Architecture and Assembly of HIV Integrase Multimers in the Absence of DNA Substrates*5

Ravi Shankar Bojja, Mark D. Andrake, George Merkel, Steven Weigand†, Roland L. Dunbrack, Jr., and Anna Marie Skalka†

From the Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Background: No full-length structure of HIV integrase alone has been reported.

Results: We elucidated the architectures of dimers and tetramers of full-length HIV-1 integrase in solution.

Conclusion: HIV apo-integrase can assemble in two alternate dimer forms: a reaching and a core-core dimer. The tetramer comprises two stacked reaching dimers, stabilized by core-core interactions.

Significance: New insights into HIV integrase architecture and its inhibition are suggested.

We have applied small angle x-ray scattering and protein cross-linking coupled with mass spectrometry to determine the architectures of full-length HIV integrase (IN) dimers in solution. By blocking interactions that stabilize either a core-core domain interface or N-terminal domain intermolecular contacts, we show that full-length HIV IN can form two dimer types. One is an expected dimer, characterized by interactions between two catalytic core domains. The other dimer is stabilized by interactions of the N-terminal domain of one monomer with the C-terminal domain and catalytic core domain of the second monomer as well as direct interactions between the two C-terminal domains. This organization is similar to the “reaching dimer” previously described for wild type ASV apoIN and resembles the inner, substrate binding dimer in the crystal structure of the PFV intasome. Results from our small angle x-ray scattering and modeling studies indicate that in the absence of its DNA substrate, the HIV IN tetramer assembles as two stacked reaching dimers that are stabilized by core-core interactions. These models of full-length HIV IN provide new insight into multimer assembly and suggest additional approaches for enzyme inhibition.

Integration of retroviral DNA into the host cell genome is an essential step in viral replication catalyzed by the viral integrase (IN)3 protein (for a recent review, see Ref. 1). The integration reaction takes place in two temporally and spatially separated steps known, respectively, as 3’-processing and end joining (or strand transfer). In the first step a di- or trinucleotide is hydrolytically removed from each viral DNA end, exposing 3’-hydroxyl groups of invariant CA dinucleotides. In the second step the two processed viral DNAs are joined to host DNA in a concerted cleavage and ligation reaction. The functional nucleoprotein complex that includes two viral DNA ends and a multimer containing at least four IN monomers is known as an intasome (2). The pathway by which intasomes are assembled from apoIN proteins is unknown. In the absence of the viral DNA substrate, soluble full-length IN proteins from different viruses can exist as monomers, dimers, or tetramers at similar concentrations. For example at ~1–4 mg/ml, the prototype foamy virus (PFV) IN is a monomer; avian sarcoma virus (ASV) IN is a dimer, and the human immunodeficiency virus (HIV) IN is a tetramer. Furthermore, the three domains of IN proteins, the N-terminal domain (NTD), catalytic core domain (core), and C-terminal domain (CTD), play important roles in the conformational dynamics and multimerization of the apo-protein that are not fully understood.

We recently employed small and wide angle x-ray scattering (SAXS/WAXS) in combination with chemical cross-linking to elucidate the full-length architecture of the apoIN monomer and dimer of ASV (3). We showed that the ASV IN dimer derives its stability from interactions of the NTD of one monomer with the CTD and core of the second monomer and from CTD-CTD contacts between the monomers; core-core interactions are only observed in tetramers. We called this unique and unexpected architecture a reaching dimer (illustrated in Fig. 1, right). We also found that the substitution of a single hydrophobic residue at an interaction hotspot in the CTD-CTD interface was sufficient to disrupt the reaching dimer and noted that a similar hydrophobic hotspot could be predicted for other retroviral IN proteins, including that of HIV. The arrangement of monomers in the ASV IN reaching dimer structure resembles that of the inner dimer of the PFV intasome that engages the viral and target DNA substrates and catalyzes the concerted joining reaction. We, therefore, proposed a dual role for hydrophobic residues in the CTD hotspot; that is, DNA binding and reaching dimer stabilization in the absence of DNA, and binding to the non-cleaved strand of viral DNA (3). To test the generality of reaching dimer assembly and to gain further
HIV-1 Integrase (Dimer and Tetramer) Solution Structures

