a nucleocapsid, characterised by a typical two canonical cysteine-histidine motifs, is not present in PFV-1, although the PFV-1 C-terminus region of Gag contains three glycine-arginine motifs called ‘GR boxes’ which are supposed to display similar properties (39;40).

The PFV-1 IN/DNA complexes, as obtained under optimal conditions for 3’-P activity (Actmax in Fig. 5), were further studied by time-resolved fluorescence anisotropy. Using short DNA substrates (15- and 21-mer), long rotational correlation times were found to be compatible with dimeric forms (Table III), suggesting that Di is the most catalytically competent form for 3’-P, as was previously found for HIV-1 IN (21;29), while large complexes are characterized by suboptimal 3’-P activity. DNA primarily promotes IN dimerization, as Di are present at concentrations in which free PFV-1 IN is monomeric (low micromolar range). Most likely, a tetramer (Di of Di) is required for integration as shown for HIV-1 when two LTR ends are simultaneously present in the context of the synaptic complex (41) or when IN is bound to a three-way junction DNA substrate mimicking an integration intermediate (42). The ability of the tetrameric species to perform 3’-P is still controversial (29;43) and the present data cannot clearly address the proper 3’-P activity of the Te as the decreasing phase is certainly related to an heterogeneous mixture of large complexes, including tetrameric forms. Recently, a kinetic study of ASV Mn2+-dependent 3’-P suggests that a Di⇌Te equilibrium accounts for the biphasic behavior of the single-turnover 3’-P with Te responsible for the fast phase (43). We found a monophasic behavior for both PFV-1 (this study; Fig. 5D) and HIV-1 (21) Mg2+-bound INs with rate constants compatible with the slow ASV phase. Apparent discrepancies between HIV-1, PFV-1 and ASV may reflect intrinsic or metal-dependent differences for Di⇌Te or/and Te⇌aggregates transitions onto DNA. However, HIV-1 and PFV-1 INs, which display similar monophasic behaviors, are characterized by different single-turnover rate constants (for a given DNA size, i.e. 21-mer DNA substrate at 50 mM NaCl, kHV-1 = 0.004 min⁻¹ (21) and kPFV-1 = 0.0028 min⁻¹ (this study)). Apparent affinities for DNA substrate were also significantly different (Kd,app = 40 and 130 nM for HIV-1 (21) and PFV-1 (Table II), respectively, for 21-mer DNA substrate at 50 mM NaCl). Although most of the enzymatic properties described here are qualitatively similar between HIV-1 and PFV-1 INs, the two enzymes display substantial quantitative differences - PFV-1 IN has lower affinity for the DNA substrate and lower catalytic efficiency as compared to HIV-1 IN -, the reasons for these differences remain to be elucidated. Although the sequence homology between HIV-1 and PFV-1 INs is low, our data suggest that both proteins, which display similar enzymatic properties, at least qualitatively, have structural homology. Considering its solubility characteristics, PFV-1 IN could therefore represent a good model for further structural studies of retroviral INs.

REFERENCES


**FOOTNOTES**

* This work was supported by grants from the TrioH European project (FP6 n°503480), ANR (06-PCVI-0015), CNRS and Institut d’Alembert.

1 The abbreviations used are: FCS, Fluorescence Correlation Spectroscopy. Fl, Fluorescein. FRET, Fluorescence Resonance Energy Transfer. HIV-1, Human Immunodeficiency Virus type-1. IN, integrase. IN³, TAMRA-labeled integrase. LTR, Long Terminal Repeat. ODN, Oligodeoxynucleotide. PFV-1, Primate Foamy Virus type-1. \( r \), steady-state fluorescence anisotropy. TFA, Time-resolved Fluorescence Anisotropy. TGase, Transglutaminase. 3'-P, 3'-processing.

