Published online 19 July 2011 *Nucleic Acids Research, 2011, Vol. 39, No. 20* ***9009–9022***
doi:10.1093/nar/gkr581

FRET analysis reveals distinct conformations of IN tetramers in the presence of viral DNA or LEDGF/p75

Jacques J. Kessl[1], Min Li[2], Michael Ignatov[3], Nikolozi Shkriabai[1], Jocelyn O. Eidahl[1], Lei Feng[1], Karin Musier-Forsyth[3], Robert Craigie[2] and Mamuka Kvaratskhelia[1,*]

[1]Center for Retrovirus Research and Comprehensive Cancer Center, College of Pharmacy, The Ohio State University, Columbus, OH 43210, [2]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, 20892, [3]Departments of Chemistry and Biochemistry, Center for Retrovirus Research and Center for RNA Biology, The Ohio State University, Columbus, OH 43210, USA

Received October 28, 2010; Revised June 29, 2011; Accepted June 30, 2011

ABSTRACT

A tetramer of HIV-1 integrase (IN) stably associates with the viral DNA ends to form a fully functional concerted integration intermediate. LEDGF/p75, a key cellular binding partner of the lentiviral enzyme, also stabilizes a tetrameric form of IN. However, functional assays have indicated the importance of the order of viral DNA and LEDGF/p75 addition to IN for productive concerted integration. Here, we employed Förster Resonance Energy Transfer (FRET) to monitor assembly of individual IN subunits into tetramers in the presence of viral DNA and LEDGF/p75. The IN–viral DNA and IN–LEDGF/p75 complexes yielded significantly different FRET values suggesting two distinct IN conformations in these complexes. Furthermore, the order of addition experiments indicated that FRET for the preformed IN–viral DNA complex remained unchanged upon its subsequent binding to LEDGF/p75, whereas pre-incubation of LEDGF/p75 and IN followed by addition of viral DNA yielded FRET very similar to the IN–LEDGF/p75 complex. These findings provide new insights into the structural organization of IN subunits in functional concerted integration intermediates and suggest that differential multimerization of IN in the presence of various ligands could be exploited as a plausible therapeutic target for development of allosteric inhibitors.

INTRODUCTION

HIV-1 integrase (IN) functions as a multimer to catalyze integration of the reverse transcribed DNA copy of the viral genome into a host chromosome [reviewed in(1)]. The enzyme stably associates with two viral DNA ends to form a large nucleoprotein complex termed pre-integration complex (PIC), which also contains a number of viral and cellular proteins contributing to the retroviral integration (2–16). Quantities of PICs extracted from the infected cells are not sufficient to perform detailed structural or even lower resolution biophysical analyses. Therefore, purified recombinant IN and model DNA substrates have been employed to better understand mechanistic and structural foundations for the retroviral integration.

Notably, recent *in vitro* studies (17,18) defined key concerted integration intermediates and provided a powerful model system closely mimicking IN interactions with viral DNA within PICs in the infected cells. The step-wise interactions between IN, viral and target DNAs proceed through formation of highly stable nucleoprotein complexes. First, a tetramer of IN associates with a pair of viral DNA ends to form stable synaptic complexes (SSC). In common with PICs isolated from infected cells, the SSCs assembled *in vitro* are resistant to treatments with high ionic strength buffers containing 1 M NaCl. The 3′-processing reaction takes place within the SSC. IN remains stably associated with the pair of viral DNA ends after capture of a target DNA and DNA strand transfer. This second stable complex is termed the strand transfer complex.

*To whom correspondence should be addressed. Tel: +1 614 292 6091; Fax: +1 614 292 7766; Email: kvaratskhelia.1@osu.edu

HIV-1 IN is comprised of three structurally and mechanistically distinct domains including the N-terminal domain (NTD) which coordinates a Zn^{2+} ion, the catalytic core domain (CCD), which contains the catalytic DDE motif and the highly basic C-terminal domain (CTD). Each of these domains contribute to functional multimerization of IN (19–22). In the absence of cognate DNA, the full-length protein can form monomers, dimers, tetramers or higher order oligomers and the relative abundance of these species depends on the protein concentration and solution conditions (20,23–25). Structural studies with full-length HIV-1 IN or its complex with cognate DNA have not been successful presumably due to limited protein solubility and inherent flexibility of the three-domain protein. Instead, atomic structures for individual protein domains have been determined (21,22,26–30), which paved the way for molecular modeling of IN–DNA interactions [see (31) for recent review]. Most recently, the co-crystal structure of prototype foamy virus (PFV) IN with cognate DNA (32) has been exploited to build a model for a tetramer of HIV-1 IN interacting with two viral DNA ends (33). In this model, two of the four IN subunits directly bind DNA. The other two protomers seem to play a supporting role and contribute to IN multimerization.

The main cellular binding partner of HIV-1 IN is the protein known as lens epithelium-derived growth factor (LEDGF/p75). LEDGF/p75 knockdown and knockout experiments revealed the importance of this cellular cofactor for effective HIV-1 integration and viral replication (3,34–36). LEDGF/p75 primarily functions during HIV-1 infection to tether PICs to active genes during integration (15,37). The cellular protein directly interacts with HIV-1 IN via its C-terminal domain, which is termed the integrase binding domain (IBD) (38,39). The N-terminal part of LEDGF/p75, which contains a PWWP domain, nuclear localization signal and dual copy of the AT-hook DNA binding motif [reviewed in (40)] tightly associates with the chromatin.

In vitro functional assays have indicated that LEDGF/p75 differentially modulates HIV-1 IN activities (41–43). The cellular protein markedly enhances integration of single viral DNA into the target DNA regardless of the order of addition of proteins and DNA substrates in the reaction mixture. In contrast, the efficiency of the biologically relevant concerted integration of two viral DNA ends strongly depends on temporal interactions of viral DNA and LEDGF/p75 with HIV-1 IN (41–43). Addition of substoichiometric concentrations of LEDGF/p75 to the preassembled IN–viral DNA complex stimulates the pair-wise integration of viral DNA substrates. The preformed IN–LEDGF/p75 complex is impaired for the biologically essential concerted integration, while still stimulating single-end integration reactions (41–43).

Biochemical analysis of HIV-1 IN interactions with LEDGF/p75 indicated that the cellular cofactor promotes IN tetramerization (41). While IN in the functional SSC is also a tetramer, mass spectrometry (MS)-based protein footprinting of IN–viral DNA and IN–LEDGF/p75 complexes uncovered the following important differences. HIV-1 IN undergoes significant conformational change upon binding with cognate DNA involving the flexible connecting segment between the CCD and the CTD (44). In contrast, such changes have not been observed upon LEDGF/p75 binding to IN (41). Taken together, these observations led us to hypothesize that tetrameric forms of IN in complex with viral DNA and LEDGF/p75 are not identical, and that LEDGF/p75 binding modulates the structures of free IN and the IN–viral DNA complex. To test this premise, we monitored how individual IN subunits are organized in their complexes with viral DNA and LEDGF/p75 using Förster Resonance Energy Transfer (FRET). Our findings reveal two distinct IN tetramers formed in the presence of viral DNA or LEDGF/p75.

MATERIALS AND METHODS

Preparation of recombinant proteins and DNA substrates

Full-length wild-type and mutated (C56S/C65S) integrase proteins were expressed in *Escherichia coli*. The point mutations were introduced in the wild-type IN sequence using a QuikChange Site-Directed Mutagenesis kit (Stratagene, CA, USA). Wild-type and mutant full-length IN proteins were purified as described previously (41,45). The isolated CTD (213–288) was prepared similarly to the reported protocol (30). Purified recombinant LEDGF/p75 and LEDGF/IBD were obtained as described previously (14,38,46). The blunt-end viral DNA substrate (~1 kb) was obtained by PCR using Phusion polymerase (New England Biolabs, MA, USA), pU3U5 plasmid (47) and the following primers: dra3 (5-GATGGTTCACGTAGT GGGCC-3) and u5r (5-ACTGCTAGAGATTTTCCACA CTG-3). Viral DNA was further purified by agarose gel prior to its use in our assays. Double-stranded DNA (3 kb) derived from pGEM (Promega, WI, USA) was used as a non-specific DNA for control experiments.

Scaling-up of SSC preparations

The reported protocol (48) for the assembly of SSCs was modified as follows to allow 100-fold scale-up. To trap the SSC and prevent formation of the strand transfer complex purified wild-type IN or the mutant (C56/65S) protein labeled with fluorophore were incubated with 80 nM 1-kb blunt-ended viral DNA and 10 μM IN strand transfer inhibitor 118-D-24 (49) obtained from the the NIH AIDS Research and Reference Reagent Program. The reaction buffer contained 20 mM HEPES, pH 7.5, 12% DMSO, 5 mM DTT, 10% PEG 6000, 10 mM $MgCl_2$, 20 μM $ZnCl_2$ and 100 mM NaCl. The reaction mixture was pre-incubated on ice for 0.5 h and then transferred to 37°C for 2 h. The mixture was centrifuged at 10 000*g* for 40 min, the supernatant was removed and the pellet was resuspended in the suspension buffer (20 μl of 20 mM HEPES, pH 7.5, containing 1 M NaCl). The solution was then loaded onto a Quantum Prep PCR Kleen Spin Column (Bio-Rad, Hercules, CA, USA) pre-equilibrated with the suspension buffer. SSCs were eluted by centrifugation at 735*g* for 2 min. Freshly prepared SSCs were used immediately for their biophysical characterization. Purities of SSCs were examined with

gel electrophoresis through a 0.8% agarose–TBE-1 M urea gel in a TBE buffer containing 1 M urea as described previously (18). This concentration of urea in the gel and running buffer reduces interaction of SSCs with agarose and improves resolution without dissociating the nucleoprotein complexes (18). SSCs and free DNA were visualized by ethidium bromide staining. IN and LEDGF/p75 proteins associated with SSCs were examined by SDS–PAGE and visualized by western blot using monoclonal antibodies against IN (8G4) (50) from the NIH AIDS Research and Reference Reagent Program and against human LEDGF from BD Biosciences.