FIGURE 1. Schema for alternate apoIN assemblies. Two possible HIV apoIN dimer forms are illustrated in this schematic. The three domains common to retroviral integrase proteins are depicted in color-coded shapes identified in the monomer as N for the red NTD, C for the green CTD, and core for the blue catalytic core domain. A small white circle symbolizes the active site in the core. The arrangement of domains in the monomer and reaching dimer are adapted from the published architectures of ASV monomers and dimers (3). The arrangement of the core-core dimer is adapted from the “outer” dimers in the crystal structure of the PFV intasome (2). A crystal structure of the full-length apoIN tetramer is not available. An architecture based on our solution studies of HIV IN is proposed in the present study (cf. Fig. 10).

insight into the modes of assembly of apoIN, we analyzed the solution structure of HIV IN using approaches that proved successful with the ASV IN protein.

EXPERIMENTAL PROCEDURES

Static/Size Exclusion Chromatography (SEC)-SAXS/WAXS of ApoHIV IN—X-ray scattering experiments were performed at the Advanced Photon Source at Argonne National Laboratories, 5ID-D beamline, Chicago, IL. Data were collected either directly from the homogeneous protein solutions or with protein fractions that were eluted at 600 μl/min from a Tricorn column (Superdex™ 200, 10/300 GL, GE Healthcare) immediately upstream of the SAXS flow cell. In the latter case, because the proteins were eluting at high concentrations, 3 scans at the retention time were averaged at an interval of 14 s. Specific details concerning protein expression and purification, experimental setup, data collection, fitting, and shape modeling are described in the supplemental Experimental Procedures.

Protein Cross-linking—Tag-less HIV IN proteins were buffer-exchanged by dialysis in 0.1 M MES-HCl, 1 m NaCl, pH 6.0, 1 mM Tris(2-carboxyethyl)phosphine, 20% glycerol (4). For wild type HIV IN cross-linking, a 1:1 mixture of unlabeled and isotopically labeled IN was first treated with 10 mM EDTA for 10–15 min on ice, and then dialyzed on ice in 0.1 M MES-HCl 1 m NaCl, pH 6.0, 20% glycerol supplemented with 2 mM DTT, 20 mM MgCl₂, and 50 μM ZnSO₄. After 60 min to allow for refolding of the NTD, the mixture was dialyzed in 0.1 M MES, pH 5.8, 1 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine, 20% glycerol. Cross-linking of the IN F181T mixtures was as described for wild type IN. The cross-linked products were separated by electrophoresis in denaturing NuPAGE 4–12% BisTris gels using MES running buffer and Coomassie Blue stain. Sample recovery was only slightly diminished by acetone precipitation. The dimer bands from all EDC reactions were excised, trypsin-digested, and analyzed for cross-links by mass spectrometry as described previously (3) and in the supplemental Experimental Procedures.

HADDOCK Docking and Fine Model Fit—to model the flexible HIV F181T IN dimer interface, we used HADDOCK docking (Guru Interface) (6) together with SANS-driven refinement parameters and distance constraints from the mass scattering analysis of the protein chemical cross-linking. Starting models for docking were based on homology with the model of the ASV IN dimer (3) using the SWISS MODEL resource (7–9), and cross-linking residues were defined to have a proximity of ~4 Å between each pair. Based on the flexibilities of the IN domains, docking was grouped into three classes that satisfied the identified chemical cross-links. Structures were selected for further refinement based on the HADDOCK score, and models were clustered with a cutoff root mean square of 10 Å that satisfied the SAXS maximum distance (D_{max}). All the final models from each group that have a maximum dimension equivalent to experimental SANS data were compared using CRYSSOL analysis fit (ATSAS) (10). In addition, P(r) functions were plotted for each model and compared with the experimental data to assess the quality of the dock model using the Igor Pro package (Irena macro) (11).

