Relationship between the Oligomeric Status of HIV-1 Integrase on DNA and Enzymatic Activity*

Received for publication, March 8, 2006, and in revised form, May 24, 2006 Published, JBC Papers in Press, June 13, 2006, DOI 10.1074/jbc.M602198200

Elvire Guiot†, Kevin Carayon†, Olivier Delelis§, Françoise Simon§, Patrick Tauc‡, Evgenii Zubin§, Marina Gottikh§, Jean-François Mouscadet†, Jean-Claude Brochon†, and Eric Deprez†‡

From the †Laboratoire de Biotechnologie et Pharmacologie Genetique Appliquee, CNRS, UMR8113, Ecole Normale Superieure de Cachan, 61 av du President Wilson, 94235 Cachan, France and the §Belozersky Institute of Physico-Chemical Biology, Moscow State University, 119992 Moscow, Russia

The 3′-processing of the extremities of viral DNA is the first of two reactions catalyzed by HIV-1 integrase (IN). High order IN multimers (tetramers) are required for complete integration, but it remains unclear which oligomer is responsible for the 3′-processing reaction. Moreover, IN tends to aggregate, and it is unknown whether the polymerization or aggregation of this enzyme on DNA is detrimental or beneficial for activity. We have developed a fluorescence assay based on anisotropy for monitoring release of the terminal dinucleotide product in real-time. Because the initial anisotropy value obtained after DNA binding and before catalysis depends on the fractional saturation of DNA sites and the size of IN-DNA complexes, this approach can be used to study the relationship between activity and binding/multimerization parameters in the same assay. By increasing the IN-DNA ratio, we found that the anisotropy increased but the 3′-ends are then used in the second reaction, strand transfer, for the covalent joining of viral and target DNAs, resulting in full-site integration. IN is sufficient for catalysis of the 3′-processing reaction in vitro, using short-length oligodeoxynucleotides (ODNs) that mimic one viral long terminal repeat (LTR) in the presence of the metallic cofactor Mg2+. This reaction generates two products: the viral DNA containing the recessed extremity and the GT dinucleotide. One of the two products, the processed viral DNA, as well as the target DNA serve as substrates for the subsequent joining reaction.

IN belongs to the superfamily of polynucleotidyl trans-ferases. Its catalytic core domain contains a triad of acidic residues constituting the D,D-35-E motif, which is strictly required for catalysis. The catalytic core establishes specific contacts with the viral DNA and, together with the C-terminal domain, is involved in DNA binding (1–4). The 3′-processing reaction is highly specific, and the terminal 13 bp of the LTR play a key role for Mg2+-dependent 3′-processing in terms of reaction specificity (4, 5). In particular, the CA sequence preceding the GT dinucleotide cleaved by IN is strictly required. Reaction specificity seems to depend on the catalytic step, because no significant difference in affinity is observed in vitro for different DNA sequences (2, 3, 5, 6). The catalytic mechanism of IN has been extensively studied, but the structural determinants of IN activity remain unclear, and controversy remains concerning the multimeric status of active IN. Single- and double-domain crystallographic structures have been produced for IN, alone or complexed with the IN-binding domain of the cellular partner lens epithelium-derived growth factor, and all these structures display a conserved dimeric structure of the catalytic core (7–11). Based on topological considerations concerning the distance separating the active sites of the two protomers, which is more than double the distance separating the two joining sites on the target DNA (separated by 5 bp), a multimeric state of an order higher than dimers, tetramers at least, is required for complete integration. However, it remains unclear whether such a higher order multimeric state is also required for 3′-processing activity and how this activity is mod-ulated by the self-association properties of IN. Moreover, most enzymatic studies are carried out with an excess of IN over DNA substrate. Taking into account the low solubility of IN,
we investigated whether high order multimers or aggregated forms of IN favored by high protein concentrations in solution were detrimental or beneficial for activity.

We developed an assay for monitoring DNA binding and subsequent 3′-processing in the same sample. This assay makes possible the separation of binding and catalytic parameters and the study of real-time kinetics. It is based on steady-state fluorescence anisotropy (r), which is sensitive to rotational diffusion, and thus suitable for studies aiming to identify structural modifications leading to a significant change in molecular size (12–15). Using a fluorescent probe covalently linked to the GT dinucleotide, this makes it possible to follow DNA binding and dinucleotide release, because both steps strongly influence molecular size of the fluorescent moiety. The anisotropy-based technique is also highly suitable for studies of the relationship between the overall size of IN-DNA complexes and activity. We found that, for low IN:DNA ratios, the r values obtained after DNA binding and before catalysis were fully predictive of subsequent IN activity, according to the fractional saturation function. For high IN:DNA ratios, anisotropy continued to increase, but 3′-processing activity decreased. Because r depends on both fractional saturation and the molecular size of complexes at saturation of DNA sites, our results show that high order multimers or aggregated states of IN are detrimental to 3′-processing activity. Activity levels were highest for non-aggregative smaller species. A more precise characterization of catalytically competent complexes by time-resolved fluorescence anisotropy (TFA) (16–20), confirmed steady-state data and led to the identification of dimeric forms as the most active forms for 3′-processing. Our results also highlight that only the IN:DNA ratio but not the IN concentration per se determines the aggregation properties and thus the activity of IN and that DNA binding stimulates the self-organization of IN to give a catalytically competent non-aggregative form.

**EXPERIMENTAL PROCEDURES**

Oligonucleotides and Nomenclature—Unlabeled and fluorescein-labeled single-stranded (ss) ODNs were purchased from Eurogentec (Liege, Belgium) (except for the 2′-aminouridine-containing ODN) and purified by electrophoresis in acrylamide gels. The specific HIV sequence was 5′-GTGTTGAAAAATCTCCTAGCAGT-3′ (the bases removed by IN are underlined). This sequence was denoted a, and the complementary non-process strand was denoted b. Fluorescein (F) was attached at the 5′- or the 3′-end of strand a or b, via a 6-carbon linker. As an example, the specific double-stranded (ds) ODN used to monitor activity was called HIV-α3F, indicating that fluorescein was attached to the 3′-end of strand a. HIV-α3F was identical except that the four terminal bases at the 3′-end of strand a were GTGT rather than the canonical sequence CAGT. The nonspecific sequence, NS-TTCC, was 5′-ACCTATGCGCCGCTAGATTCC-3′ (strand a). Two other nonspecific ODNs with different 3′-ends of strand a were derived: NS-CACC and NS-CAGT. ds ODNs were obtained by mixing equimolar amounts of complementary strands in 20 mM Tris-HCl (pH 7.2), 100 mM NaCl. The mixture was heated to 85°C for 5 min, and annealing was allowed by slow cooling to 25°C.