Size exclusion chromatography

Experiments were performed using HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) at 1 ml/min in buffer containing 50 mM HEPES (pH 7.4), 750 mM NaCl and 10% glycerol. The column was calibrated with the following proteins: bovine thyroglobulin (670 000 Da), bovine α-globulin (158 000 Da), chicken ovalbumin (44 000 Da), horse myoglobin (17 000 Da), vitamin B12 (1350 Da). These proteins were detected by absorbance at 280 nm.

Site-selective labeling of HIV-1 IN with fluorophores

We used the commercially available Alexa Fluor 488 and 568 (Invitrogen, Carlsbad, CA, USA) as a donor (D) and acceptor (A) pair with a Förster distance (R_0) of 62 Å. In parallel preparations, 2 μM of Alexa 488 or Alexa 568 maleimide were incubated with 8 μM mutant (C56/65 S) protein for 60 min at room temperature in a buffer containing 50 mM HEPES, pH 7.4, 1 M NaCl. The fluorophore/protein molar ratio of 1:4 was optimal to effectively label IN and avoid protein precipitation. Reactions were quenched by addition of 2 mM β-mercaptoethanol. Labeled proteins were then extensively dialyzed at 4°C against a buffer containing 50 mM HEPES, pH 7.4, 1 M NaCl, 7.5 mM CHAPS and 2 mM β-mercaptoethanol. A degree of labeling of 25 ± 2% was achieved in all reactions as determined by measuring absorbance of the dye (Alexa 488 at 493 nm and Alexa 568 at 575 nm) and the protein at 280 nm. The following extinction coefficients were used: $\varepsilon_{493} = 72\,000$ for Alexa 488, $\varepsilon_{575} = 92\,000$ for Alexa 568 and $\varepsilon_{280} = 50800$ for IN(C56/65 S).

Steady-state FRET measurement

Fluorescence was recorded at 25°C using a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, CA, USA). The assay buffer contained 20 mM HEPES, pH 7.5, 1 M NaCl and 1 mM DTT. The donor fluorescence was excited at 488 nm and fluorescence intensities were monitored in the range of 500–650 nm. Two fluorescence spectra were obtained for each set of experiments: (i) the D alone, where IN(C56/C65S) labeled with Alexa 488 was mixed with the unlabeled protein and (ii) the D–A pair, where two preparations of IN(C56/C65S), one labeled with Alexa 488 and another with Alexa 568 were mixed.

Time-resolved FRET measurement

Fluorescence decays were measured by time-correlated single-photon counting using laser excitation (LifeSpec-red, Edinburgh Instruments). The donor fluorescence was excited at 467 nm and the emission was collected at 520 nm with a polarizer at the magic angle (55°). All measurements were performed at 25°C using Versafluo cuvettes (BioRad, Hercules, CA, USA) with a 3×3 mm path length. Two fluorescence decay curves were obtained for each set of energy transfer experiments: (i) the D alone, where IN(C56/C65S) labeled with Alexa 488 was mixed with the unlabeled protein and (ii) the D–A pair, where two preparations of IN(C56/C65S) one labeled with Alexa 488 and another with Alexa 568 were mixed. The fluorescence intensity decays were deconvolved assuming a sum of exponentials:

$$I(t) = \sum_i a_i \exp\left(-\frac{t}{\tau_i}\right)$$

where τ_i and α_i are lifetime components and their relative amplitudes, respectively. The goodness of the fit was judged by reduced chi-square and autocorrelation function of weighted residuals. The amplitude-weighted lifetime $<\tau>$ for the donor only or the donor and acceptor pair was calculated by:

$$\langle\tau\rangle = \frac{\sum_1^n a_i\tau_i}{\sum_1^n a_i}$$

The average energy transfer efficiency E was calculated as follows:

$$E = 1 - \frac{\langle\tau_{\mathrm{DA}}\rangle}{\langle\tau_D\rangle}$$

where $<\tau_D>$ and $<\tau_{DA}>$ are the amplitude-averaged excited state lifetimes of the donor in the absence and presence of an acceptor, respectively.

The average distance r between the donor and acceptor was calculated by:

$$r = R_0\sqrt[6]{\frac{\langle\tau_D\rangle}{\langle\tau_D\rangle - \langle\tau_{\mathrm{DA}}\rangle} - 1}$$

where R_0 is the constant Förster distance (62 Å for the Alexa 488–Alexa 568 pair). Time-resolved (tr) fluorescence anisotropies of IN(C280-D) in its free form and complexed with LEDGF/p75 or viral DNA yielded fast rotational correlation times (see 'Results' section and Supplementary data) suggesting that the fluorophore tethered to IN exhibited significant degree of free motion. Therefore, in our calculations a dipole orientation factor of 2/3 was assumed.

Molecular modeling

Structures of individual IN domains were obtained from the two domain structures of HIV-1 IN (21,22). The CCDs were used as the common alignment feature and the CCD–CCD interfaces were maintained throughout

our modeling studies due to their essential role in IN multimerization and IN–LEDGF/p75 binding. To generate a molecular model for the HIV-1 IN–viral DNA complex, we used the structure of the PFV intasome (32) as the scaffold. We then employed our FRET results to position CTDs in the supporting subunits, which are absent in the reported crystal structure (32). Next, LEDGF/IBD was docked onto the SSC by establishing its interactions with the CCDs of one dimer and the NTD of another dimer similar to the published structures (51,52). For these studies, we used the online homology modeling server 'SWISSModel' (53) and the software 'Modeller' (54).

To build the full-length HIV-1 IN complex with LEDGF/IBD we analyzed available crystal structures of the two domain fragments of HIV-2 and maedi-visna virus (MVV) INs in the complex with LEDGF/IBD (52,55). The cellular protein interacted with the HIV-2 NTD–CCD dimer (55), while it stabilized a tetrameric form of the MVV NTD–CCD (52). Our published biochemical data (41) have shown that full-length LEDGF/p75 promotes HIV-1 IN tetramerization. Therefore, in our model LEDGF/IBD (Figure 7B) bridges between the two dimers in agreement with our MS footpriniting data (41) and the MVV structure (52). To extend these studies by including CTDs, which are missing in the published two domain structures (52,55), we considered our FRET data for modeling of the full-length IN tetramer. The molecular shape and global dimensions for the full-length IN tetramer in the complex with LEDGF/IBD obtained by small angle X-ray scattering (SAXS) experiments (56) were also employed as an additional constraint for our modeling experiments. The loops connecting IN domains were refined with the 'loop refinement' function in 'Modeller' software. The 'COOT' software (57) was used for building the model, resolving local clashes and optimizing the side chain conformations of the protein. The 'minimization' function of an Insight II software package (Accelrys Inc., CA, USA) was used to refine final models.

RESULTS

Interactions between IN and LEDGF/p75 can be monitored at sufficiently high concentrations of these proteins for their biochemical and biophysical analysis. In contrast, SSCs have previously been productively assembled only at very dilute concentrations of donor DNA and IN, which allowed biochemical characterization of concerted integration intermediates (17,18), but were not sufficient for biophysical studies. Therefore, our initial experiments focused on scaling-up purification of SSCs. We estimated that ~100-fold scale-up of the previously described preparations was necessary to enable FRET experiments.

Our experimental strategy for preparation of SSCs is depicted in Figure 1A. Reported reaction conditions for assembly of SSCs (17,18,48) served as a starting point. Recombinant IN and a long viral donor DNA in combination with IN strand transfer inhibitor 118-D-24 (49) allowed us to effectively trap the SSC. The inhibitor effectively impairs binding of the target DNA to the SSC and prevents formation of the strand transfer complex (18,58). To increase the yield of SSCs, reaction volumes were increased 20-fold from 25 µl used in previous studies (17,18,48) to 500 µl in our assays. Using the same buffer conditions as previously reported (17,18,48) IN concentrations in the reaction mixture were optimized as shown in Figure 1B.

Under these conditions, IN forms both specific and non-specific complexes with donor DNAs. To delineate between these, we exploited the intrinsic property of SSCs, which unlike non-specific IN–DNA interactions, are resistant to high ionic strength conditions. The mixture was subjected to treatments with 1 M NaCl followed by size exclusion spin column chromatography. SSCs and free DNAs were readily eluted from the column due to their large molecular weights, while free IN, which dissociated from non-specific DNA sites under high ionic strength conditions, was retained by the column. The obtained fractions were analyzed by SDS–PAGE to monitor relative quantities of IN in the complexes and by non-denaturing agarose gel electrophoresis to determine the purity of the final products.

Figure 1B shows that the optimal concentration range of IN for the assembly of SSCs under these conditions is 200–400 nM. At these concentrations, free IN is predominantly a dimer (59,60) (see also Supplementary Figure S1 and Supplementary Table S1). At higher protein concentrations, IN forms tetramers (~2 µM) with subsequent concentration increments leading to formation of higher order oligomers and protein precipitation. To delineate the role of IN tetramers in SSC formation, we compared the data in Figure 1B with IN 3′-processing activities (Supplementary Figure S2) as pre-assembled IN tetramers are active in this reaction (41). These experiments revealed a sharp contrast between the 3′-processing activities and the formation of SSCs. The highest 3′-processing activities were detected with 800–1600 nM IN, whereas these protein concentrations were very ineffective for the SSC assembly. These results are consistent with our earlier observations that the highly dynamic interplay between individual IN subunits is essential for productive concerted integration, and that a preformed IN tetramer lacks sufficient flexibility to form the fully functional nucleoprotein complex (41).

Figure 1C compares IN interactions with cognate donor and non-specific DNAs. In line with the previous report (17) IN formed SSCs resistant to 1 M NaCl treatment only with cognate DNA substrate and not with non-specific DNA (compare lanes 6 and 7 with lanes 4 and 5 in Figure 1C).

Non-denaturing agarose gel electrophoresis results (Figure 1D) demonstrate successful scale-up of SSCs (compare lanes 3–5). In the 100-fold scale-up (lane 5), the band corresponding to the SSC was readily detectable with ethidium bromide staining with only residual amounts of dimerized SSCs being observed. These optimized preparations of SSCs were employed for further biophysical analysis.