RESULTS

Preparation and Characterization of Wild Type HIV ApoIN—HIV IN is notoriously difficult to maintain at moderate to high concentrations because of its tendency to aggregate. After investigating a variety of buffer conditions we, as others (12), found that solubility and stability could be optimized by the inclusion of 1 M urea during protein purification. Analysis of wild type HIV IN (8 μM) by CD spectroscopy revealed no significant differences in the α-helical structural elements of the protein in the region of interest (218–223 nm) in the presence of 0–1 M urea; denaturation was not detected until the urea concentration reached 2 M or higher (Fig. 2A). In addition, our enzymatic assays showed that single viral DNA end processing and joining activities were unaffected in the presence of 1 M urea; concerted integration into a target DNA was unaffected by 250 mM urea, and this activity was still detectable in 1 M urea (Fig. 2B–D). We concluded from these results that HIV IN retains its native structure in the presence of this chaotropic reagent at concentrations employed in our biophysical analyses.

Static and dynamic light scattering analyses were used to determine molecular mass and to gauge the homogeneity of our HIV IN protein preparations. We observed that wild type HIV IN exhibited properties expected of a homogeneous tetramer in
the presence of 250 mM and 1 M urea at protein concentrations in the 1–2 mg/ml range (Table 1). In contrast, the protein appeared to be a mixture of tetramers and larger aggregates at similar concentrations in the absence of urea (data not shown).

Disruption of Hydrophobic Interactions in the Core-Core Interface of HIV IN Blocks Tetramer but Not Dimer Formation—We have demonstrated that in the absence of DNA, full-length ASV IN forms two distinct subunit interfaces. A reaching dimer interface is stabilized by CTD-CTD interactions and interactions of the NTD from one monomer with the CTD and core domain of the second monomer. A second interface stabilized by core-core domain interactions is observed in ASV IN tetramers, which we and others have shown are required for catalysis of concerted integration but not 3’-end processing of viral DNA (3). In the case of HIV, Neamati and co-workers (13, 14) have noted that a four-tiered π interaction among several conserved hydrophobic amino acids at the core-core interface (Trp-132, Met-178, Phe-181, Phe-185) is likely to be critical for its stability. Although molecular masses were not determined for proteins with non-conservative substitutions of these residues, these authors showed that such changes resulted in loss of the single end-joining activity of HIV IN but had less effect on 3’ end-processing, as might be expected for some capacity to form dimers (15, 16).

By analogy with ASV IN, we hypothesized that substitution of one or more of these conserved hydrophobic residues in the HIV IN core domain might block formation of tetramers but not reaching dimers. To test this idea, we introduced three independent, non-conservative substitutions for residue Phe-181 in wild type HIV IN and analyzed these proteins by dynamic light scattering. The results showed that the HIV IN derivatives with either threonine (1.2 mg/ml) or alanine (3.2 mg/ml) at position 181 were, indeed, homogenous dimers in the presence of 1M urea; the protein that contained glycine (2.3 mg/ml) at this position had properties expected for a mixture of dimers and tetramers (Table 1). Our enzymatic assays showed that the IN F181A derivative was essentially inactive for 3’ end-processing, but F181T and F181G retained ~12% of the catalytic rate exhibited by wild type HIV IN (Fig. 3A). Partial processing activity for the F181G substitution was reported previously (13). These data indicate that the HIV IN protein can form partially active dimers when core-core interactions are disrupted.

Destabilization of NTD Structure Blocks Formation of Reaching Dimers but Not Core-Core Dimers—The NTD of HIV IN contains a conserved Zn²⁺ binding motif (HH-CC), and the presence of this ion is required for conformational integrity of this domain (17–19). As the reaching dimer of ASV IN is stabilization.
HIV-1 Integrase (Dimer and Tetramer) Solution Structures