**FIGURE LEGENDS**

**Fig. 1. Rotational and translational dynamics of PFV-1 IN.**

---

**A)** Top, rotational correlation time (\( \theta \)) distribution of IN (grey: 2 \( \mu \)M; white: 125 \( \mu \)M) at 20 °C. Fluorescence anisotropy decays were obtained by monitoring the intrinsic Trp fluorescence. One additional short correlation time (\( \approx 400 \) ps) was found under both concentration conditions (not shown). Bottom, corresponding experimental fluorescence anisotropy decays (grey: 2 \( \mu \)M; black: 125 \( \mu \)M) with line fits (black: 2 \( \mu \)M; white: 125 \( \mu \)M) resulting from maximum-entropy method analysis (see Experimental procedures for details). The lifetime distributions of PFV-1 IN, recovered from intensity decays (not shown), were as follows: \( \tau_1 = 0.28 \pm 0.03 \) ns (29%), \( \tau_2 = 1.09 \pm 0.09 \) ns (28%), \( \tau_3 = 2.6 \pm 0.3 \) ns (20%), \( \tau_4 = 6.65 \pm 0.25 \) ns (23%) for 2 \( \mu \)M and \( \tau_1 = 0.24 \pm 0.01 \) ns (21%), \( \tau_2 = 0.94 \pm 0.1 \) ns (31%), \( \tau_3 = 2.4 \pm 0.2 \) ns (21%), \( \tau_4 = 6.6 \pm 0.1 \) ns (27%) for 125 \( \mu \)M. **B)** Long correlation time as a function of [IN]. **C)** Analysis of PFV-1 IN by size-exclusion chromatography. Experiments were performed at room temperature using a Superdex 200 column. Three concentrations of IN were loaded on the column. Only the monomeric species were detected using 2 \( \mu \)M. An additional peak, corresponding to the dimeric form, appeared for the two highest IN concentrations (20 and 60 \( \mu \)M). No other species were detected. \( V_0 \) and \( V_e \) are the exclusion volume of the column and the elution volume of the protein, respectively. The calibration sample (•) is composed of globular proteins of known MW (see Experimental procedures). Circles indicate the position of elution peaks for various samples of PFV-1 IN, in the presence of either 0.05 (●) or 0.5 (○) M NaCl. The concentration of the loaded sample is explicitly mentioned in the figure. **D)** SELDI analysis of trypsin digestion of the unlabeled tagged (top) and TAMRA-labeled (bottom) PFV-1 IN. The peptides obtained after proteolysis were analyzed using a SEND-ID protein chip. Spectra are magnified in the region of interest, corresponding to the expected mass for the modified C-term peptide (NELEPKPQOQM, MW 1860.16 Da) for the TAMRA-labeled protein. The entire spectra did not show any other differences between the two profiles (data not shown).
TAMRA-labeled IN was characterized by two peaks (MW 1860.2 Da and 1876.9 Da). The first one is consistent with the expected theoretical value for the TAMRA-labeled C-term peptide (i.e. a single fluorophore was found as expected on the C-term specific tag.), and the second one (ΔMW = 16.7 Da) corresponds to the same peptide containing an oxidized methionine. 

**Fig. 2.** Differential DNA-binding kinetics of HIV-1 and PFV-1 INs. Fl-labeled DNA that mimics either the HIV-1 or the PFV-1 U5 LTR extremity (double-stranded 21-mer, 4 nM) was mixed with HIV-1 (☐) or PFV-1 IN (■), respectively, in buffer A+50 mM NaCl. Protein concentrations are indicated in the figure. Steady-state fluorescence anisotropy values (r) were measured at 25°C as described in Experimental procedures. \( \Delta r(t) = r(t) - r_{\text{free DNA}} \).

**Fig. 3.** FRET between DNA substrates and PFV-1 IN. FRET between Fl-labeled DNA substrates (donor) of various lengths and IN\(^T\) (acceptor) was measured as a function of [IN\(^T\)], as described in Experimental procedures. 

A) The different DNA substrates used in the FRET study. The Fl donor was attached at the 5'-end of either the A or B strand, giving X-mer AF5 or BF5, respectively (X = 21, 45, 100 or 300).

B) Quenching donor (qD) between 21-mer AF5 and IN\(^T\) (measured at 520 nm). The plateau value depends on the labeled/unlabeled IN ratio (indicated on the plot). Right, emission spectra for the various mixtures of labeled and unlabeled IN (total concentration, 1 µM) in presence of 20 nM 21-mer AF5.

C) Comparison of qD values between the different DNA substrates. [IN\(^T\)] = 100 (☐), 200 (■) or 300 nM (●).

D) Summary of the differential quenching amplitude between the two DNA ends as a function of the DNA size for 200 nM IN\(^T\). All FRET experiments were performed in 20 mM Hepes pH 7.5 containing 1 mM DTT, 10 mM MgCl\(_2\) and 50 mM NaCl.

**Fig. 4.** Comparison of HIV-1 and PFV-1 IN DNA-binding properties. IN (A & B, PFV-1; C & D, HIV-1) was incubated with 4 nM Fl-labeled 45-mer DNA (A & C, cognate sequence; B & D, random sequence) and varying concentrations of unlabeled 45-mer DNA competitor (●, cognate sequence; ○, random sequence) in buffer A+50 mM NaCl (See Experimental procedures for ODN sequences). Fluorescence anisotropy was measured as described in Experimental procedures. \( \Delta r = r_{\text{complex}} - r_{\text{free DNA}} \). [IN] = 700 nM (PFV-1) or 400 nM (HIV-1).

**Fig. 5.** Influences of DNA size and ionic strength on the PFV-1 IN 3'-P activity. 

A & B) Ionic strength effect. 3'-P activity was determined using the 21-mer DNA substrate as a function of [IN] (A) or the \( \Delta r_{t=0} \) value (B). [NaCl] = 50 (■), 150 (○) or 200 mM (●). \( \Delta r_{t=0} = r_{t=0} - r_{\text{free DNA}} \). \( r_{t=0} \) represents the r value after the DNA-binding step (before the start of the reaction).