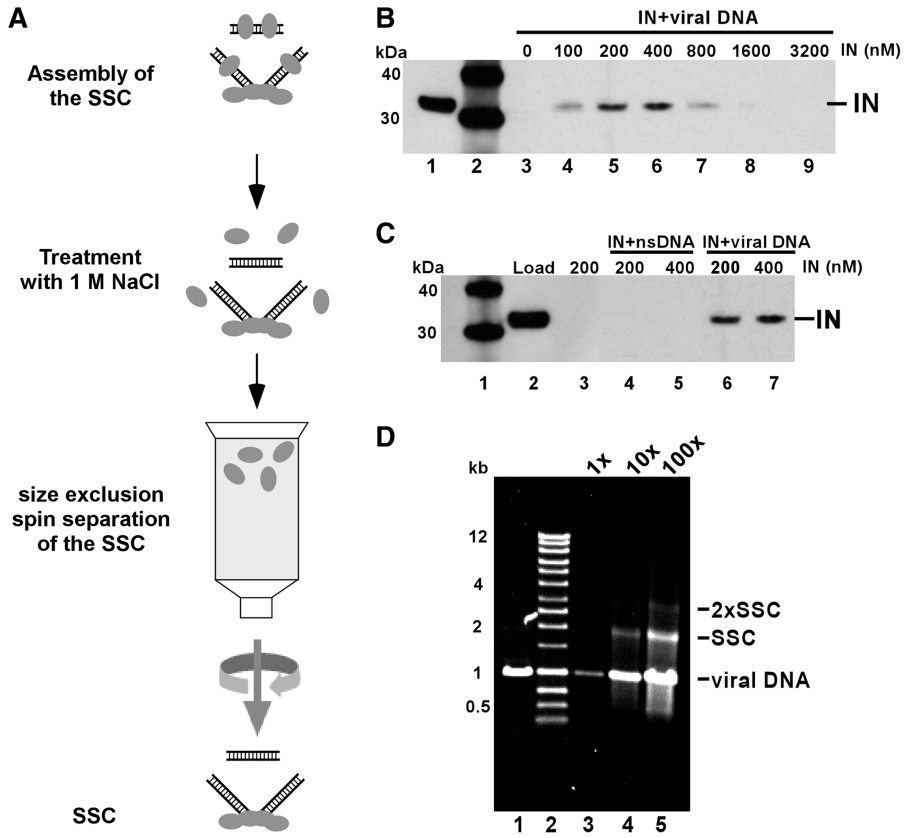

Figure 1. Scaled-up preparations of the SSC. (**A**) Experimental design. (**B**) Optimization of IN concentrations for the SSC assembly. Purified SSCs were subjected to SDS–PAGE and the IN band was visualized by western blot. Lane 1, IN load; lane 2, protein markers: MagicMark XP Western Protein Standard (Invitrogen, Carlsbad, CA, USA); lanes 3–9, SSCs assembled with increasing IN concentrations. At the optimal protein concentrations (200–400 nM), ~20% of total IN was assembled in the SSC. (**C**) Comparison of HIV-1 IN interactions with specific and non-specific DNA (nsDNA). Lane 1, protein markers: MagicMark XP Western Protein Standard (Invitrogen, Carlsbad, CA, USA); lane 2, IN load; lane 3, no DNA was included in the reaction mixture; lanes 4 and 5, SSC assembly with nsDNA; lanes 6 and 7, SSC assembly with viral DNA. (**D**) Agarose gel electrophoresis: Lane 1, viral DNA alone; lane 2, DNA markers: GeneMate Quanti-Marker 1 kb (BioExpress, Kaysville, UT, USA); lane 3, the initial (1×) scale for SSC preparations as reported previously (17); lane 4, 10-fold scale-up; lane 5, 100-fold scale-up.

Previous reports (42,43) indicated the importance of the order of viral DNA and LEDGF/p75 addition to IN for effective concerted integration. Particularly puzzling has been the observation that the preformed IN–LEDGF/p75 complex is selectively defective for concerted integration (41–43). Moreover, the mechanism behind these observations has remained obscure. LEDGF/p75 exhibits dual activities, with its N-terminal domain tightly binding DNA and its C-terminal IBD directly interacting with HIV-1 IN. Each of these properties of the full-length protein could potentially affect SSC formation by different mechanisms. For example, we previously demonstrated that increasing concentrations of LEDGF/p75 effectively competed with HIV-1 IN for viral DNA binding and inhibited the 3′-processing reaction (41). In contrast, LEDGF/IBD strongly modulated dynamic interplay between individual IN subunits and stimulated the 3′-processing reaction but potently impaired concerted integration (41).

To differentiate between these two activities of LEDGF/p75 and to examine how its direct interaction with IN could affect the formation of the SSC, we employed both, the full-length protein and LEDGF/IBD in our studies. Addition of LEDGF/p75 to free IN with subsequent exposure of protein–protein complexes to donor DNA effectively impaired formation of SSCs (Figure 2A, lane 5). Agarose gel electrophoresis results (Figure 2B) corroborated with the western blot data. No SSCs were observed when viral DNA was exposed to preformed IN complexes with LEDGF/p75 (Figure 2B, lane 4). Very similar results were obtained when the above experiments were conducted with LEDGF/IBD instead of the full-length protein (data not shown). Therefore, we conclude that direct interactions of LEDGF/p75 with HIV-1 IN modulate the structure of the retroviral enzyme in a way that impairs formation of the SSC.

We next examined whether under our reaction conditions, LEDGF/p75 associated with the SSC (Figure 2C and D). The SSCs prepared according to Figure 1A were incubated with LEDGF/p75 in a binding buffer containing 750 mM NaCl to prevent non-specific association of the full-length cellular protein with viral DNA. While LEDGF/p75 potently binds DNA in low ionic strength buffers, these interactions are inhibited at NaCl concentrations >200 mM (61). Indeed, no binding of

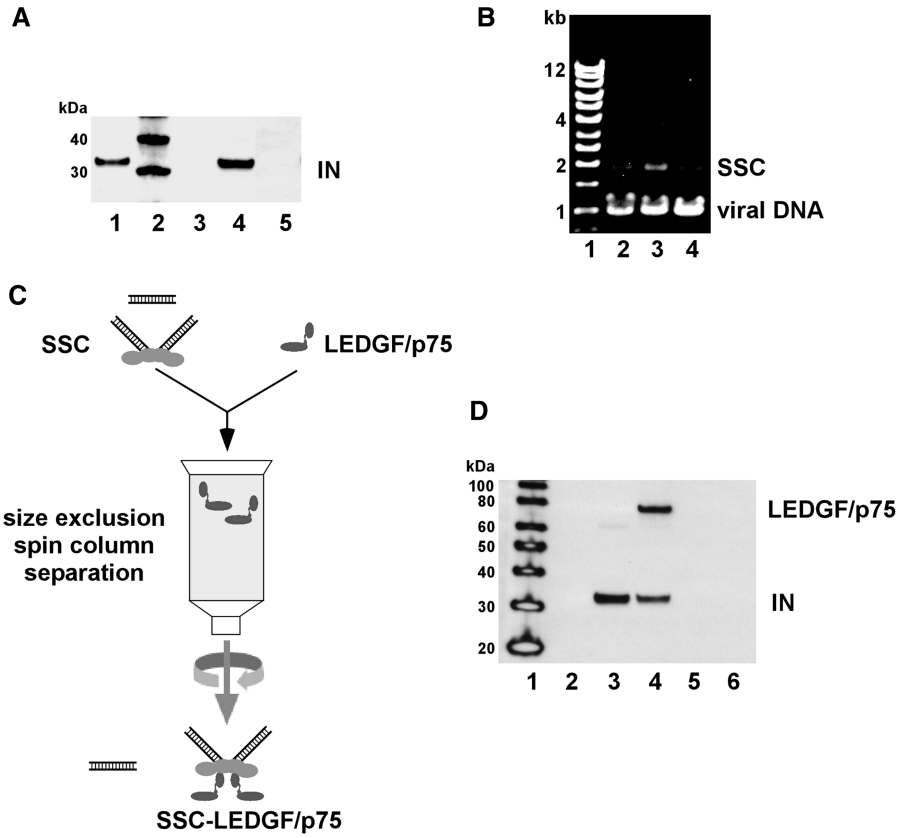

Figure 2. Effects of the order of viral DNA and LEDGF/p75 additions to HIV-1 IN on the SSC assembly. (**A**) SDS–PAGE analysis of SSCs. Lane 1: 1/10 of IN load, lane 2: protein markers: MagicMark XP Western Protein Standard (Invitrogen, Carlsbad, CA, USA), lane 3: no DNA was included in the reaction mixture, lane 4: the SSC assembly with IN and viral DNA, lane 5: LEDGF/p75 was pre-incubated with IN and then viral DNA was added to the reaction. IN was visualized by western blot using the respective antibody as described in 'Materials and Methods' section. (**B**) Non-denaturing agarose gel electrophoresis. Lane 1, DNA markers: GeneMate Quanti-Marker 1 kb (BioExpress, Kaysville, UT, USA); lane 2, no IN was included in the reaction mixture; lane 3, the SSC assembly with IN and viral DNA; lane 4, LEDGF/p75 was pre-incubated with IN and then viral DNA was added to the reaction. Free DNA and the SSC were visualized by ethidium bromide staining. (**C**) Experimental design to probe LEDGF/p75 interactions with the SSC. (**D**) SDS–PAGE analysis of LEDGF/p75 interactions with the SSC. Lane 1, protein markers: MagicMark XP Western Protein Standard (Invitrogen, Carlsbad, CA, USA); lanes 2–6, the following samples were incubated in the buffer containing 750 mM NaCl and then subjected to size exclusion chromatography as shown in (C): IN alone (lane 2), the purified SSC (lane 3), LEDGF/p75 plus the SSC (lane 4), LEDGF/p75 alone (lane 5), LEDGF/p75 plus viral DNA (lane 6). IN and LEDGF/p75 were visualized by western blot using respective antibodies as described in 'Materials and Methods' section.

LEDGF/p75 with viral DNA was detected in our experiments (Figure 2D, lane 6). In contrast, 750 mM NaCl did not significantly interfere with LEDGF/p75 binding to IN (see Supplementary Figure S1) and the cellular protein effectively interacted with the SSC (Figure 2C, lane 4). Collectively, these results indicate that under our reaction conditions, LEDGF/p75 associated with the SSC through its biologically relevant interactions with IN.