Lized by interactions of the NTD with the core and CTD of the second IN monomer, we reasoned that disruption of the NTD structure would prevent formation of an HIV IN reaching dimer but not a dimer formed by core-core interactions (illustrated in Fig. 1, left). To remove the NTD-bound Zn$^{2+}$ ion, wild type HIV-1 IN (1.5 mg/ml) was dialyzed overnight in our standard buffer (including 1 M NaCl and 1 M urea) supplemented with 10 mM EDTA and then concentrated to 5 mg/ml. A sample of this protein was then applied to a SEC column that had been pre-equilibrated with the same EDTA-supplemented buffer. A homogeneous peak of protein was eluted from this column with retention time (24.3 min) expected for a dimer. The F181T derivative eluted as a dimer in SEC (24.75 min) in the absence of EDTA treatment. When the IN F181T was treated with EDTA and chromatographed in the presence of EDTA, its retention time was 26.2 min, consistent with a monomer. Light scattering analysis of untreated IN F181T and EDTA-treated wild type IN confirmed that both were dimers (Tables 1 and 2). These results support the hypothesis that destabilization of the NTD by removal of Zn$^{2+}$ ion leads to disruption of the reaching dimer interface but will allow core-core stabilized dimers to assemble. Further support for this interpretation comes from the demonstration by Hare et al. (20) that an E11K substitution, which disrupts a salt bridge between the NTD and Lys-186, results in a shift in the equilibrium of multimeric forms of full-length wild type HIV-1 IN from tetramers to a dimer-monomer mixture. Our light scattering data, at higher concentrations (Table 1), also revealed dimers, and an expected reduction in enzymatic activity of this derivative is illustrated by analysis of 3'/5' processing (Fig. 3 B).

SAXS Analysis of Wild Type HIV IN and the F181T and E11K Dimers—The homogeneous preparations of wild type HIV IN tetramers and derivative dimers were next analyzed by SAXS at protein concentrations ranging from 1 to 2 mg/ml. A summary of the SAXS-determined parameters and apparent multimeric state of each of the proteins analyzed is provided in Table 2. Consistent with the light-scattering results, SAXS data for all proteins in the Guinier regions and Kratky plots confirmed the absence of aggregation or unfolding of IN (data not shown). A PRIMUS analysis (10) of scattering intensity versus $q$ range of 0.01–0.04 for four independent wild type IN concentrations also showed that there was no concentration-dependent aggregation in the range tested (data not shown).

The scattering profiles for wild type HIV IN and the F181T and E11K dimers are shown in Fig. 4 A; their pair distance dis-

![Figure 3](https://example.com/figure3.png)

**FIGURE 3.** Processing activities of HIV-1 IN Phe-181- and E11K-substituted derivatives. **A**, shown are Phe-181-substituted proteins compared with wild type IN. Proteins were treated with 10 mM EDTA overnight at 4 °C in storage buffer to optimize activity (19) (see supplemental Experimental Procedures for details). **B**, shown is a comparison of wild type HIV IN and the Glu-11-substituted protein. The fluorescence-based assay for single 3' end-processing and details of data acquisition and processing were as described (44).

---

**TABLE 1** Light scattering data for wild type HIV IN and derivatives

| Protein                        | [Protein]$^a$ | MW-I$^b$ | $S^c$  | PD$^d$ | Pd-D$^e$ | Comment | $R_h^f$ |
|-------------------------------|---------------|----------|--------|--------|----------|---------|--------|
| HIV-1 IN wild type            | 1.63          | 167      | 7.7    | 56     | 10       | Tetramer | 62     |
| HIV-1 IN wild type + Mg$^{2+}$| 1.58          | 132      | 4.2    | 45     | 11       | Tetramer | 65     |
| HIV-1 IN wild type + EDTA (SEC)| 1.2           | 50       | 7.3    | 77     | 10       | Dimer   | 43     |
| HIV-1 IN D64N                 | 2.48          | 149      | 2.2    | 29     | 18       | Tetramer | 70     |
| HIV-1 IN D64N + Mg$^{2+}$     | 2.4           | 156      | 3.5    | 30     | 27       | Tetramer | 76     |
| HIV-1 IN F181T                | 1.5           | 68       | 12     | 57     | 9        | Dimer   | 46     |
| HIV-1 IN F181T + Mg$^{2+}$    | 1.2           | 76       | 0.8    | 53     | 19       | Dimer   | 44     |
| HIV-1 IN F181A                | 3.2           | 64       | 5.4    | 57     | 11       | Dimer   | 44     |
| HIV-1 IN F181G                | 2.3           | 116      | 8.5    | 57     | 11       | Mixture | 57     |
| HIV-1 IN E11K                | 1.85          | 76       | 9.5    | 44     | 26       | Dimer   | 60     |

$^a$ Protein concentration in mg/ml, measured using molar extinction coefficient of the respective IN protein.