C) DNA size effect. The response of the 3'-P activity to IN concentration was studied with: 15-mer (○), 21-mer (■), 45-mer (□) and 65-mer (▲) mimicking the U5 LTR extremity. [NaCl] = 50 mM. Activities reported in the Y-axis (panels A-C) correspond to a 420 min-incubation with 12 nM Fl-labeled DNA at 37°C, and were calculated as described in Experimental procedures. 

D) Time course of 3'-P for 15 → 65-mer (symbols: see panel C). Experimental conditions correspond to optimal conditions in Fig. 5C for each DNA size.

E) \( \text{Act}_{\text{max}}/\text{Act}_{\text{plat}} \) ratio as a function of the DNA length. \( \text{Act}_{\text{max}} \) and \( \text{Act}_{\text{plat}} \) are defined in the text and in Fig. 5C.
### TABLE I: Comparison of long correlation times obtained for PFV-1 and HIV-1 INs

<table>
<thead>
<tr>
<th>IN</th>
<th>Theoretical $\theta_{long}$\textsuperscript{a} monomeric form</th>
<th>Experimental $\theta_{long}$\textsuperscript{a} monomeric form</th>
<th>Experimental $\theta_{long}$\textsuperscript{b} higher-order oligomeric forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFV-1</td>
<td>21.5 ns</td>
<td>22.5 – 30 ns\textsuperscript{b}</td>
<td>47 – 53 ns\textsuperscript{c}</td>
</tr>
<tr>
<td>HIV-1</td>
<td>15 ns</td>
<td>18 – 20 ns\textsuperscript{d}</td>
<td>70 – 90 ns + &gt;100 ns\textsuperscript{e}</td>
</tr>
</tbody>
</table>

\textsuperscript{a}assuming partial specific volume (v) of 0.73 mL/g and average hydration of 0.4 gH\textsubscript{2}O/g\textsubscript{protein} (44;45). $\theta_{long}$ is related to the volume of the rotating unit (V) by $\theta = \eta V/kT = \eta Mv/kT$ where $\eta$ is the viscosity, k the Boltzmann constant, T the temperature (K) and M the molecular weight.

\textsuperscript{b}in the 1-5 µM range.

\textsuperscript{c}monodisperse solution of dimeric species above 100 µM.

\textsuperscript{d}at submicromolar concentrations in the presence of detergent (7).

\textsuperscript{e}mainly tetrameric ($\theta = 70-90$ ns) in the absence of detergent below 200 nM; aggregation ($\theta > 100$ ns) occurs above 200 nM (8;26).

### TABLE II: Apparent K\textsubscript{d} values of PFV-1 IN as a function of DNA size and ionic strength\textsuperscript{a}

<table>
<thead>
<tr>
<th>DNA size, bp</th>
<th>[NaCl], mM</th>
<th>K\textsubscript{d,app}, µM</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>50</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>n.d\textsuperscript{b}</td>
</tr>
<tr>
<td>21</td>
<td>50</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>3.5</td>
</tr>
<tr>
<td>45</td>
<td>50</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>1.6</td>
</tr>
<tr>
<td>65</td>
<td>50</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.97</td>
</tr>
<tr>
<td>100</td>
<td>50</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>n.d\textsuperscript{b}</td>
</tr>
<tr>
<td>300</td>
<td>50</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>n.d\textsuperscript{b}</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Fl-labeled ODNs of various sizes mimicking the cognate U5 DNA end (12 nM) were mixed with increasing concentrations of unlabeled PFV-1 IN in the presence of 10 mM Mg\textsuperscript{2+}. DNA-binding was measured by recording the steady-state anisotropy value as described in Experimental procedures.

\textsuperscript{b}not determined.
Table III: Long correlation times of PFV-1 IN-DNA complexes

<table>
<thead>
<tr>
<th>DNA size, bp</th>
<th>[PFV-1 IN], nM</th>
<th>$\theta_{\text{long}}, \text{ns}$ $^{\text{b}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>1200</td>
<td>63 ± 6</td>
</tr>
<tr>
<td>21</td>
<td>400</td>
<td>53 ± 8</td>
</tr>
<tr>
<td>21</td>
<td>600</td>
<td>61 ± 5</td>
</tr>
<tr>
<td>21</td>
<td>1200</td>
<td>&gt; 100</td>
</tr>
</tbody>
</table>

$^{a}$TFA experiments were performed with 12 nM DNA in buffer A+50 mM NaCl at 20°C as described in Experimental procedures.

$^{b}$The correlation time distributions of free ODNs as a function of the DNA size were previously published (21). $\theta_{\text{long}} = 6$ and 9.5 ns at 20°C for free ODN$_{15\text{-mer}}$ and free ODN$_{21\text{-mer}}$, respectively.
Figure 1

A

B

C

D

E

17
Figure 2
Figure 4