The uridine-containing ODN (corresponding to strand a of the DNA substrate HIV-αUF) contains a fluorescein attached to the 2′-amino group of the 3′-terminal 2′-aminouridine. It was synthesized as follows: the solid support was prepared as previously described (21). Briefly, 2′-deoxy-5′-O-4,4′-dime-thoxytrityl-2′-O-trifluoracetamidouridine (0.2 mmol) was co-evaporated with pyridine (3 × 5 ml) and dissolved in dry pyridine (2 ml). Succinylated long-chain alkylamine-controlled pore glass (500 Å) (200 mg), 2,4,6-trisopropylbenzenesulfonyl chloride (0.6 mmol), and 1-methylimidazole (1.2 mmol) were added, and the mixture was incubated for 3 h at 25°C. The support was then filtered and washed successively with pyridine, CH$_2$Cl$_2$, and ether. Starting from 2′-deoxy-5′-O-4,4′-dimethoxytrityl-2′-O-trifluoracetamidouridine-derivatized long-chain alkylamine-controlled pore glass (500 Å), a 21-mer ODN was assembled on an ABI394B DNA synthesizer, by the phosphoramidite method, according to the manufacturer’s recommendation. Protected 2′-O-deoxyribonucleoside phosphoramidites and S-ethylthiotetrazole were purchased from Glen Research. Ammonia was used for cleavage from the support and deprotection overnight at 55°C. The reaction mixture was then analyzed by reverse-phase high-performance liquid chromatography in ion-pair mode. For fluorescein labeling, 1 mg of fluorescein isothiocyanate dissolved in 20 μl of dimethylformamide was added to a solution of 40 nmol of 2′-amino ODN in 130 μl of sodium carbonate-bicarbonate (1 M, pH 9):water (v/v, 5:8) (22, 23). The mixture was incubated at 25°C in the dark for 18 h and then loaded onto an NAP5-Sephadex G-25 column (Amersham Biosciences) pre-equilibrated in water. The ODN was eluted with water and analyzed by reversed-phase high-performance liquid chromatography in ion-pair mode.

**IN Purification and Standard Analysis of the 3′-Processing Reaction**—HIV-1 IN (32 kDa) was purified under native conditions, as previously described (24). Standard IN assay, using $^3$P-labeled ODN and gel electrophoresis, was carried out as previously described (24). Gels were analyzed on a STORM 840™ PhosphorImager (Amersham Biosciences) and quantified with ImageQuaNT™ 4.1 software. The 3′-processing activity was calculated as follows: activity (%) = 19-mer/(21-mer + 19-mer) × 100.

**Steady-state Fluorescence Anisotropy Assay**—Steady-state anisotropy values were recorded on a Beacon 2000 instrument (PanVera, Madison, WI), in a cell thermostatically held at 25 or 37°C, for the DNA-binding step and 3′-processing reaction, respectively. Unless otherwise stated, we studied the formation of IN-DNA complexes by incubating fluorescein-labeled ds ODNs with IN in 20 mM Tris (pH 7.2), 1 mM dithiothreitol, 20 mM NaCl, 5 mM MgCl$_2$ (the sample volume was 200 μl). The fractional saturation function (Y) was calculated as follows,

$$Y = \frac{r - r_{ODN}}{r_{max} - r_{ODN}} \times 100$$  \hspace{1cm} (Eq. 1)

where $r_{max}$ and $r_{ODN}$ are the anisotropies of IN-bound and free ODN, respectively (no significant concomitant change in fluorescence intensity was observed). After the DNA-binding step, the temperature was raised from 25 to 37°C for monitoring of
the catalytic process. The 3′-processing activity was assessed by quantifying the decrease in $r$. Two independent methods were used for quantification as follows. (i) In fixed-time experiments, the reaction was stopped by adding SDS (0.25% final), disrupting all the IN-DNA complexes in the sample. In such experiments, the solution contained two fluorescent species: the nonprocessed ODN and the fluorescein-labeled dinucleotide released by the cleavage reaction. The fraction of dinucleotides released ($F_{\text{dinu}} = [\text{GT}]/[\text{DNA}]_{\text{total}}$) is given by Equation 2,

$$F_{\text{dinu}} = \frac{r_{\text{NP}} - r}{r_{\text{NP}} - r_{\text{dinu}}} \quad \text{(Eq. 2)}$$

where $r_{\text{NP}}$ and $r_{\text{dinu}}$ are the anisotropy values for pure solutions of non-processed ds ODN and dinucleotide, respectively (fluorescence did not change significantly during the reaction). We used the 5′-GT-3′F dinucleotide (Eurogentec) to determine $r_{\text{dinu}}$. (ii) In real-time conditions, an additional fluorescent population corresponding to IN complexed with the unprocessed ds ODN, is present in the sample. In this case, $F_{\text{dinu}}$ was calculated as follows,

$$F_{\text{dinu}} = \frac{r_{\text{NP}} - r}{r_{\text{max}} - r_{\text{dinu}}} \quad \text{(Eq. 3)}$$

where $r_{\text{max}}$ is the characteristic $r$ value obtained for optimal activity, and $r_{\text{NP}}$ is the $r$ value obtained at the end of the DNA-binding step (before the start of the reaction). The 3′-processing activity obtained with Equations 2 and 3 is referred to as Activity$_{\text{SS/DS}}$ and Activity$_{\text{real-time}}$, respectively. Activity$_{\text{real-time}}$ was not used if $r_{\text{NP}}$ was higher than 0.22 (aggregation of IN on DNA not negligible). We analyzed single-turnover kinetics using the Equations 4 and 5,

$$\ln(1 - F_{\text{dinu}}) = -k_{\text{obs}} \times t \quad \text{(Eq. 4)}$$

with

$$k_{\text{obs}} = k_{\text{chemistry}}/(K_{d,\text{app}}[\text{IN}]_0 + 1) \quad \text{(Eq. 5)}$$

where $k_{\text{chemistry}}$ is the single-turnover rate constant, and $K_{d,\text{app}}$ is the apparent $K_d$ (25). Reactions were conducted in the presence of Mg$^{2+}$ (not Mn$^{2+}$) to limit nonspecific hydrolysis products (cleavages at positions −3, −4, etc.), because anisotropy cannot discriminate between these small products (which are minor products with Mg$^{2+}$ but not with Mn$^{2+}$) and the specific GT product. The $r$ values for the TAMRA-labeled DNA substrate ($\lambda_{\text{ex}} = 562 \text{ nm}$ and $\lambda_{\text{em}} = 582 \text{ nm}$) were recorded on a Cary Eclipse spectrofluorometer (Varian, Mulgrave, Australia) in polarization mode and were compared with values obtained with the same instrument and HIV-a3F as the substrate.

$K_{d,\text{app}}$ as a function of ODN length were determined by competition experiments. Fluorescein-labeled ds ODN HIV-a3F (4 nm) was preincubated with various concentrations of unlabeled ss or ds HIV ODN (from 10- to 45-mers) in 20 mM Tris buffer (pH 7.2) supplemented with 1 mM dithiothreitol, 20 mM NaCl, and 5 mM MgCl$_2$. IN (150 nm final concentration) was then added, and $r$ values were recorded. $\Delta r (= r - r_{\text{ODN}})$ was plotted against competitor concentration to determine $K_{d,\text{app}}$ (concentration of competitor decreasing the initial $\Delta r$ value by 50%).

**Time-resolved Fluorescence Experiments**—Time-resolved fluorescence parameters (lifetimes and correlation times) were obtained from the two polarized fluorescence decays $I_i(t)$ and $I_\perp(t)$, using the time-correlated single photon counting technique. The instrumentation setup was essentially similar to those previously described (16, 17), with modifications: 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 second harmonic generator tuned to 490 nm. The emission monochromator (ARC SpectraPro-150) was set to 530 nm ($\Delta \lambda = 15 \text{ nm}$). The two polarized components were collected alternately over a period of 30 s (total count of $I_o$: 15,000,000). The reaction mixture contained 20 mM Tris, pH 7.2, 20 mM NaCl, 5 mM MgCl$_2$, 1 mM dithiothreitol. The correlation time ($\theta$) distributions of free ss or ds ODNs were obtained at ODN concentrations of 10 nm or 0.5 $\mu$m. All ODNs (from 10- to 45-mer) mimicked the U5-end of the HIV-1 DNA and were fluorescein-labeled at the 5′-end of strand $\alpha$. IN-DNA complexes were analyzed using ds HIV-a5F and various IN-DNA ratios, from 40 to 400. We analyzed both decays, $I_i(t)$ and $I_\perp(t)$, by the maximum entropy method (26). Fluorescence anisotropy decay is described by Equations 6 and 7.

$$r(t) = \sum_{i=1}^{n} \rho_i \times e^{-t/\theta_i} \quad \text{(Eq. 6)}$$

with

$$\sum_{i=1}^{n} \rho_i = \rho_0 \quad \text{(Eq. 7)}$$

where $\theta_i$ is the individual rotational correlation time, and $\rho_i$ is the associated amplitude. $\rho_0$ was found to lie between 0.36 and 0.37. Normalization of $\theta$ for a given temperature was performed using,

$$\theta = \eta V/kT \quad \text{(Eq. 8)}$$

where $\eta$ is the viscosity, $V$ is the volume of the rotating unit, $k$ is the Boltzmann constant, and $T$ is the temperature (K).