$^b$ Apparent molecular mass (MW-I) was determined by static light scattering and calculated using DynaPro Version 5 software.

$^c$ Percentage standard deviation, $S$.

$^d$ The % polydispersity of the sample was determined using a cumulants analysis.

$^e$ Polydispersity index.

$^f$ Hydrodynamic radius, $R_h$, of the apparent multimer calculated from differential coefficient of the Stokes Einstein equation.

---

## REFERENCES

1. A. Hare, *et al.* (2020) J. Biol. Chem. 295, 3454-3464.
2. B. Smith, *et al.* (2020) Protein Sci. 29, 2245-2258.
3. C. Zhang, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 14563-14572.
4. D. Lee, *et al.* (2020) Cell 182, 1654-1668.
5. E. K. Johnson, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 80-85.
6. F. S. White, *et al.* (2020) J. Mol. Biol. 430, 678-690.
7. G. M. Whitesides, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
8. H. Zhang, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 80-85.
9. I. Lee, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 6542-6547.
10. J. M. Blakely, *et al.* (2020) J. Mol. Biol. 430, 1-12.
11. K. Honda, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
12. L. Wang, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
13. M. H. Kim, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
14. N. R. Kevrekidis, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
15. O. Tanaka, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
16. P. Tseng, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
17. Q. Wang, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
18. R. Zhang, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
19. S. Zhang, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
20. T. Honda, *et al.* (2020) Proc. Natl. Acad. Sci. U. S. A. 117, 12342-12347.
distribution $P(r)$ functions using data to a $q_{\text{max}}$ of 0.4 Å$^{-1}$ revealed no major deviations in the $D_{\text{max}}$ as a function of concentration in the range of 1–1.5 mg/ml (Fig. 4B). SAXS scattering parameters for IN F181T, either as a stock preparation or from direct analysis of a size exclusion column fraction (SEC-SAXS), show average $R_g$ values of 33–36 Å, and $D_{\text{max}}$ values of 107–113 Å, respectively, whereas the E11K protein showed an $R_g$ of 39 Å and $D_{\text{max}}$ of 124 Å. An $R_g$ of 40 Å and $D_{\text{max}}$ of 117 Å were obtained with the wild type HIV IN.

| Protein                        | [Protein] | $R_g$ | $D_{\text{max}}$ | $I(0)$ | Volume | Apparent multimer (LS) | MW-I² |
|-------------------------------|-----------|-------|------------------|--------|--------|------------------------|-------|
| HIV-1 IN wild type            | 1.7       | 40 ± 1| 117 ± 2          | 0.053–0.06 | 260 | Tetramer               | 168   |
| HIV-1 IN wild type + Mg²⁺     | 1.6       | 41 ± 1| 132 ± 1          | 0.06   | 305 | Tetramer               | 132   |
| HIV-1 IN D64N                 | 2.8       | 44 ± 2| 142 ± 3          | 0.66   | 281 | Tetramer               | 149   |
| HIV-1 IN D64N + Mg²⁺          | 2.3       | 45 ± 1| 147 ± 1          | 0.07   | 288 | Tetramer               | 156   |
| HIV-1 IN F181T                | 1.9       | 33 ± 2| 107 ± 3          | 0.03   | 150 | Dimer                  | 60    |
| HIV-1 IN F181T + Mg²⁺         | 1.9       | 34 ± 1| 103 ± 1          | 0.027  | 166 | Dimer                  | 76    |
| HIV-1 IN WT EDTA(SEC)³        | 1.2       | 33 ± 2| 100 ± 2          | 0.04   | 233 | Dimer                  | 50    |
| HIV-1 IN F181T EDTA(SEC)⁵     | 1.9       | 29 ± 1| 88 ± 1           | 0.025  | 82  | ND                     | ND    |
| HIV-1 IN E11K                 | 1.8       | 39 ± 1| 124 ± 1          | 0.032  | 179 | Dimer                  | 76    |
| HIV-1 IN EDTA(SEC)            |           |       |                  |        |       |                        |       |
| HIV-1 IN F181T EDTA(SEC)      |           |       |                  |        |       |                        |       |

⁴SAXS scattering data obtained from IN proteins at the listed concentrations were processed with the program IRENA to determine the radius of gyration ($R_g$) and the maximum length of the scattering multimer ($D_{\text{max}}$) and $I(0)$.