To gain structural insight into how LEDGF/p75 affects IN conformations, we employed protein–protein FRET. The experimental strategy for FRET studies is outlined in Figure 3. Two separate preparations of IN were used: one labeled with the D and the other with the A fluorophores. The IN concentration range of 200–400 nM was employed, which is optimal for assembly of the SSC (Figure 1B). At these concentrations, unliganded IN is predominantly a dimer (59,60) (see also Supplementary Figure S1). Upon binding to viral DNA ends, two separate dimers of IN assemble into a tetramer (17). LEDGF/p75 also promotes IN tetramerization (41,62) (see also Supplementary Figure S1). Assembly of two dimers labeled with D and A fluorophores into tetramers in the presence of viral DNA or LEDGF/p75 is expected to yield a FRET signal. The goal of our experiments was to compare average FRET values for IN–viral DNA and IN–LEDGF/p75 complexes and thereby determine whether IN tetramers formed in these complexes differed from one another.

A crucial step for effective FRET experiments is to site-selectively tether D and A dyes to IN preparations. The Alexa fluorophores chosen for these studies contain reactive maleimide groups enabling covalent attachment to surface Cys residues. Wild-type HIV-1 IN contains 6 Cys residues that present a challenge for site-specific labeling. Of these, C40 and C43 coordinate the structural Zn^{2+} ion and C130 is partially buried in the structure and not surface accessible. Therefore, these residues were expected to be chemically inert. In contrast, C56, C65 and C280 are surface exposed and could readily react with maleimide groups. Published site-directed

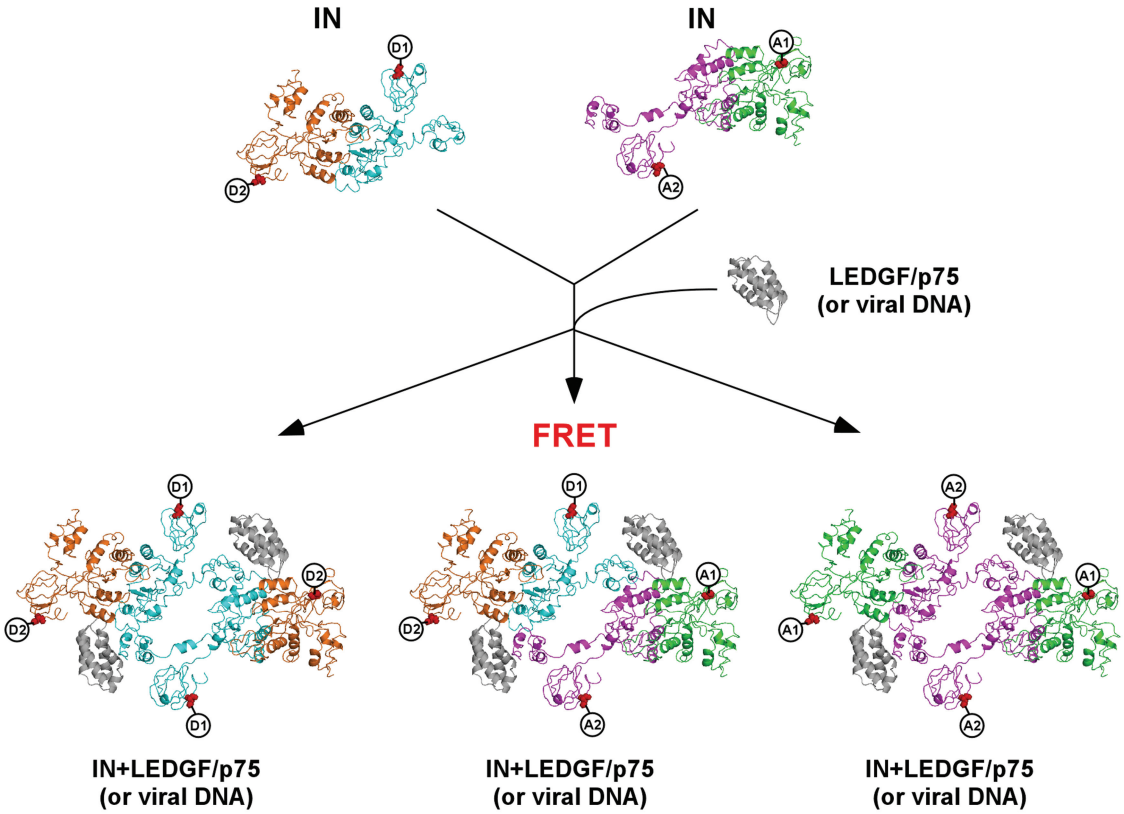

Figure 3. Scheme illustrating design of protein–protein FRET experiments. Two IN proteins are prepared in parallel: one labeled with the D probe and another with the A probe. The protein concentration range in the reaction mixture is 200–400 nM, where free IN is predominantly a dimer. Two sites (A1/A2 and D1/D2) are labeled in each dimer. Two IN preparations are mixed in ice-cold buffer to minimize the subunit exchange between free dimeric IN proteins. Subsequent addition of viral DNA or LEDGF/p75 promotes IN tetramerization. Three different populations of IN–viral DNA or IN–LEDGF/p75 complexes are formed. Of these, only the complex containing D1–D2 and A1–A2 pairs yields FRET. For the cartoon, a molecular model of full-length HIV-1 IN in complex with LEDGF/IBD was employed.

mutagenesis studies (63,64) showed that each of the three surface cysteines could individually be mutated to Ser without significantly compromising IN catalytic activities *in vitro* or viral replication in infected cells. In addition, we also noted that C56 and C65 are proximal to the viral DNA binding channel (33), and the placement of fluorophores at these locations could potentially interfere with protein–DNA interactions. In contrast, C280 is significantly removed from both viral DNA and LEDGF/p75 binding sites. Therefore, to accomplish the selective tethering of fluorophores, we mutated C56 and C65 to Ser and exploited the reactivity of the native C280 residue sulfhydryl. The data presented in Figure 4A demonstrate that C280 was indeed specifically targeted by the dye. The C56/C65S variant was readily labeled by both dyes, while no reactivity was observed in the case of the triple C56/65/280 S variant. Importantly, the fluorophore labeled mutant protein used in our FRET studies readily interacted with LEDGF/p75, formed stable synaptic complexes (Figure 4B) and displayed concerted integration activity (Figure 4C).

Prior to proceeding with FRET measurements, we monitored time-resolved anisotropies of the D labeled IN variants, both free and complexed with viral DNA or LEDGF/p75 using time-correlated single photon counting. Time-resolved anisotropy curves were well-fit by a single exponential function. Rotational correlation times of 2.2, 2.7 and 2.6 ns were measured for IN alone, IN complexed with viral DNA and IN bound to LEDGF/p75, respectively. These fast rotational correlation times suggest that the probe retained a significant degree of free motion upon its tethering to C280. Importantly, very similar values were observed for IN–LEDGF/p75 and IN–viral DNA complexes (Supplementary Figure S3), indicating that neither LEDGF/p75 nor viral DNA significantly altered the conformational freedom of the probe on IN. These control experiments assured us that FRET values obtained for unliganded IN and its complexes with LEDGF/p75 and viral DNA can reliably be compared with one another.

To measure FRET between individual IN subunits in the context of various complexes, we initially conducted steady-state (ss) measurements, which reveal average FRET intensities. In the absence of a binding partner, the IN dimers mixed together did not exhibit any detectable FRET signal (Figure 5A). This was anticipated as at 200 nM concentrations of IN(C280-A) and IN(C280-D) in the reaction mixture, the protein was predominantly a dimer (Supplementary Figure S1). Furthermore, under such conditions we did not expect to observe significant

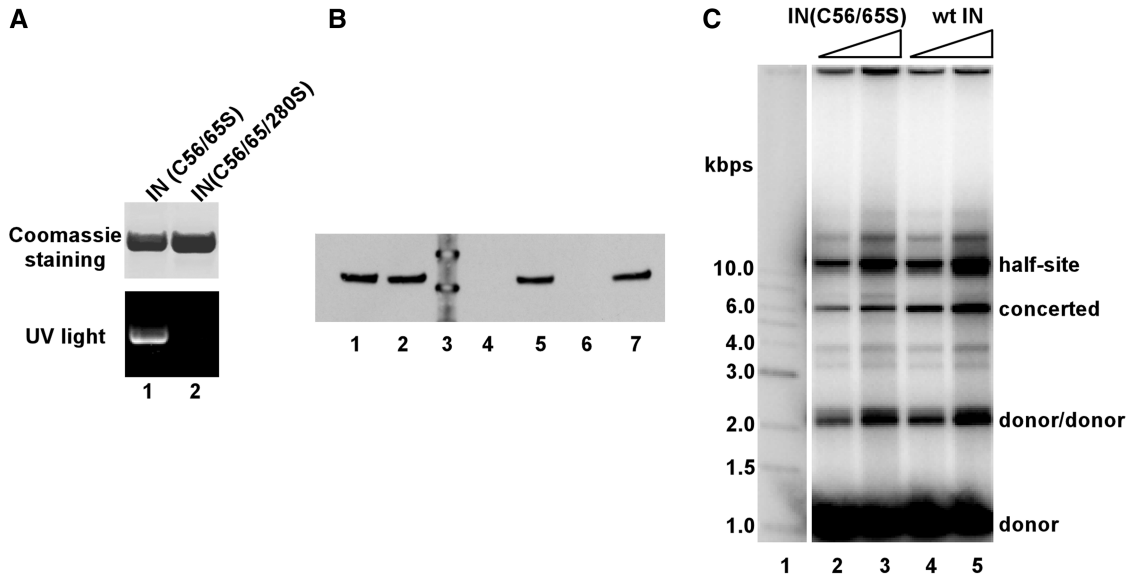

Figure 4. Site-selective labeling of HIV-1 IN with a fluorophore. (**A**) In parallel experiments, C56/65S and C56/65/280S mutants were subjected to treatment with Alexa 488 maleimide. The reactions were quenched with DTT and subjected to SDS–PAGE. Images of the same gel following coomassie staining (upper panel) and UV-light exposure (lower panel) are shown. No fluorescence signal was detected for the C56/65/280S protein (lane 2), while the C56/65S mutant (lane 1) was effectively labeled with the dye. (**B**) Assembly of SSCs with the labeled mutant IN. Lane 1, wild-type IN load; lane 2, load of the IN (C56/65S) mutant; lane 3, protein markers; lane 4, no DNA control; lane 5, the SSC assembly with wild-type IN and viral DNA; lane 6, the IN (C56/65S) mutant without DNA; lane 7, the SSC assembly with the IN (C56/65S) mutant. (**C**) Concerted integration assays of wild-type and mutant (C56/65S) IN proteins: Lane 1: DNA markers, lanes 2 and 3: increasing concentration of the IN (C56/65S) mutant, lanes 4 and 5: wild-type IN activities.