**RESULTS**

**Monitoring 3′-Processing Activity by Steady-state Fluorescence Anisotropy**—Fluorescence anisotropy measurements are based on the principle of photoselective excitation of a fluorophore by a polarized light, providing information about rotational motions of the fluorophore or fluorescently labeled molecule between photon absorption and emission. Some events such as overall rotational diffusion or flexibility are major causes of light depolarization. High levels of anisotropy are generally associated with large molecules or complexes characterized by slow rotational diffusion or low flexibility level. In this study, we used an extrinsic fluorophore covalently linked to DNA to monitor the binding of IN to viral DNA substrate and the subsequent 3′-processing reaction, in the same assay. Both DNA binding and 3′-processing would be expected to have a
The initial was started by shifting the sample to a temperature of 37°C, and as previously reported (12, 27). The processing reaction conditions of Fig. 2). Equilibrium was typically reached after 15 min, as previously reported (12). The processing reaction was started by shifting the sample to a temperature of 37°C, and the initial r values obtained therefore correspond to $r_t = 0$. In the presence of the divalent cationic cofactor Mg$^{2+}$, all tested ODNs gave similar $r_t = 0$ values (0.22), indicating that fluorescein position had no effect on DNA binding by IN (Fig. 2B).

We tested 21-mer DNA substrates labeled with fluorescein at one end, in all possible combinations (5’- or 3’-extremity of the a or b strand), or labeled on the 2’-amino group of a 3’-terminal 2’-aminouridine (Fig. 2A). The free ds ODNs were characterized by r values of 0.060–0.130 at 25°C (0.045–0.105 at 37°C), depending on the location of fluorescein, whereas the fluorescein-labeled dinucleotide was characterized by r = 0.02. DNA binding on the addition of IN was monitored at 25°C and led to a significant increase in r value (>0.2 in the experimental conditions of Fig. 2). Equilibrium was typically reached after 15 min, as previously reported (12, 27). The processing reaction was started by shifting the sample to a temperature of 37°C, and the initial r values obtained therefore correspond to $r_t = 0$. In the presence of the divalent cationic cofactor Mg$^{2+}$, all tested ODNs gave similar $r_t = 0$ values (0.22), indicating that fluorescein position had no effect on DNA binding by IN (Fig. 2B).
decreases in \( r \) or \( \Delta r \), and these decreases were strictly related to the presence of the metallic cofactor. In contrast, HIV-\( a5F \), HIV-\( b5F \), and HIV-\( b3F \) displayed no significant decrease in \( r \) or \( \Delta r \) value, under either real-time (Fig. 2) or fixed-time conditions (data not shown) (3′-processing activity was normally detected by a gel-electrophoresis method for HIV-\( b5F \) and HIV-\( b3F \); HIV-\( a5F \) was not tested). The result obtained with these three ODNs indicates that anisotropy is not sensitive enough to differentiate between the DNA substrate (ds ODN 21/21) and the first reaction product (ds ODN 19/21). Furthermore, under real-time conditions, this result suggests that the processed DNA product remains tightly bound to the enzyme after 3′-processing. This tight binding together with the slow catalytic step could be responsible for the observed single-turnover property of IN, even under conditions of excess DNA substrate (25). In contrast, the terminal dinucleotide is released normally from IN-DNA complexes after 3′-processing, and our results indicate that the decrease in \( r \) is related to the formation of the GT product. Thus, anisotropy is reliable for monitoring the 3′-processing reaction, but only if fluorescein is attached to the GT dinucleotide.

The 3′-processing reaction is strongly sequence-dependent, and the endonucleolytic site includes the crucial conserved CA dinucleotide immediately preceding the GT dinucleotide. We assessed sequence specificity, using the anisotropy-based assay and the following ODNs: HIV-\( a3F \) corresponds to the wild-type (wt) sequence, whereas HIV-GTGT-\( a3F \) is a variant in which the 3′-terminal CAGT sequence is replaced by GTGT. We also tested three nonspecific sequences, NS-TTCC-\( a3F \), NS-CACC-\( a3F \), and NS-CAGT-\( a3F \). We found that the DNA-binding step was not influenced by sequence. Indeed, the final \( r \) values obtained after DNA binding were similar for all ODNs, irrespective of sequence context, confirming that in vitro assays primarily reveal the nonspecific DNA binding mode of IN. In contrast, sequence had a major effect on \( r \) in catalysis conditions (Fig. 3A). Only HIV-\( a3F \) gave a large decrease in \( \Delta r \). This decrease was abolished by replacement of the 3′-terminal sequence CAGT by GTGT or the use of nonspecific sequences, even with the CAGT sequence at the 3′-end. These results confirm that the CA dinucleotide is strictly required but not sufficient for activity, particularly in the presence of Mg\(^{2+} \), which gives more stringent conditions than Mn\(^{2+} \) (4, 5). We also investigated the effect of a styrylquinoline compound, FZ41, which inhibits 3′-processing activity (30). The anisotropy-based approach has been shown to be suitable for assays of the effects of inhibitors on the formation and stability of IN-DNA complexes, independently of catalysis, and has been successfully used in studies of competitive inhibitors preventing IN-DNA interactions (12). Here, we investigated both steps, DNA binding and catalysis, in the presence of inhibitor. The IC\(_{50} \) value for 3′-processing was found to be 0.74 μM (Fig. 3B), consistent with the value determined by standard gel-electrophoresis (30). In the same experiment, the study of the DNA-binding step revealed the competitive nature of FZ41: \( r_c = 0 \) decreased with increasing FZ41 concentration (not shown), suggesting that FZ41 primarily prevents the IN-DNA interaction.