⁵The volume of the Situs-derived envelope was calculated with Chimera software (UCSF).

⁶The apparent multimer was determined from light scattering (LS).

⁷The MW-I as determined by DynaPro Dynamics software is expressed in kilodaltons.

⁸EDTA-treated IN proteins separated by SEC were directly injected into the SAXS beam line.

FIGURE 4. SAXS data and envelopes for wild type HIV IN and the F181T and E11K derivatives. A, shown is intensity versus $Q$ (Å$^{-1}$) at similar protein concentrations (1.6 mg/ml). B, $P(r)$ functions show distinct curves but similar $D_{\text{max}}$. C, shown are two orthogonal views of the SAXS-determined envelopes of the wild type HIV IN tetramer and the F181T and E11K dimers. Data were obtained as described under “Experimental Procedures” (see also supplemental Fig. S2).
monomers were also analyzed by SAXS. Envelopes derived for these EDTA-treated proteins are shown in Fig. 6 and, judging from their Kratky plots (not shown), both of these proteins have lost some secondary structure compared with untreated proteins, as expected for destabilized NTDs (18). SAXS parameters for the EDTA-derived wild type IN dimer show values of 33 Å for the \(R_g\) and 100 Å for \(D_{\text{max}}\), which are quite similar to the values of 36 and 113 Å, respectively, for the SEC-isolated IN F181T dimer. However, the volume of the wild type EDTA-derived dimer is larger than that of the IN F181T dimer, indicating that the conformations are likely to be distinct (Table 2). Indeed, the envelopes derived for the wild type IN EDTA dimer and the F181T IN dimer are quite different, consistent with the notion that there are two distinct modes of stabilization of their respective dimer interfaces. The difference in contours of the wild type EDTA dimer and the E11K dimer likely reflects the disorder of the NTD associated with the removal of \(Zn^{2+}\) ion. The EDTA-treated HIV IN F181T monomer exhibited reduced values for \(R_g\) and \(D_{\text{max}}\) compared with the untreated protein. These results are consistent with a monomer with disordered NTD (Fig. 6).

Identification of Intersubunit Proximities in the HIV IN Reaching Dimer—Previously we used isotopic labeling followed by chemical cross-linking and mass spectrometry to map the interacting interfaces in ASV IN dimers and tetramers (3). In the present study we employed similar methods to identify the intermolecular proximities of protein domains in the HIV IN F181T dimer, and the wild type IN dimer(s) that exist at low protein concentration (supplemental Table S1).

The strategy was to mix equal amounts of separate preparations of unlabeled and isotopically labeled lysine \((^{13}\text{C},^{15}\text{N})\) and arginine \((^{13}\text{C},^{15}\text{N})\) proteins and allow the mixtures to equilibrate such that they formed mixed multimers (5). In preliminary tests, we found that labeled and unlabeled monomers of the F181T derivative did not exchange as readily as those of wild type IN, indicating that the F181T dimer was somewhat more stable than the wild type. To facilitate exchange we treated the F181T protein with 10 mM EDTA to form monomers; mixed dimers were readily assembled upon the addition of \(Zn^{2+}\) ion.

SAXS Analysis of EDTA-treated HIV IN Proteins—EDTA-treated, SEC purified wild type HIV IN dimers and F181T
through slow dialysis (see “Experimental Procedures”). After treatment with the EDC cross-linking reagent, the wild type IN or F181T IN products were separated by electrophoresis in a denaturing gel. Protein excised from the dimer band was then subjected to trypsin digestion and mass spectrometry. Inter-subunit cross-linked peptides are recognized uniquely by their hybrid mass. The observed mass differences for cross-linked labeled and unlabeled tryptic peptides were consistent with the expected values of $K_{d}/H_{11001} = 8.014$ and $R_{d}/H_{11001} = 10.008$, where $K$ and $R$ are masses of unlabeled lysine and arginine, respectively.