subunit exchange between two IN preparations for the following reasons. The dissociation constant for IN dimers has been reported to be in the subnanomolar range (59,60). Therefore, effective exchange of individual subunits between dimeric forms of IN can only be observed at subnanomolar to low nM protein concentrations (60,65). In contrast, IN forms stable dimers at the protein concentrations of 200–400 nM used here (Supplementary Figure S1). At the same time, this concentration range is low enough to avoid tetramer formation. IN tetramers and effective exchange of the stable dimers between tetrameric forms of IN can be detected at ~2 μM protein (41). Therefore, under our assay conditions, the background FRET due to subunit exchange was very minimal (Figure 5A).

Addition of LEDGF/p75 promoted formation of a tetramer by bridging two IN dimers (41) (see also Supplementary Figure S1) and resulted in significant energy transfer (Figure 5B). In parallel reactions, viral DNA was added to the IN(C280-A) and IN(C280-D) mixture to form the SSC (Figure 5C). While IN in the context of the SSC is also a tetramer, the IN–viral DNA complex displayed a significantly higher FRET than IN–LEDGF/p75 (see overlay of the spectra in Figure 5D). These results suggest that tetrameric forms of IN in IN–LEDGF/p75 and IN–viral DNA complexes differ.

We also examined the samples using tr-FRET (Figure 6 and Supplementary Figure S4). We noted that donor only controls, where IN(C280-D) was mixed with the unlabeled protein (Figure 6A) and then incubated with LEDGF/p75 (Figure 6B) or viral DNA (Figure 6D), yielded complex decay curves (Supplementary Figure S5) suggesting that the fluorophore tethered to IN adopts multiple conformations. Potential asymmetric arrangements of individual subunits within multimeric IN could contribute to this. Alternatively, local environment at C280 could allow the fluorophore to adopt multiple conformations. To delineate between these possibilities, we used the isolated CTD as a reliable control. Consistent with previous reports (29,30) isolated CTD formed dimers in our experiments as judged by size exclusion chromatography (data not shown). Each symmetrical subunit of this protein fragment contains a single Cys residue at the position corresponding to C280 in the full-length protein (29,30). Similarly to full-length IN, the isolated CTD also yielded three exponential decay curves. These findings suggest that the local environment at the tethering site contributes to multiple conformations of the fluorophore. Our results are reminiscent of the published (66) tr-FRET analysis of Trp residues in proteins indicating that surface tryptophans typically adopt different conformations and yield multi-exponential decays curves. In common with the donor alone, experiments analysis of D–A pairs (Figure 6 and Supplementary Figure S4) yielded the decay curves that were best fit to a three exponential decay (Supplementery Figure S5). No FRET was detected when IN(C280-A) was mixed with IN(C280-D) (Figure 6A). In contrast, these proteins yielded FRET when incubated with LEDGF/p75 or viral DNA (Figure 6B and D). Average distances calculated from tr-FRET results were ~81 Å for the IN–LEDGF/p75 complex and ~69 Å for the SSC (Table 1) indicating distinct IN conformations in these complexes.

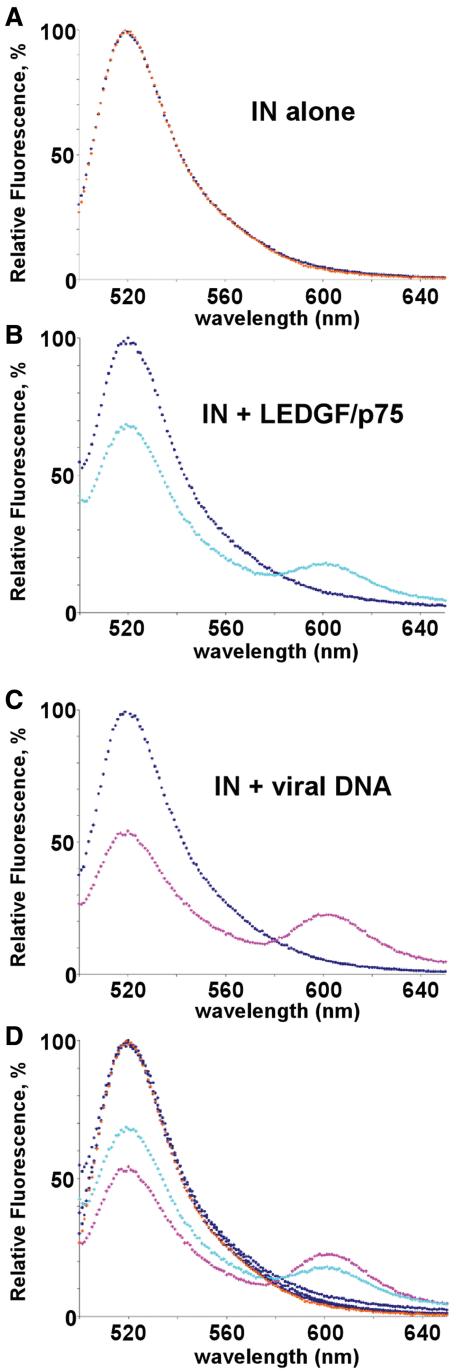

Figure 5. Ss-FRET plots for IN–viral DNA and IN–LEDGF/p75 complexes. (**A**) IN alone, (**B**) IN–LEDGF/p75 complex, (**C**) IN–viral DNA complex, (**D**) overlay of the spectra from experiments shown in panels A–C. For control experiments, Alexa 488-labeled IN(C56/65S) was mixed with the unlabeled protein (blue circles in A) and then incubated with LEDGF/p75 (blue circles in B) or viral DNA (blue circles in C). The fluorescence intensities for the D–A pairs are depicted with diamonds and color coded as follows: IN alone, orange; IN–LEDGF/p75 complex, cyan; IN–viral DNA complex, magenta. Fluorescence quenching at 520 nm and concomitant increase of emission intensities at 610 nm demonstrate FRET. The spectra was normalized by defining the maximum intensity of the donor fluorophore in each donor alone experiment as 100%.

We then extended the tr-FRET experiments to test more complex interactions involving IN, viral DNA and LEDGF/p75. The following two pathways for the assembly of large nucleoprotein complexes were considered. First, the IN–LEDGF/p75 complex was preformed and viral DNA was then added. The fluorescence decay profile for this complex was virtually identical to that for the IN–LEDGF/p75 complex (Figure 6 and Table 1). In the second set of experiments, we first obtained the SSC and then exposed it to LEDGF/p75. The tr-FRET data for this large nucleoprotein complex was very similar to that of the SSC (Figure 6 and Table 1). The above FRET experiments were also conducted with LEDGF/IBD and the data (see Supplementary Figure S4) closely mirrored those obtained with full-length LEDGF/p75 (Figure 6). Taken together, these data show that the conformation of IN tetramer depends on the order of ligand addition.

Our FRET results together with available crystallographic data were employed to generate molecular models for HIV-1 IN interactions with viral DNA and LEDGF/IBD (Figure 7 and Supplementary Figure S6). Figure 7A and B depict two separate conformations of IN tetramers. The model in Figure 7A was generated stepwise by first modeling HIV-1 IN interactions with viral DNA and then docking LEDGF/IBD into the nucleoprotein complex. Interactions between two functional HIV-1 IN subunits with viral DNA (Figure 7A) were modeled based on the crystal structure of PFV IN in the complex with cognate DNA (32) and the resulting nucleoprotein interactions were similar to those proposed recently (33). However, published studies (32,33) did not define NTDs and CTDs in the supporting two subunits. Our FRET data provided complementary information and enabled us to model interactions of the full-length IN tetramer with viral DNA ends (Figure 7A and Supplementary Figure S6). As shown in Supplementary Figure S6 (left panels), there are four possible D–A distances. Of these, distances between D2–A2 and D1–A1 pairs are identical due to the 2-fold symmetry between two IN dimers, while the distance between D2–A1 pairs exceeds an effective FRET range ($>2 \times R_0$) and would not affect our FRET measurements. Therefore, average FRET distances are likely to be derived from the following three distances D1–A2, D2–A2 and D1–A1. In fact, FRET measurements and molecular modeling results suggest that these three distances are very similar. For example, the D1–A2 distance of ~68 Å is in excellent agreement with the structure-based modeling using the PFV intasome as a template. This distance together with the average FRET distance measurement of ~69 Å for this complex (Table 1) and the limited length of the loop connecting the CCD and CTD, provided us with significant constraints to position CTDs in supporting subunits as shown in Figure 7A and Supplementary Figure S6 (left panel).