Altogether, our data show that the decrease in \( r \) during incubation at 37 °C is related to specific 3′-processing activity and that the anisotropy-based method is reliable for inhibitor characterization. In our experiments, strand transfer products, as quantified by gel-electrophoresis, represent 5–8% of the total products, regardless of fluorescein labeling (data not shown). Only, the strand transfer of one processed DNA into a fluorescently labeled DNA (not yet processed) may underestimate the 3′-processing activity determined by anisotropy. Consequently, the half-transfer reaction is more problematic at the beginning of the reaction, and its influence on \( r \) tends to disappear during 3′-processing. However, this effect is probably small, particularly in the real-time assay, in which the high \( r \) values characterizing the IN-DNA complex cannot be severely influenced by strand transfer.
Quantification Analysis and Characterization of the 3’-Processing Reaction—The 3’-processing activity of IN can be calculated by real-time assay or by quenching of the reaction with SDS. Both these approaches allow quantification of the fraction of released dinucleotides. Fig. 4 displays one example of the time course of GT release in real-time conditions. Activity increased linearly with time over 180 min. This apparent linearity is compatible with single-turnover conditions (excess of IN over DNA substrate), when the rate constant of cleavage is low (25). Under single-turnover conditions, the time course of product formation provides a measure of the actual chemical cleavage step, which is not affected by subsequent turnovers. The kinetic profile, as determined by anisotropy measurement, gave $k_{\text{chemistry}} = 0.0035 \text{ min}^{-1}$, consistent with the value obtained with the standard gel-electrophoresis method ($k_{\text{chemistry}} = 0.004 \text{ min}^{-1}$) (25). A similar result was obtained using the anisotropy method with SDS (data not shown). Indeed, Activity_{real-time} and Activity_{SDS} were well correlated (Fig. 4B). These results demonstrated that (i) all the GT dinucleotides produced by the cleavage reaction were released from the IN-DNA complexes in the real-time assay and were therefore not trapped within the complexes, (ii) the decrease in $r$ in real-time conditions was due exclusively to the release of dinucleotides from the IN-DNA complexes, rather than another phenomenon, such as dissociation of the enzyme from the DNA substrate or product. The single-turnover rate constant, as determined with the anisotropy-based test, was similar to that determined by gel electrophoresis, suggesting that the presence of fluorescein at the DNA 3’-extremity has no effect on DNA-binding or the subsequent catalytic step. Comparison of the 3’-processing activity of non-fluorescent and fluorescein-labeled U5-substrates by gel electrophoresis confirmed that fluorescein has no influence on activity (Fig. 4C). This result sharply contrasts with that obtained using a TAMRA-labeled ODN: this ODN was also tested by both anisotropy and gel-electrophoresis methods and, although TAMRA did not influence the number of IN-DNA complexes, 3’-processing activity was significantly decreased by a factor of 5 (data not shown). Note that the $r$ parameter is not sensitive to the precise positioning of IN on the DNA, and this positioning may be disrupted by the TAMRA probe in the absence of a strong effect on overall DNA-binding isotherms.

The Initial Steady-state Anisotropy Value Is Not Predictive of 3’-Processing Activity—The final $r$ value obtained after the DNA-binding step corresponds to the initial $r$ value for activity ($r_f = 0$) and is related, in first approximation, to the fractional
The IN 3′-Processing Reaction Mechanism

saturation function (Y). Under single-turnover conditions, activity is expected to increase with Y, and the fraction of released dinucleotides should therefore be proportional to the number of IN-DNA complexes initially formed in the sample. Interestingly, a plot of 3′-processing activity against \( r_\text{t=0} \) gave a clear bell-shaped curve (Fig. 5A). The various \( r_\text{t=0} \) values were obtained by varying the initial IN:DNA ratio from 4.2 to 400, corresponding to different mixtures of IN (from 50 to 800 nM) and ODN (from 2 to 12 nM). The optimal ratio was found to be \( r_\text{t=0} \approx 0.226 \) (Fig. 5B) and corresponds to an \( r_\text{t=0} \) value of 0.226. Larger values (up to 0.272) led to a significantly decrease in the 3′-processing activity. Increases in \( r_\text{t=0} \) may occur for two reasons: (i) the number of free DNA molecules decreases with increasing IN:DNA ratio and (ii) self-assembly of IN on DNA may further increase the size of the complexes. The first phase \( (r_\text{t=0} < 0.226) \) most likely corresponds to an increase in the number of complexes, increasing both \( r_\text{t=0} \) and product formation. In the second phase \( (r_\text{t=0} > 0.226) \), the activity drops off suggesting that high order multimeric forms of IN on DNA, possibly aggregates, formed under conditions of high IN:DNA ratio and characterized by high \( r \) values, are less active than lower order multimeric forms. Furthermore, the \( r_\text{max} \) value corresponding to saturation ([DNA] \( _\text{free} \approx 0 \)) must be below 0.272, and, thus, 0.226 is the best estimate for \( r_\text{max} \).

To get deeper insight into the bell-shaped phenomenon, we investigated the relationship between the fraction of product obtained at the end of the reaction and the \( Y \) parameter, which is related to the number of DNA-substrate molecules initially bound to IN (at \( t = 0 \)). The initial number of complexes was varied, and activity at \( L_\text{t} \) (corresponding to the maximum amount of released product) was determined and plotted as a function of \( Y \) (Fig. 6, A–C). The fraction of product was found to increase linearly with the initial number of IN-DNA complexes, as expected under single-turnover conditions. Optimal conditions, corresponding to an IN:DNA ratio of \( \sim 40 \) (characterized by \( r = 0.226 \)), gave a final activity of 92%. This result confirms that complete saturation of the DNA substrate can be achieved at anisotropy values much lower than 0.272, and is compatible with \( r_\text{max} = 0.226 \). The activity per complex, specific activity \( (\text{activity}/Y) \), was then calculated for each IN:DNA ratio (Fig. 6D). Specific activity decreased as a function of IN:DNA ratio (Fig. 6D), suggesting that aggregation occurs as soon as IN is present in excess over DNA and that the activity of large complexes (IN\text{agg}-DNA) is lower than that of the lower order multimeric form (IN\text{act}-DNA). At ratios exceeding 200:1, the samples contained only IN\text{agg}-DNA complexes \( (r = 0.272; \text{see also TFA experiments below}) \). However, these complexes retained significant levels of 3′-processing activity \( (=35\% \text{ the activity of IN\text{act}-DNA complexes}) \). The number of aggregated forms of IN increased continuously with the IN:DNA ratio, suggesting that, at the peak of the bell-shaped curve, the sample already contained some IN\text{agg}-DNA complexes, together with the IN\text{act}-DNA complexes. From Fig. 6, we can estimate that IN\text{agg}-DNA complexes in these conditions accounted for \( \sim 11–15\% \) of total complexes. We made use of the additivity law of anisotropy to calculate the theoretical \( r_\text{max} \) value (0.218) characterizing the fully active complex (IN\text{act}-DNA), in the absence of both free DNA and aggregated complexes. Thus, the bell-shaped curve shown in Fig. 5A results from two opposing trends: \( Y \) and [IN\text{agg}-DNA] both increase as a function of IN:DNA ratio and are favorable and unfavorable, respectively, for 3′-processing activity.

Characterization of the IN\text{act}-DNA and IN\text{agg}-DNA Complexes Using TFA—Under site saturation conditions, \( r \) values reflect the molecular size of the complexes in a purely indicative manner. We evaluated the molecular sizes of active IN-DNA complexes in greater detail by studying representative samples corresponding to different \( r \) values, as shown in Fig. 5A, independently by TFA. This approach allows the determination of rotational correlation times (\( \theta \)), which are directly related to hydrodynamic volume according to Equation 8. We first characterized ss and ds ODNs of various lengths (10- to 45-mer) in the absence of IN (Table 1). All ODNs were characterized by short \( \theta \) values (\( \theta_1 \) and \( \theta_2 \)), which can reasonably be assigned to rotation of the fluorescein moiety around the linker at the 5′-end of the DNA and the flexibility of the linker itself. Unlike \( \theta_1 \) and \( \theta_2 \), longer \( \theta \) values (\( \theta_3 \) and \( \theta_4 \)) strongly depend on length and thus account for the overall rotational motion of the ODN. For each length, the longest correlation time was 2–2.7 times longer for ds than for ss ODNs. For example, the \( \theta_4 \) values for
The IN 3′-Processing Reaction Mechanism

![Graph showing product formation](image)

**FIGURE 6. Product formation is related to the initial fractional saturation function.** A, 3′-processing kinetics for different IN:DNA ratios. Activity was measured using different ratios (below 40:1, to limit aggregation). HIV-a3F concentration was 4 nM with 10 (black squares), 20 (white squares), 30 (black circles), 40 (white circles), 50 (black triangles), 60 (white triangles), or 70 nM (gray circles) IN. MgCl₂ concentration was 10 mM. B, 3′-processing kinetics in conditions of increasing ionic strength: 30 mM (black circles), 35 mM (white squares), or 40 mM MgCl₂ (black squares). Concentrations of ODN and IN were 4 nM and 150 nM, respectively. Activities in A and B were calculated using $\Delta \tau_{obs-process}$, C, 3′-processing activity at $t_\text{max}$ (Act$_{max}$) as a function of initial fractional saturation function. Act$_{max}$ corresponds to the maximal fraction of released dinucleotides for given conditions of IN:DNA ratio or MgCl₂ concentration (plateau in A and B) and was estimated by fitting: $\text{Act} = \text{Act}_{max} \times (1 - e^{-\frac{t}{\text{Kd}}})$. Y was calculated using $\text{t}_{\text{max}} = 0.226$. D, specific activity ($\text{A} = \text{activity} / Y$) as a function of the IN:DNA ratio. Total activity was normalized using fractional saturation function ($Y$), giving 3′-processing activity per complex.

**TABLE 1**

Correlation times ($\theta$) obtained for fluorescein-labeled ODNs of various lengths at 25°C

| ODN     | Single-stranded$^a$ | Double-stranded |
|---------|---------------------|-----------------|
| 10-mer  | $\theta_1 = 0.12 \pm 0.02 (0.385)$ | $\theta_1 = 0.23 \pm 0.05 (0.48)$ |
|         | $\theta_2 = 0.37 \pm 0.06 (0.385)$ | $\theta_2 = 1.05 \pm 0.51 (0.31)$ |
| 13-mer  | $\theta_1 = 1.48 \pm 0.08 (0.23)$ | $\theta_1 = 3.50 \pm 1.17 (0.21)$ |
|         | $\theta_2 = 0.38 \pm 0.02 (0.285)$ | $\theta_2 = 0.84 \pm 0.02 (0.26)$ |
| 21-mer (HIV-a5F) | $\theta_1 = 1.64 \pm 0.05 (0.215)$ | $\theta_1 = 3.83 \pm 0.38 (0.25)$ |
|         | $\theta_2 = 0.15 \pm 0.04 (0.28)$ | $\theta_2 = 0.16 \pm 0.03 (0.425)$ |
| 36-mer  | $\theta_1 = 0.42 \pm 0.08 (0.405)$ | $\theta_1 = 0.47 \pm 0.11 (0.26)$ |
|         | $\theta_2 = 1.48 \pm 0.25 (0.16)$ | $\theta_2 = 1.75 \pm 0.56 (0.19)$ |
| 45-mer  | $\theta_1 = 4.27 \pm 0.63 (0.155)$ | $\theta_1 = 8.58 \pm 2.31 (0.125)$ |
|         | $\theta_2 = 0.12 \pm 0.03 (0.35)$ | $\theta_2 = 0.17 \pm 0.02 (0.46)$ |
|         | $\theta_3 = 0.39 \pm 0.06 (0.39)$ | $\theta_3 = 0.65 \pm 0.17 (0.22)$ |
|         | $\theta_4 = 1.62 \pm 0.48 (0.17)$ | $\theta_4 = 2.70 \pm 0.56 (0.17)$ |

$^a$ $\rho_i/\rho_0$ ratios are indicated in parentheses.

The 21-mer ODN was 4.3 and 8.6 ns for ss and ds molecules, respectively. We found that the 21-mer ds ODNs HIV-a5F, HIV-b3F, HIV-b5F, NS-TTCC-a3F, and NS-CACC-a3F displayed similar $\theta$ distributions and were characterized by $r$ values of $0.045-0.055$ (not shown). In contrast, the 21-mer ds ODNs HIV-a3F, HIV-a1F, HIV-GTGT-a3F, and NS-CAGT-a3F, which had higher $r$ values (0.105), gave similar correlation times to the previous series, but with different $\rho_i/\rho_0$ ratios (data not shown). For example, HIV-a3F had a long $\theta_1$ (9.6 ns), similar to that of HIV-a5F (8.6 ns), but the flexibility of the fluorophore was significantly lower, as demonstrated by the higher $\rho_i/\rho_0$ ratio ($= 0.5$) than for HIV-a5F ($\rho_i/\rho_0 = 0.125$) (the fractional amplitude of fast motions is lower for HIV-a3F than for HIV-a5F). This decrease in flexibility seems to be related to the presence of a 3′-terminal GT and accounts for the higher $r$ value (similar results were obtained using TAMRA; data not shown). Finally, the 21-mer ODN was the most suitable substrate for studies of IN-DNA complexes by TFA, because longer ODNs (36 or 45 bp) had a long $\theta$, which could complicate the analysis of IN-DNA complexes. Moreover, shorter ODNs (10 or 13 bp) were not chosen for reasons of affinity, because we found that ODN length had a strong effect on $K_d$ (Fig. 7).

We then applied TFA to different samples corresponding to various IN:DNA ratios. We used HIV-a5F to prevent analytical problems due to the activity-dependent decrease in $r$ when fluorescein is linked to the GT dinucleotide. The activity observed at $t_\text{max}$ for an IN:DNA ratio of 40:1 (92%) suggested that only a small amount of free ODN was present in this sample. We therefore restricted our analysis to samples corresponding to the peak of the bell-shaped curve and to the decreasing phase, so as to avoid studying mixtures of free and IN-bound ODNs. At 37°C, two short $\theta$ (<2.5 ns) were resolved upon the addition of IN to ODN (Table 2). The distribution in the short $\theta$ domain was not dependent on the IN:DNA ratio. In contrast, the characteristic long $\theta$ at 25°C of 8.6 ns found for free DNA was replaced by a $\theta_{25°C} \approx 37.8$ ns (26.8 ns at 37°C), depending on the IN:DNA ratio (Fig. 8 and Table 2). The sample corresponding to the peak of the bell-shaped curve (IN:DNA 40:1; $r = 0.226$) was rather monodisperse, as suggested by the small difference between the maximum (26.8 ns at 37°C) and the center of gravity (31.4 ns at 37°C) of the peak in the $\theta$ distribution. As discussed above, samples corresponding to the TFA distribution shown in Fig. 8A contained mostly non-aggregated active forms of IN (85–89%). These complexes are characterized by a long correlation time of ~38 ns. This value is consistent with hydrodynamic properties of a dimeric form of IN on the DNA (16, 17, 31). The shift of the center of gravity toward longer $\theta$ values probably results from the presence of small numbers of IN$_\text{agg}$-DNA complexes. Further increases in the IN:DNA ratio displaced the
distribution toward longer $\theta$ values (Fig. 8, B–E). Our data confirmed that the sample corresponding to the highest $r$ value (0.272) consisted essentially of aggregated forms of IN on DNA (Fig. 8E), as demonstrated by the broad distribution above 100 ns (17).

**DISCUSSION**

We used a continuous fluorescence assay to monitor the HIV-1 IN 3′-processing reaction. We found that the processed DNA product remained tightly bound to IN, thereby limiting enzyme turnover, whereas the other reaction product (GT dinucleotide) was rapidly released from the processed DNA-IN complex. This assay made it possible to characterize standard equilibrium/kinetic parameters and to obtain insight into the relationship between the size of IN-DNA complexes and catalytic activity. We studied 3′-processing under various IN:DNA ratio conditions: optimal activity was obtained at a ratio of 40:1. Higher ratios, associated with high $r$ values, led to a decrease in product formation. The oligomeric status of active IN was further investigated by TFA. The best condition for 3′-processing activity corresponds to a majority of complexes in the sample characterized by a long correlation time of $\sim$38 ns. This value is compatible with a dimeric form of IN bound to one short DNA substrate mimicking a single viral DNA end. The aggregation of IN on DNA was found to be responsible for the decrease in the 3′-processing activity.