It should be noted that PFV IN does not interact with LEDGF/IBD and the published model of HIV-1 intasome (33) did not address its interactions with the key cellular cofactor. Our data show that LEDGF/p75 potently binds the SSC through its biologically relevant site on HIV-1 IN

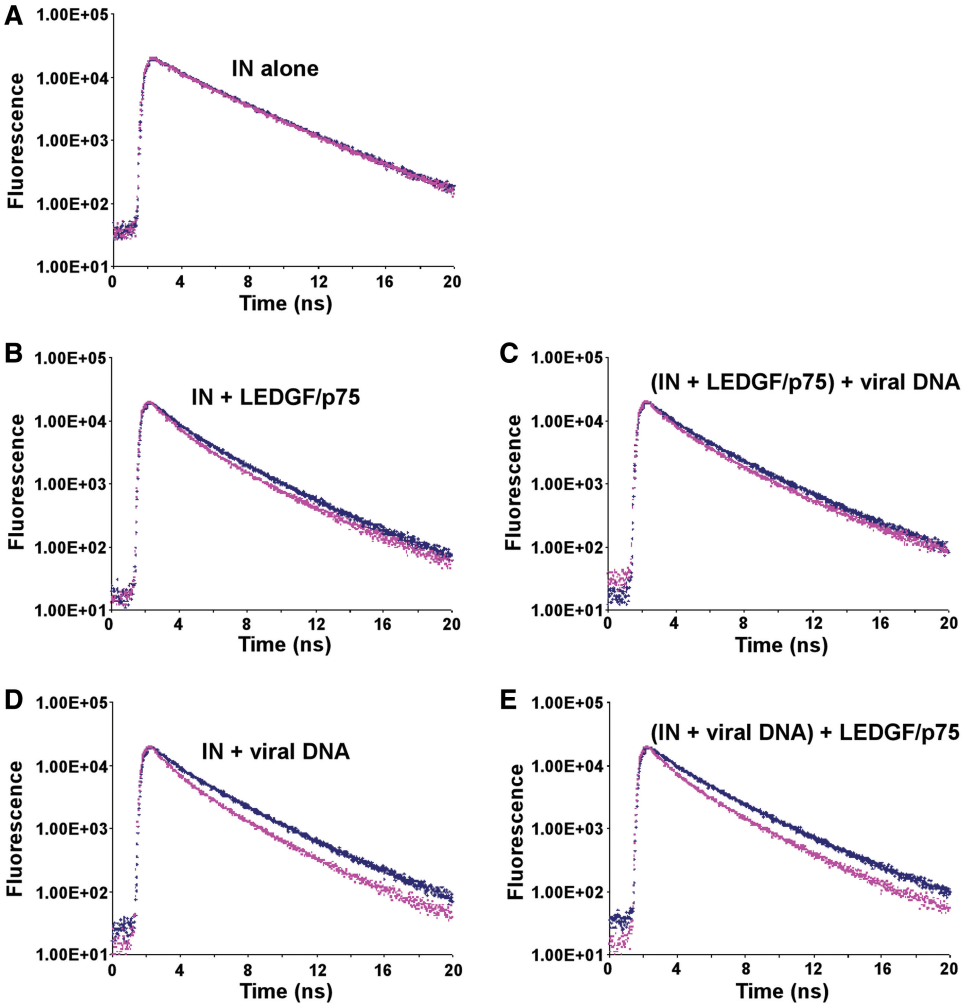

Figure 6. Tr-fluorescence decay plots for IN–viral DNA and IN–LEDGF/p75 complexes. (**A**) IN alone; (**B**) IN–LEDGF/p75 complex, (**C**) IN–LEDGF/p75 complex was preformed and then exposed to viral DNA; (**D**) IN–viral DNA complex; (**E**) IN–viral DNA complex was preformed and then LEDGF/p75 was added to the nucleoprotein complex. Blue plots show fluorescence decays for donor only control, where Alexa 488-labeled IN(C56/65S) was mixed with the unlabeled protein. Magenta plots show data for the D–A pairs.

Table 1. Tr-FRET measurements

	τ_D, ns	τ_{DA}, ns	E	r, Å
SSC	1.51 (0.03)	0.99 (0.03)	0.34 (0.02)	69.2 (0.9)
SSC+LEDGF/p75	1.65 (0.10)	1.11 (0.08)	0.33 (0.01)	70.1 (0.4)
IN+LEDGF/p75	1.50 (0.01)	1.23 (0.01)	0.18 (0.01)	80.0 (0.6)
(IN+LEDGF/p75)+ viral DNA	1.49 (0.01)	1.23 (0.01)	0.17 (0.01)	80.5 (0.6)

τ_D is the average excited state lifetime of the donor, τ_{DA} is the average excited state lifetime of the donor–acceptor pair, E is the average energy transfer efficiency, r is the average calculated distance. The number in parenthesis is standard deviation of three independent experiments.

(Figure 2D) and that the cellular protein does not alter the pre-formed architecture of the nucleoprotein complex (Figure 6 and Table 1). In complete agreement with these experimental data, we were able to dock two LEDGF/IBD molecules onto the SSC without altering the pre-existing IN–viral DNA interactions (Figure 7A). In particular, each LEDGF/IBD engages the CCDs of one dimer and establishes additional contacts with the NTD from another dimer, which ensures high-affinity binding of the cellular cofactor to the complex (41). Taken together, the experimental results presented here have allowed us to extend previous modeling of HIV-1 IN–viral DNA interactions (33) by building a ternary complex between viral DNA, full-length IN tetramer and LEDGF/IBD (Figure 7A).

Figure 7B shows an alternative conformation of the IN tetramer in complex with LEDGF/IBD. To create this model, we considered the following main criteria. Our published biochemical studies (41) have shown that interactions between the NTD and the CCD are important for IN tetramerization and high-affinity LEDGF/p75 interactions. In agreement with these findings, subsequent crystallographic studies (52) with the two domain construct of MVV IN have demonstrated that LEDGF/IBD bridges between the NTD of one dimer and the CTD of another dimer. Therefore, these interactions were included

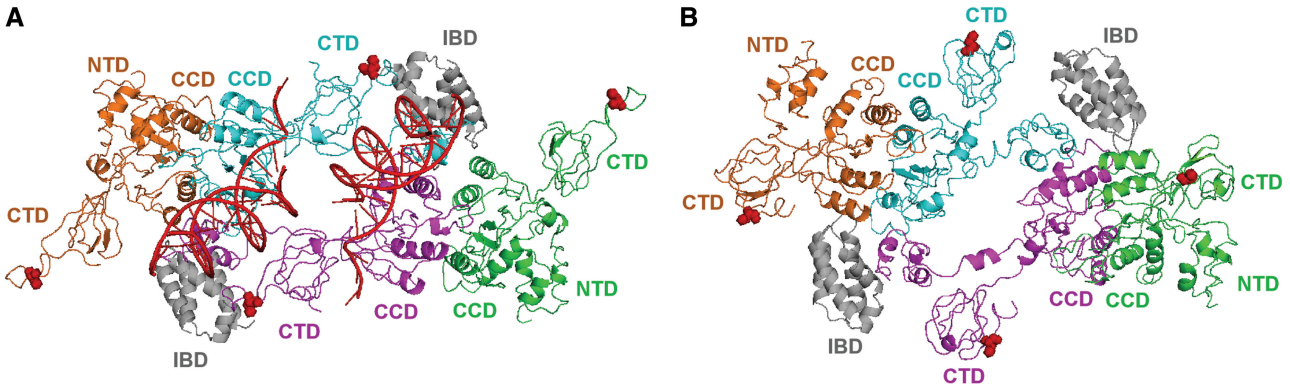

Figure 7. Molecular modeling of IN, viral DNA and LEDGF/IBD interactions. (**A**) The assembly of the fully functional nucleoprotein complex. First, IN interacts with viral DNA to form the SSC. Then, LEDGF/IBD tightly binds the SSC by bridging between the two dimers. Four individual subunits of IN are colored orange, cyan, magenta and green. The cyan and magenta protomers directly interact with viral DNA, while green and orange subunits play supporting roles. LEDGF/IBD is depicted in gray. (**B**) IN interactions with LEDGF/IBD. Colors for IN subunits and LEDGF/IBD are the same as in A. Locations of C280 in each subunit are shown by red spheres.

in our model. Published studies (41,52), however, did not address the positioning of the CTDs in the protein–protein complex. Therefore, our FRET data (Table 1) were employed to create the model between full-length IN and LEDGF/IBD (Figure 7 and Supplementary Figure S6). Additional constraints were provided by SAXS data (56), which revealed the molecular shape and global dimensions for the full-length HIV-1 IN complex with LEDGF/IBD. The resulting model and calculated distances for D–A pairs are given in Figure 7A and Supplementary Figure S6 (right panel).

DISCUSSION

Our FRET studies show that IN tetramers formed in IN–viral DNA and IN–LEDGF/p75 complexes are distinct. The order of addition experiments (Figure 6) further underscore the conclusion that there are different pathways for IN multimerization. Accordingly, we propose two separate models for IN tetramers (Figure 7). The most noticeable difference between these tetramers is differential positioning of CTDs. In the IN–viral DNA complex (Figure 7A), the two functional CTDs tightly interact with the ends of viral DNA and are positioned immediately adjacent to the CCDs. Our earlier mass spectrometry-based protein footprinting studies demonstrated that the CTD undergoes a significant conformational change upon formation of the IN–viral DNA complex, while such changes are not observed in the IN–LEDGF/p75 complex (44). Our FRET results show that the CTDs are positioned closer to each other in the presence of viral DNA than LEDGF/p75 (Figures 5 and 6; Table 1). We propose that such conformational flexibility of the CTDs is crucial for the effective assembly of the SSC. In contrast, the CTDs in the IN tetramer formed upon binding to LEDGF/IBD may not be sufficiently flexible to fully engage viral DNA in a manner that would ensure formation of the SSC. For example, bound LEDGF/IBD could indirectly limit repositioning of the CTD upon subsequent interactions of the protein–protein complex with viral DNA. At the same time, it should be noted that preformed IN–LEDGF/IBD can bind viral DNA to form a less stable nucleoprotein complex, which effectively catalyzes the 3′-processing reaction but fails to carry out the concerted integration (41). In line with these observations, both viral and target DNA can be modeled in the IN–LEDGF/IBD complex (data not shown). However, the DNA binding cleft in such a complex would differ from that observed with the SSC (Figure 7A). In other words, we propose that there are different modes of viral DNA binding to HIV-1 IN, and only the SSC is capable of productive concerted integration.