**Kinetic Analysis of 3′-Processing Activity by the Anisotropy-based Method**—The real-time assay described here is based on the use of a DNA duplex fluorescein-labeled on the terminal-3′ GT. It allows quantification of the cleavage reaction resulting in physical separation of the processed 19/21-mer DNA (first product) and the fluorescently labeled dinucleotide (second product). IN assays based on fluorescence intensity (fluorescence resonance energy transfer or fluorescence quenching) have been described in previous studies (32–35). Unlike electrophoresis-based methods, which can be used to analyze discrete time points only, fluorescence methods present the significant advantage of allowing continuous monitoring of the reaction for kinetic studies of IN in the same sample. However, only the anisotropy-based method can be used to study both the DNA-binding step and subsequent activity in the same assay. Using various DNA sequences, we showed that IN does not demonstrate highly selective binding to the viral sequence over other sequences, but, in contrast, displayed strong selectivity for catalysis (Fig. 3). This result is not subject to possible bias due to indirect comparison, as might be the case when DNA binding and catalysis are studied by different methods. It suggests that IN first binds viral DNA in a nonspecific manner and that subsequent relaxation leads to the formation of a specific and catalytically competent complex before catalysis. This relaxation step must be very slow to account for the low single-turnover rate constant as previously reported (25).

The anisotropy-based assay is reliable, because $r$ decreased only in the presence of Mg$^{2+}$ with the wt DNA sequence and only when fluorescein was attached to the GT 3′-extremity and was prevented by a competitive inhibitor of IN. The additivity law of anisotropy renders the quantification of activity easy and both real-time and fixed-time studies can be carried out. The real-time assay is possible, because IN-DNA complexes with DNA substrate (21/21-mer) or product (processed 19/21-mer) are highly stable in solution (Fig. 2), and the observed decrease in $r$ is therefore due solely to the cleavage reaction and not to dissociation of the protein from DNA. Unlike many standard methods, which quantify the 19/21-mer product under denaturing conditions, the real-time assay allows quantification of the GT product under native conditions. Our kinetic analysis gave a catalytic rate of 0.2 h$^{-1}$, which is similar to the values published in previous studies (25, 36, 37). The strong correlation between real-time and fixed-time data (Fig. 4B) and similarities between the time courses of GT and 19/21-mer formation demonstrate that the GT dinucleotide is rapidly released from IN after the 3′-processing reaction. This release is consistent with the low apparent affinity expected for such a small ligand (see Fig. 7, $K_a > 100 \mu M$). This situation is very different to that for the other reaction product (19/21-mer), which remains tightly bound to IN after catalysis. Finally, kinetic analysis showed that the catalytic rate constant was not significantly affected by the presence of fluorescein at the end of the DNA molecule, whereas inhibition was observed with TAMRA.

![FIGURE 7. Apparent $K_d$ values of IN as a function of DNA length (number of nucleotides). Values of $K_d$ were determined for ss (white squares) or ds ODNs (black squares) as indicated under “Experimental Procedures.”](image)

**TABLE 2**

| Correlation times | 40:1 | 67:1 | 100:1 | 200:1 | 400:1 |
|-------------------|------|------|-------|-------|-------|
| $\theta_1$        | 0.29 (0.31) | 0.19 (0.38) | 0.17 (0.33) | 0.15 (0.30) | 0.19 (0.26) |
| $\theta_2$        | 1.93 (0.23) | 2.13 (0.14) | 2.14 (0.17) | 2.6 (0.15) | 2.5 (0.15) |
| $\theta_3$, maximum | 26.8 (0.46) | 38.5 (0.48) | 53 (0.50) | N.D (0.55) | N.D (0.59) |
| $\theta_3$, conjugate of gravity | 31.4 | 53.3 | >90 | >120 | >150 |

* $\mu_1/\mu_2$ ratios are indicated in parentheses.
* Normalization of the $\theta_1$ value for a temperature of 25 °C (according to Equation 8), allowing direct comparison with $\theta$ values in Table 1.
showed that the local motion of fluorescein or TAMRA was more restricted when these fluorophores were covalently linked to a GT 3'-extremity, suggesting a direct interaction between the fluorophore and DNA. In the case of fluorescein, this interaction perturbed neither DNA binding nor cleavage. In contrast, the additional methyl groups and (or) charge effects of TAMRA may prevent the correct positioning of IN on the cleavage site or phosphodiester bond hydrolysis itself.

**Dimeric Forms of IN Are Optimal for the 3'-Processing of One DNA Extremity**—We found that the optimal conditions for 3'-processing were obtained for a majority of dimeric forms bound to the DNA ($\theta_{\text{long}} = 38$ ns), consistent with complementation studies between individually inactive IN mutants, suggesting that at least the dimeric form is catalytically active for 3'-processing (38, 39). The optimal IN:DNA ratio (40:1) reported here corresponds to conditions used in most IN assays: IN in excess over DNA and DNA concentration maintained below the $K_d$ value. This ratio therefore does not indicate the final stoichiometry and is highly dependent on experimental conditions. For example, increasing ionic strength shifts the $K_d$ toward higher values and influences the optimal ratio (data not shown).

Under experimental conditions in which DNA and IN concentrations were maintained above the $K_d$ value, an optimal ratio of 2:1 was found, consistent with the stoichiometric binding of two IN protomers per DNA end (25). We and others (12, 40) have previously observed cooperative binding of IN to DNA. Consequently, using DNA concentration below the $K_d$ value, the IN:DNA ratio determines the number of complexes formed in solution as well as the stoichiometry of these complexes. Therefore, below the optimal ratio for activity, it is very likely that, in addition to free DNA, the solution contains a mixture of monomeric and dimeric forms of IN bound to the DNA substrate. Indeed, we have previously found that, using suboptimal IN:DNA ratios (10:1), the complexes exhibit lower $\theta$ values consistent with a monomer-dimer equilibrium at 37 °C (normalized $\theta_{\text{long}}$, 24–27 ns) and monomeric IN at 25 °C (normalized $\theta_{\text{long}}$, 16–20 ns) (16). Here, it is highly suggested that the optimal activity originates in two critical events: DNA binding and cooperativity. From this point of view, steady-state anisotropy is then particularly helpful to establish direct relationships between DNA binding parameters, the specific activity of the complexes, and their hydrodynamic properties as measured by TFA.

Our conclusion that the dimeric form of IN is active for 3'-processing under native conditions strongly parallels those obtained by Parissi and coworkers (41) with purified cross-linked multimers. The aggregation of IN on DNA was found to be strongly dependent on the IN:DNA ratio and was detrimental for 3'-processing, as previously reported for concerted integration (42). The bell-shaped curve obtained in Fig. 5 most likely originates from competition between the cooperative binding of IN on DNA leading to a dimeric form (specific IN-IN interaction, which is favored by increasing the IN:DNA ratio and enhances 3'-processing activity) and aggregation (nonspecific interactions also favored by increasing the IN:DNA ratio but detrimental for 3'-processing). The bell-shaped behavior is also compatible with the results of a recent cross-linking study (41) suggesting that monomers and tetramers display neither 3'-processing or half-strand transfer (transfer of a single viral DNA end).
activity, whereas dimers are competent for both these reactions, and only tetramers catalyze full-site concerted integration. Our results suggest that, in vitro, IN aggregates do not much hinder 3′-processing under IN:DNA ratio conditions leading to the formation of dimers, but that such aggregates may impede tetrameric organization in a highly competitive manner. However, TFA experiments did not indicate whether the tetrameric form on one LTR end actually existed, because the intermediate θ distributions (Fig. 8, B–D) probably corresponded to mixtures of different multimeric forms (polydisperse solutions). It is therefore difficult to assess 3′-processing activity of tetramers, because IN:DNA ratios shifting dimer-tetramer equilibrium toward tetramers also gave rise to IN aggregation. Although recombinant IN alone is sufficient for concerted integration in vitro (41–44), this reaction remains much less efficient with recombinant IN than with the purified pre-integration complex (45) or virion-derived IN (46). The main limiting factor may be the multimeric organization of IN to yield the correct tetrameric form on DNA compatible with full integration process: aggregation is probably more limiting for DNA integration, which requires tetramers, than for 3′-processing, which requires dimers. Moreover, samples containing many aggregated forms displayed significant 3′-processing activity (one-third that of dimeric IN). The reason for this residual activity remains unclear.