The biological relevance of our findings is corroborated by the following data. PICs isolated from LEDGF/p75 knockout cells exhibit normal levels of DNA strand transfer activity *in vitro*, suggesting that IN and viral DNA can effectively assemble in the cytoplasm of infected cells in the absence of the cellular cofactor (15). Furthermore, LEDGF/p75 knockout did not affect nuclear import of PICs (15). Instead, the cellular cofactor has been shown to bind PICs at a later stage in the nucleus and navigate them to active genes on the chromatin (15,37). While LEDGF/p75 is a predominantly nuclear protein, it has also been suggested that low endogenous amounts of the cellular cofactor in the cytoplasm could engage PICs (2). Our findings do not preclude this possibility but rather outline a necessary order of pre-integration events in which IN engages viral DNA ends prior to recruitment of LEDGF/p75 into the PIC.

An alternative sequence of events could be detrimental for HIV-1 integration. For example, overexpression of LEDGF/IBD that lacked a nuclear localization signal and therefore interacted with HIV-1 IN in the cytoplasm, effectively impaired retroviral integration (3). One possible explanation for this is a potential competition between LEDGF/IBD and endogenous LEDGF/p75. However, this mechanism alone cannot explain the findings that LEDGF/IBD was significantly more effective at suppressing HIV-1 replication in LEDGF/p75 knockdown cells compared with cells containing normal levels of the cellular cofactor (3). Instead, our FRET results suggest

that LEDGF/IBD can stabilize an alternative tetrameric conformation of IN, which is defective in concerted integration. This notion is further supported by the observation that overexpressed LEDGF/IBD stabilized the IN structure in infected cells and protected the retroviral protein from proteasomal degradation (67). It should also be noted that IN bound to LEDGF/IBD can still interact with viral DNA and carry out 3′-processing reactions (41). Moreover, LEDGF/IBD is not expected to interfere with the nuclear import of PICs as LEDGF/p75 is not involved in this process (15). In contrast, in line with *in vitro* observations (41), cell-based assays have shown that the IN complex with LEDGF/IBD is defective in concerted integration (3). Taken together, these studies indicate that modulation of IN structure prior to its binding to viral DNA is detrimental for retroviral integration.

Understanding of alternative pathways of IN multimerization is highly important for the development of allosteric inhibitors. We recently proposed (65) a mechanism for inhibiting HIV-1 IN that would mimic LEDGF/IBD effects, and reported the discovery of a small molecule that interacted with K173 at the IN dimer interface and stabilized an inactive multimeric conformation of the protein *in vitro*. More recently, allosteric inhibitors have been described that impair IN interactions with LEDGF/p75 in infected cells (68). These compounds bind at the IN dimer interface, occupying the LEDGF/IBD binding pocket. In general, detailed analysis of available structures of the HIV-1 IN CCD dimer revealed two separate adjacently positioned cavities suitable for small molecule binding (65). Moreover, several compounds have been reported to target these sites (69,70). Further studies in this direction are warranted as the HIV-1 IN multimer is an unexploited therapeutic target and allosteric drugs binding these sites are likely to be effective against HIV-1 strains resistant to current therapies.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGMENTS

We are grateful to Dr Christopher McKee for critical reading of the manuscript and helpful comments.

FUNDING

The National Institutes of Health (grants AI062520, AI081581 and AI077341 to M.K., AI077387 to K.M.-F.); the intramural research program of the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health and by the NIH AIDS Targeted Antiviral Program (to R.C.). Funding for open access charge: National Institutes of Health (grant AI062520).

Conflict of interest statement. None declared.

REFERENCES

1. Brown,P.O. (1997) Integration. In Coffin,J.M., Hughes,S.H. and Varmus,H.E. (eds), *Retroviruses*. Cold Spring Harbor Laboratory, Plainview, NY, pp. 161–204.
2. Llano,M., Vanegas,M., Fregoso,O., Saenz,D., Chung,S., Peretz,M. and Poeschla,E.M. (2004) LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J. Virol.*, **78**, 9524–9537.
3. Llano,M., Saenz,D.T., Meehan,A., Wongthida,P., Peretz,M., Walker,W.H., Teo,W. and Poeschla,E.M. (2006) An essential role for LEDGF/p75 in HIV integration. *Science*, **314**, 461–464.
4. Lewinski,M.K., Yamashita,M., Emerman,M., Ciuffi,A., Marshall,H., Crawford,G., Collins,F., Shinn,P., Leipzig,J., Hannenhalli,S. *et al.* (2006) Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathog.*, **2**, e60.
5. Farnet,C.M. and Haseltine,W.A. (1990) Integration of human immunodeficiency virus type 1 DNA *in vitro*. *Proc. Natl Acad. Sci. USA*, **87**, 4164–4168.
6. Bukrinsky,M.I., Sharova,N., McDonald,T.L., Pushkarskaya,T., Tarpley,W.G. and Stevenson,M. (1993) Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc. Natl Acad. Sci. USA*, **90**, 6125–6129.
7. Buckman,J.S., Bosche,W.J. and Gorelick,R.J. (2003) Human immunodeficiency virus type 1 nucleocapsid zn(2+) fingers are required for efficient reverse transcription, initial integration processes, and protection of newly synthesized viral DNA. *J. Virol.*, **77**, 1469–1480.
8. Carteau,S., Batson,S.C., Poljak,L., Mouscadet,J.F., de Rocquigny,H., Darlix,J.L., Roques,B.P., Kas,E. and Auclair,C. (1997) Human immunodeficiency virus type 1 nucleocapsid protein specifically stimulates Mg2+-dependent DNA integration *in vitro*. *J. Virol.*, **71**, 6225–6229.
9. Carteau,S., Gorelick,R.J. and Bushman,F.D. (1999) Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase *in vitro*: stimulation by the viral nucleocapsid protein. *J. Virol.*, **73**, 6670–6679.
10. Chen,H. and Engelman,A. (1998) The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc. Natl Acad. Sci. USA*, **95**, 15270–15274.
11. Lee,M.S. and Craigie,R. (1994) Protection of retroviral DNA from autointegration: involvement of a cellular factor. *Proc. Natl Acad. Sci. USA*, **91**, 9823–9827.
12. Lee,M.S. and Craigie,R. (1998) A previously unidentified host protein protects retroviral DNA from autointegration. *Proc. Natl Acad. Sci. USA*, **95**, 1528–1533.
13. Cherepanov,P., Maertens,G., Proost,P., Devreese,B., Van Beeumen,J., Engelborghs,Y., De Clercq,E. and Debyser,Z. (2003) HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem*, **278**, 372–381.
14. Cherepanov,P., Sun,Z.Y., Rahman,S., Maertens,G., Wagner,G. and Engelman,A. (2005) Solution structure of the HIV-1 integrase-binding domain in LEDGF/p75. *Nat. Struct. Mol. Biol.*, **12**, 526–532.
15. Shun,M.C., Raghavendra,N.K., Vandegraaff,N., Daigle,J.E., Hughes,S., Kellam,P., Cherepanov,P. and Engelman,A. (2007) LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev*, **21**, 1767–1778.
16. Engelman,A. (2006) Host cell factors and HIV-1 integration. *Future HIV Ther*, **1**, 415–426.
17. Li,M., Mizuuchi,M., Burke,T.R. Jr and Craigie,R. (2006) Retroviral DNA integration: reaction pathway and critical intermediates. *EMBO J*, **25**, 1295–1304.
18. Li,M. and Craigie,R. (2009) Nucleoprotein complex intermediates in HIV-1 integration. *Methods*, **47**, 237–242.
19. Andrake,M.D. and Skalka,A.M. (1995) Multimerization determinants reside in both the catalytic core and C terminus of avian sarcoma virus integrase. *J. Biol. Chem*, **270**, 29299–29306.
20. Jenkins,T.M., Engelman,A., Ghirlando,R. and Craigie,R. (1996) A soluble active mutant of HIV-1 integrase: involvement of both

the core and carboxyl-terminal domains in multimerization. *J. Biol. Chem.*, **271**, 7712–7718.
21. Chen,J.C., Krucinski,J., Miercke,L.J., Finer-Moore,J.S., Tang,A.H., Leavitt,A.D. and Stroud,R.M. (2000) Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc. Natl Acad. Sci. USA*, **97**, 8233–8238.
22. Wang,J.Y., Ling,H., Yang,W. and Craigie,R. (2001) Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J*, **20**, 7333–7343.
23. van Gent,D.C., Elgersma,Y., Bolk,M.W., Vink,C. and Plasterk,R.H. (1991) DNA binding properties of the integrase proteins of human immunodeficiency viruses types 1 and 2. *Nucleic Acids Res*, **19**, 3821–3827.
24. Vincent,K.A., Ellison,V., Chow,S.A. and Brown,P.O. (1993) Characterization of human immunodeficiency virus type 1 integrase expressed in Escherichia coli and analysis of variants with amino-terminal mutations. *J. Virol.*, **67**, 425–437.
25. Deprez,E., Tauc,P., Leh,H., Mouscadet,J.F., Auclair,C. and Brochon,J.C. (2000) Oligomeric states of the HIV-1 integrase as measured by time-resolved fluorescence anisotropy. *Biochemistry*, **39**, 9275–9284.
26. Cai,M., Zheng,R., Caffrey,M., Craigie,R., Clore,G.M. and Gronenborn,A.M. (1997) Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *Nat. Struct. Biol.*, **4**, 567–577.
27. Dyda,F., Hickman,A.B., Jenkins,T.M., Engelman,A., Craigie,R. and Davies,D.R. (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science*, **266**, 1981–1986.
28. Goldgur,Y., Dyda,F., Hickman,A.B., Jenkins,T.M., Craigie,R. and Davies,D.R. (1998) Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. *Proc. Natl Acad. Sci. USA*, **95**, 9150–9154.
29. Eijkelenboom,A.P., Lutzke,R.A., Boelens,R., Plasterk,R.H., Kaptein,R. and Hard,K. (1995) The DNA-binding domain of HIV-1 integrase has an SH3-like fold. *Nat. Struct. Biol.*, **2**, 807–810.
30. Lodi,P.J., Ernst,J.A., Kuszewski,J., Hickman,A.B., Engelman,A., Craigie,R., Clore,G.M. and Gronenborn,A.M. (1995) Solution structure of the DNA binding domain of HIV-1 integrase. *Biochemistry*, **34**, 9826–9833.
31. Kessl,J.J., McKee,C.J., Eidahl,J.O., Shkriabai,N., Katz,A. and Kvaratskhelia,M. (2009) HIV-1 Integrase-DNA Recognition Mechanisms. *Viruses*, **1**, 713–736.
32. Hare,S., Gupta,S.S., Valkov,E., Engelman,A. and Cherepanov,P. (2010) Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature*, **464**, 232–236.
33. Krishnan,L., Li,X., Naraharisetty,H.L., Hare,S., Cherepanov,P. and Engelman,A. (2010) Structure-based modeling of the functional HIV-1 intasome and its inhibition. *Proc. Natl Acad. Sci. USA*, **107**, 15910–15915.
34. De Rijck,J., Vandekerckhove,L., Gijsbers,R., Hombrouck,A., Hendrix,J., Vercammen,J., Engelborghs,Y., Christ,F. and Debyser,Z. (2006) Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. *J. Virol*, **80**, 11498–11509.
35. Vandekerckhove,L., Christ,F., Van Maele,B., De Rijck,J., Gijsbers,R., Van den Haute,C., Witvrouw,M. and Debyser,Z. (2006) Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. *J. Virol*, **80**, 1886–1896.
36. Hombrouck,A., De Rijck,J., Hendrix,J., Vandekerckhove,L., Voet,A., De Maeyer,M., Witvrouw,M., Engelborghs,Y., Christ,F., Gijsbers,R. *et al.* (2007) Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. *PLoS Pathog*, **3**, e47.
37. Ciuffi,A., Llano,M., Poeschla,E., Hoffmann,C., Leipzig,J., Shinn,P., Ecker,J.R. and Bushman,F. (2005) A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med*, **11**, 1287–1289.
38. Cherepanov,P., Devroe,E., Silver,P.A. and Engelman,A. (2004) Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J. Biol. Chem.*, **279**, 48883–48892.
39. Vanegas,M., Llano,M., Delgado,S., Thompson,D., Peretz,M. and Poeschla,E. (2005) Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. *J. Cell Sci.*, **118(Pt 8)**, 1733–1743.
40. Engelman,A. and Cherepanov,P. (2008) The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. *PLoS Pathog*, **4**, e1000046.
41. McKee,C.J., Kessl,J.J., Shkriabai,N., Dar,M.J., Engelman,A. and Kvaratskhelia,M. (2008) Dynamic modulation of HIV-1 integrase structure and function by cellular lens epithelium-derived growth factor (LEDGF) protein. *J. Biol. Chem.*, **283**, 31802–31812.
42. Pandey,K.K., Sinha,S. and Grandgenett,D.P. (2007) Transcriptional coactivator LEDGF/p75 modulates human immunodeficiency virus type 1 integrase-mediated concerted integration. *J. Virol.*, **81**, 3969–3979.
43. Raghavendra,N.K. and Engelman,A. (2007) LEDGF/p75 interferes with the formation of synaptic nucleoprotein complexes that catalyze full-site HIV-1 DNA integration in vitro: implications for the mechanism of viral cDNA integration. *Virology*, **360**, 1–5.
44. Zhao,Z., McKee,C.J., Kessl,J.J., Santos,W.L., Daigle,J.E., Engelman,A., Verdine,G. and Kvaratskhelia,M. (2008) Subunit-specific protein footprinting reveals significant structural rearrangements and a role for N-terminal Lys-14 of HIV-1 Integrase during viral DNA binding. *J. Biol. Chem*, **283**, 5632–5641.
45. Cherepanov,P. (2007) LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity *in vitro*. *Nucleic Acids Res*, **35**, 113–124.
46. Maertens,G., Cherepanov,P., Pluymers,W., Busschots,K., De Clercq,E., Debyser,Z. and Engelborghs,Y. (2003) LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J. Biol. Chem*, **278**, 33528–33539.
47. Cherepanov,P., Surratt,D., Toelen,J., Pluymers,W., Griffith,J., De Clercq,E. and Debyser,Z. (1999) Activity of recombinant HIV-1 integrase on mini-HIV DNA. *Nucleic Acids Res*, **27**, 2202–2210.
48. Kotova,S., Li,M., Dimitriadis,E.K. and Craigie,R. (2010) Nucleoprotein intermediates in HIV-1 DNA integration visualized by atomic force microscopy. *J. Mol. Biol*, **399**, 491–500.
49. Zhang,X., Pais,G.C., Svarovskaia,E.S., Marchand,C., Johnson,A.A., Karki,R.G., Nicklaus,M.C., Pathak,V.K., Pommier,Y. and Burke,T.R. (2003) Azido-containing aryl beta-diketo acid HIV-1 integrase inhibitors. *Bioorg. Med. Chem. Lett*, **13**, 1215–1219.
50. Nilsen,B.M., Haugan,I.R., Berg,K., Olsen,L., Brown,P.O. and Helland,D.E. (1996) Monoclonal antibodies against human immunodeficiency virus type 1 integrase: epitope mapping and differential effects on integrase activities *in vitro*. *J. Virol.*, **70**, 1580–1587.
51. Cherepanov,P., Ambrosio,A.L., Rahman,S., Ellenberger,T. and Engelman,A. (2005) Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. *Proc. Natl Acad. Sci. USA*, **102**, 17308–17313.
52. Hare,S., Di Nunzio,F., Labeja,A., Wang,J., Engelman,A. and Cherepanov,P. (2009) Structural basis for functional tetramerization of lentiviral integrase. *PLoS Pathog*, **5**, e1000515.
53. Arnold,K., Bordoli,L., Kopp,J. and Schwede,T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, **22**, 195–201.
54. Eswar,N., Webb,B., Marti-Renom,M.A., Madhusudhan,M.S., Eramian,D., Shen,M.Y., Pieper,U. and Sali,A. (2006) Comparative protein structure modeling using Modeller. *Curr. zProtoc. Bioinformatics*, Chapter 5, Unit 5 6.
55. Hare,S., Shun,M.C., Gupta,S.S., Valkov,E., Engelman,A. and Cherepanov,P. (2009) A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. *PLoS Pathog.*, **5**, e1000259.
56. Gupta,K., Diamond,T., Hwang,Y., Bushman,F. and Van Duyne,G.D. (2010) Structural properties of HIV integrase. Lens epithelium-derived growth factor oligomers. *J. Biol. Chem*, **285**, 20303–20315.

57. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
58. Grobler,J.A., Grobler,J.A., Stillmock,K., Hu,B., Witmer,M., Felock,P., Espeseth,A.S., Wolfe,A., Egbertson,M. *et al.* (2002) Diketo acid inhibitor mechanism and HIV-1 integrase: implications for metal binding in the active site of phosphotransferase enzymes. *Proc. Natl Acad. Sci. USA*, **99**, 6661–6666.
59. Tsiang,M., Jones,G.S., Hung,M., Samuel,D., Novikov,N., Mukund,S., Brendza,K.M., Niedziela-Majka,A., Jin,D., Liu,X. *et al.* (2010) Dithiothreitol causes HIV-1 integrase dimer dissociation while agents interacting with the integrase dimer interface promote dimer formation. *Biochemistry*, **50**, 1567–1581.
60. Tsiang,M., Jones,G.S., Hung,M., Mukund,S., Han,B., Liu,X., Babaoglu,K., Lansdon,E., Chen,X., Todd,J. *et al.* (2009) Affinities between the binding partners of the HIV-1 integrase dimer-lens epithelium-derived growth factor (IN dimer-LEDGF) complex. *J. Biol. Chem.*, **284**, 33580–33599.
61. Botbol,Y., Raghavendra,N.K., Rahman,S., Engelman,A. and Lavigne,M. (2008) Chromatinized templates reveal the requirement for the LEDGF/p75 PWWP domain during HIV-1 integration *in vitro*. *Nucleic Acids Res.*, **36**, 1237–1246.
62. Michel,F., Crucifix,C., Granger,F., Eiler,S., Mouscadet,J.F., Korolev,S., Agapkina,J., Ziganshin,R., Gottikh,M., Nazabal,A. *et al.* (2009) Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. *EMBO J*, **28**, 980–991.
63. Bischerour,J., Leh,H., Deprez,E., Brochon,J.C. and Mouscadet,J.F. (2003) Disulfide-linked integrase oligomers involving C280 residues are formed *in vitro* and *in vivo* but are not essential for human immunodeficiency virus replication. *J. Virol.*, **77**, 135–141.
64. Zhu,K., Dobard,C. and Chow,S.A. (2004) Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase. *J. Virol.*, **78**, 5045–5055.
65. Kessl,J.J., Eidahl,J.O., Shkriabai,N., Zhao,Z., McKee,C.J., Hess,S., Burke,T.R. Jr, Kvaratskhelia,M. *et al.* (2009) An allosteric mechanism for inhibiting HIV-1 integrase with a small molecule. *Mol. Pharmacol.*, **76**, 824–832.
66. Huang,F., Lerner,E., Sato,S., Amir,D., Haas,E. and Fersht,A. (2009) Time-resolved FRET study shows a compact denatured state of the B domain of Protein A. *Biochemistry*, **48**, 3468–3476.
67. Llano,M., Delgado,S., Vanegas,M. and Poeschla,E.M. (2004) Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase. *J. Biol. Chem.*, **279**, 55570–55577.
68. Christ,F., Voet,A., Marchand,A., Nicolet,S., Desimmie,B.A., Marchand,D., Bardiot,D., Van der Veken,N.J., Van Remoortel,B., Strelkov,S.V. *et al.* (2010) Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat. Chem. Biol.*, **6**, 442–448.
69. Molteni,V., Greenwald,J., Rhodes,D., Hwang,Y., Kwiatkowski,W., Bushman,F.D., Siegel,J.S. and Choe,S. (2001) Identification of a small-molecule binding site at the dimer interface of the HIV integrase catalytic domain. *Acta Crystallogr. D Biol. Crystallogr.*, **57(Pt 4)**, 536–544.
70. Al-Mawsawi,L.Q., Fikkert,V., Dayam,R., Witvrouw,M., Burke,T.R. Jr, Borchers,C.H. and Neamati,N. (2006) Discovery of a small-molecule HIV-1 integrase inhibitor-binding site. *Proc. Natl Acad. Sci. USA*, **103**, 10080–10085.