Li and Craigie (44) recently showed that blunt-ended DNAs are more efficient substrates for concerted integration than preprocessed DNAs, whereas half-integration is more efficient with preprocessed DNA. These data, together with our results and those of Parissi et al. (41), suggest that dimers of IN are necessary and sufficient for 3′-processing and half-integration and that the 3′-processing step probably promotes a conformational change of IN compatible with concerted integration. This conformational change may play a key role in promoting the IN tetramerization required for concerted integration. IN is known to exist in a monomer-dimer-tetramer equilibrium in solution (16, 17, 24, 38). The multimeric form is modified on DNA, and the final quaternary structure depends on the IN:DNA ratio and probably also on the nature of the DNA substrate. From this point of view, the tetrameric organization of IN (a dimer of dimers) may be stabilized by the simultaneous presence of two LTR ends (47) and it highly suggested that the condensation-distortion properties of DNA play a key role in bringing the two ends of the viral DNA close enough in space for a dimer of dimers to form. A dimer at each LTR end, as described here, is consistent with current models based on cross-linking (48) or molecular dynamics data (49), and the 5-bp separation at target level is compatible with a tetrameric complex rather than a dimer. Several models of the tetrameric active complex exist for concerted integration: the 5-bp separation is directly compatible with a tetrameric complex displaying 2-fold symmetry, as modeled by McCammon and coworkers (31) or a tetrameric complex assuming target DNA bending (48, 50) or a hinging motion of the dimer interface (7). The formation of a dimer of dimers is probably mediated by the C-terminal domain, as suggested by the x-ray structure of the catalytic core/C-terminal double domain (9). Moreover, residue Trp-235 in the C-terminal domain has a significant effect on concerted integration but not on half-integration (44). The resulting symmetric organization, dispensable for 3′-processing but not for integration, may be favored or stabilized by the symmetric base sequence of the target DNA. Such symmetry in integration sites has recently been reported (51, 52).

It is noteworthy that the single-turnover rate constant of IN is particularly slow, for example about one thousandth that for restriction enzyme HincII (53). This indicates that a pre-catalysis step has a limiting effect independent from that for product release. We found that dimers were the major oligomeric forms of IN on one LTR end. IN free in solution was mostly tetrameric in the 100–200 nM concentration range and aggregated at concentrations above 250 nM (24). This suggests that the dissociation of high order multimeric forms precedes or occurs simultaneously to DNA binding, as previously reported (16) (Fig. 9). Parissi and coworkers (41) carried out a kinetic analysis of the IN-DNA complexes formation and found that the presence of one LTR induces the binding of one IN monomer (M) prior to formation of the IN dimer (D). The conversion of high order multimeric forms of IN into monomers may account for the rather slow DNA-binding step (12, 25, 27). We have previously shown that IN binds cooperatively to DNA with a Hill number compatible with our results and those of Parissi et al. (41), suggest that the catalytic rate is directly compatible with a tetrameric complex assuming target DNA bending (48, 50) or a hinging motion of the dimer interface (7). The formation of a dimer of dimers is probably mediated by the C-terminal domain, as suggested by the X-ray structure of the catalytic core/C-terminal double domain (9). Moreover, residue Trp-235 in the C-terminal domain has a significant effect on concerted integration but not on half-integration (44). The resulting symmetric organization, dispensable for 3′-processing but not for integration, may be favored or stabilized by the symmetric base sequence of the target DNA. Such symmetry in integration sites has recently been reported (51, 52).

It is noteworthy that the single-turnover rate constant of IN is particularly slow, for example about one thousandth that for restriction enzyme HincII (53). This indicates that a pre-catalysis step has a limiting effect independent from that for product release. We found that dimers were the major oligomeric forms of IN on one LTR end. IN free in solution was mostly tetrameric in the 100–200 nM concentration range and aggregated at concentrations above 250 nM (24). This suggests that the dissociation of high order multimeric forms precedes or occurs simultaneously to DNA binding, as previously reported (16) (Fig. 9). Parissi and coworkers (41) carried out a kinetic analysis of the IN-DNA complexes formation and found that the presence of one LTR induces the binding of one IN monomer (M) prior to formation of the IN dimer (D). The conversion of high order multimeric forms of IN into monomers may account for the rather slow DNA-binding step (12, 25, 27). We have previously shown that IN binds cooperatively to DNA with a Hill number compatible with our results and those of Parissi et al. (41), suggest that the catalytic rate is directly compatible with a tetrameric complex assuming target DNA bending (48, 50) or a hinging motion of the dimer interface (7). The formation of a dimer of dimers is probably mediated by the C-terminal domain, as suggested by the X-ray structure of the catalytic core/C-terminal double domain (9). Moreover, residue Trp-235 in the C-terminal domain has a significant effect on concerted integration but not on half-integration (44). The resulting symmetric organization, dispensable for 3′-processing but not for integration, may be favored or stabilized by the symmetric base sequence of the target DNA. Such symmetry in integration sites has recently been reported (51, 52).

It is noteworthy that the single-turnover rate constant of IN is particularly slow, for example about one thousandth that for restriction enzyme HincII (53). This indicates that a pre-catalysis step has a limiting effect independent from that for product release. We found that dimers were the major oligomeric forms of IN on one LTR end. IN free in solution was mostly tetrameric in the 100–200 nM concentration range and aggregated at concentrations above 250 nM (24). This suggests that the dissociation of high order multimeric forms precedes or occurs simultaneously to DNA binding, as previously reported (16) (Fig. 9). Parissi and coworkers (41) carried out a kinetic analysis of the IN-DNA complexes formation and found that the presence of one LTR induces the binding of one IN monomer (M) prior to formation of the IN dimer (D). The conversion of high order multimeric forms of IN into monomers may account for the rather slow DNA-binding step (12, 25, 27). We have previously shown that IN binds cooperatively to DNA with a Hill number compatible with our results and those of Parissi et al. (41), suggest that the catalytic rate is directly compatible with a tetrameric complex assuming target DNA bending (48, 50) or a hinging motion of the dimer interface (7). The formation of a dimer of dimers is probably mediated by the C-terminal domain, as suggested by the X-ray structure of the catalytic core/C-terminal double domain (9). Moreover, residue Trp-235 in the C-terminal domain has a significant effect on concerted integration but not on half-integration (44). The resulting symmetric organization, dispensable for 3′-processing but not for integration, may be favored or stabilized by the symmetric base sequence of the target DNA. Such symmetry in integration sites has recently been reported (51, 52).

It is noteworthy that the single-turnover rate constant of IN is particularly slow, for example about one thousandth that for restriction enzyme HincII (53). This indicates that a pre-catalysis step has a limiting effect independent from that for product release. We found that dimers were the major oligomeric forms of IN on one LTR end. IN free in solution was mostly tetrameric in the 100–200 nM concentration range and aggregated at concentrations above 250 nM (24). This suggests that the dissociation of high order multimeric forms precedes or occurs simultaneously to DNA binding, as previously reported (16) (Fig. 9). Parissi and coworkers (41) carried out a kinetic analysis of the IN-DNA complexes formation and found that the presence of one LTR induces the binding of one IN monomer (M) prior to formation of the IN dimer (D). The conversion of high order multimeric forms of IN into monomers may account for the rather slow DNA-binding step (12, 25, 27). We have previously shown that IN binds cooperatively to DNA with a Hill number compatible with our results and those of Parissi et al. (41), suggest that the catalytic rate is directly compatible with a tetrameric complex assuming target DNA bending (48, 50) or a hinging motion of the dimer interface (7). The formation of a dimer of dimers is probably mediated by the C-terminal domain, as suggested by the X-ray structure of the catalytic core/C-terminal double domain (9). Moreover, residue Trp-235 in the C-terminal domain has a significant effect on concerted integration but not on half-integration (44). The resulting symmetric organization, dispensable for 3′-processing but not for integration, may be favored or stabilized by the symmetric base sequence of the target DNA. Such symmetry in integration sites has recently been reported (51, 52).
The IN 3′-Processing Reaction Mechanism

with the binding of monomers and the presence of a dimeric form of IN on DNA (12). We suggest that relaxation of the IN-DNA complex leading to the formation of a competent and specifically bound dimeric form (D∗) is probably the limiting step accounting for the slow single turnover of 3′-processing. High IN concentrations lead to aggregation on DNA, as shown here by TFA and in other studies by means of a fluorescence correlation spectroscopy approach (54). However, we found that maximum 3′-processing activity was not controlled by absolute IN concentration but instead depended on the IN:DNA ratio. The critical IN concentration for solubility is valuable only for a given DNA substrate concentration, because aggregation properties on DNA are controlled mainly by the IN:DNA ratio. Mostly dimeric forms bound to DNA were found with, for example, a 500 nM IN solution in the presence of 12 nM DNA substrate, although DNA-free IN essentially formed aggregates at concentrations >250 nM. This does not imply that this solution is monodisperse, because IN concentration exceeds DNA concentration, but this result is consistent with the dissociation effect of DNA on high multimeric forms of IN reported in conditions of DNA excess (16). Thus, the solubility of IN may be increased by increasing DNA concentration or, alternatively, DNA binding may promote a specific conformation of IN compatible with catalysis.

Acknowledgments—We thank Dr. H. Leh for helpful discussions and G. Mbemba for technical assistance.

REFERENCES

1. Lutke, R. A., Vink, C., and Plasterk, R. H. (1994) *Nucleic Acids Res.* 22, 4125–4131
2. Engelman, A., Hickman, A. B., and Craigie, R. (1994) *J. Virol.* 68, 5911–5917
3. Jenkins, T. M., Esposito, D., Engelman, A., and Craigie, R. (1997) *EMBO J.* 16, 6849–6859
4. Esposito, D., and Craigie, R. (1998) *EMBO J.* 17, 5832–5843
5. Agapkina, J., Smolov, M., Barbe, S., Zubin, E., Zatsepin, T., Deprez, E., Leh, H., Brodin, P., Zucchi, E., Auclair, C., and Mouscadet, J. F. (2000) *Biochemistry* 39, 9285–9294
6. Smolov, M., Gottikh, M., Tashlitski, V., Korolev, S., Demidylvuk, I., Brochon, J. C., Mouscadet, J. F., and Deprez, E. (2006) *FEBS J.* 273, 1137–1151
7. Brochon, J. C. (1994) *Methods Enzymol.* 240, 262–311
8. Pinskaya, M., Romanova, E., Volkov, E., Deprez, E., Leh, H., Brochon, J. C., Mouscadet, J. F., and Gottikh, M. (2004) *Biochemistry* 43, 8735–8743
9. Hazuda, D. J., Wolfe, A. L., Hastings, J. C., Robbins, H. L., Graham, P. L., LaFemina, R. L., and Emini, E. A. (1994) *J. Biol. Chem.* 269, 3999–4004
10. Vink, C., Lutke, R. A., and Plasterk, R. H. (1994) *Nucleic Acids Res.* 22, 4103–4110
11. Zouhiri, F., Mouscadet, J. F., Meekour, K., Desmaele, D., Savoure, D., Leh, H., Subra, F., Le Bret, M., Auclair, C., and D’Angelo, J. (2000) *J. Med. Chem.* 43, 1533–1540
12. Podtelezhnikov, A. A., Gao, K., Bushman, F. D., and McCammon, J. A. (2003) *Biopolymers* 68, 110–120
13. Lee, S. P., Censullo, M. L., Kim, H. G., Knutson, J. R., and Han, M. K. (1995) *Anal. Chem.* 227, 295–301
14. Hawkins, M. E., Pfleiderer, W., Mazumder, A., Pommier, Y. G., and Balis, F. M. (1995) *Nucleic Acids Res.* 23, 2872–2880
15. Wang, Y., Klock, H., Yin, H., Wolff, K., Biez, K., Niswonger, K., Matzen, J., Gunderson, D., Hale, J., Lesley, S., Kuhren, K., Caldwell, J., and Brinker, A. (2005) *J. Biol. Chem. Struct.* 10, 456–462
16. Bera, S., Vora, A. C., Chiu, R., Heyduk, T., and Grandgenett, D. P. (2005) *Biochemistry* 44, 15106–15114
17. Tramontano, E., Colla, P. L., and Cheng, Y. C. (1998) *Biochemistry* 37, 7257–7283
18. Lee, S. P., Kim, H. G., Censullo, M. L., and Han, M. K. (1995) *Biochemistry* 34, 10205–10214
19. Engelman, A., Bushman, F. D., and Craigie, R. (1993) *EMBO J.* 12, 3269–3275
20. van den Bent, F. M., Vos, A., and Plasterk, R. H. (1999) *J. Virol.* 73, 3176–3183
21. Pemberton, I. K., Buckle, M., and Buc, H. (1996) *J. Biol. Chem.* 271, 1498–1506
22. Fauro, A., Calmes, C., Desjoubert, C., Castroviejo, M., Caumont-Sarcos, A., Tarrago-Litvak, L., Litvak, S., and Parissi, V. (2005) *Nucleic Acids Res.* 33, 977–986
23. Sinha, S., Pursley, M. H., and Grandgenett, D. P. (2002) *J. Virol.* 76, 3105–3113
24. Sinha, S., and Grandgenett, D. P. (2005) *J. Virol.* 79, 8208–8216
25. Li, M., and Craigie, R. (2005) *J. Biol. Chem.* 280, 29334–29339
26. Farnet, C. M., and Bushman, F. D. (1997) *Cell* 88, 483–492
27. Goodarzi, G., Im, G. J., Brackmann, K., and Grandgenett, D. (1995) *J. Virol.* 69, 6090–6097
28. Li, M., Mizuuchi, M., Burke, T. R., and Craigie, R. (2006) *EMBO J.* 25, 1295–1304
48. Gao, K., Butler, S. L., and Bushman, F. (2001) *EMBO J.* **20**, 3565–3576
49. Chen, A., Weber, I. T., Harrison, R. W., and Leis, J. (2006) *J. Biol. Chem.* **281**, 4173–4182
50. Wielens, J., Crosby, I. T., and Chalmers, D. K. (2005) *J. Comput. Aided Mol. Des.* **19**, 301–317
51. Wu, X., Li, Y., Crise, B., Burgess, S. M., and Munroe, D. J. (2005) *J. Virol.* **79**, 5211–5214
52. Holman, A. G., and Coffin, J. M. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 6103–6107
53. Etzkorn, C., and Horton, N. C. (2004) *Biochemistry* **43**, 13256–13270
54. Vercammen, J., Maertens, G., Gerard, M., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2002) *J. Biol. Chem.* **277**, 38045–38052