The reagent used in these experiments, EDC, promotes formation of trypsin-resistant, irreversible cross-links between the carboxyl groups of acidic amino acids, such as aspartate and glutamate, with the side chains of lysine residues that act as salt bridge partners. Gel electrophoresis of the IN F181T derivative includes an N-terminal extension of three amino acid residues, Gly, His, and Met, which remains after removal of the His tag. This Gly-1 is also capable of forming cross-links to the core domain at Glu-157 and Glu-170. The C-terminal tail extremities of both the labeled and unlabeled monomers were found to

![Diagram](image-url)
form cross-links to G1 in the other subunit. As the labeled IN only incorporates isotopes of Lys and Arg, the origin of the cross-linked tryptic fragments that include the extreme tail, with sequence QDED, could not be determined. G1 of the unlabeled monomer of F181T IN is observed to form cross-links with labeled IN at helical regions of the NTD at positions Glu-10, Glu-11, and Glu-13. Glu-35 from unlabeled IN is cross-linked with G1 of the labeled IN. These cross-links reveal the proximities of the NTD helical regions in the unlabeled and labeled IN subunits.

In addition to cross-links to secondary structural elements of the NTD, core, and CTD tail ends, G1 was observed to cross-link at CTD β-sheet elements at positions Asp-229 and Glu-246 in IN F181T. Furthermore, Glu-212 from both labeled and unlabeled IN was found to cross-link with G1 on unlabeled and labeled IN, respectively (Fig. 7B). Helical regions of the NTD from the unlabeled IN at positions Asp-6, Glu-11, Glu-13, and Glu-35 were also cross-linked to the labeled lysines in either the core or CTD regions. Cross-links between β-sheet regions in the CTDs of the unlabeled and labeled F181T IN were observed between residues Glu-212 and Glu-246, with Lys-236, Lys-240, and Lys-264 (Fig. 7B). Overall these results indicate that the HIV IN F181T dimer is stabilized by interactions of the NTD of the one monomer subunit with the core, CTD, and NTD of the second subunit. In addition, CTD-CTD interactions are also detected between the two subunits. Finally, no cross-links were observed between the core domains of the monomer subunits, indicating that this dimer is not stabilized by core-core interactions but rather more closely resembles the reaching dimer architecture observed with wild type ASV IN. Results from mass spectrometry analysis of cross-linked peptides formed with IN F181T at 250 nM protein concentration are summarized in Fig. 7C and supplemental Table S1D. As expected, fewer interactions were detected under these conditions. However, all of the cross-links found in this sample were included in the set obtained at the higher protein concentration (Fig. 7B).

Results from mass spectrometry analysis of cross-linked peptides from the dimer band of wild type HIV IN treated with EDC at 450 nM concentration are summarized in Fig. 7D and supplemental Table S1E. We detected 17 cross-links between labeled and unlabeled IN. Although we again observed prominent interactions of the NTD with the core and CTD, the cross-links are at different positions than observed with the IN F181T dimer. For example, the CTD to CTD cross-links in the wild type IN dimer were Glu-212—Lys-211, Asp-229—Lys-264, Asp-232—Lys-264, and Lys-264—Asp-232, whereas with IN F181T, the cross-links were Glu-212—Lys-236, Glu-246—Lys-240, and Glu-246—Lys-264 (compare Figs. 7, B and D). Other distinct features of the wild type dimer were observed in the region of the core-CTD interface, including cross-links of Glu-170—Lys-264, Asp-229—Lys-159, and Asp-167—Lys-264, and NTD-core and NTD-CTD cross-links at Glu-35—Lys-160 and Asp-3—Lys-264. Furthermore, core-core cross-links at Glu-157—Lys-186 and Asp-167—Lys-160 were interactions observed only with the wild type dimer. Most notably, unlike the F181T dimer, no cross-links were observed between the NTDs from each interacting monomer in the wild type dimer.

A number of the intermolecular cross-links observed with the wild type HIV IN were mutually exclusive, likely arising from a mixture of oligomers. The possibility that some of the cross-links (e.g. core-core adducts) reflect inclusion in the SDS-PAGE dimer band of products from cross-linked tetramers cannot be excluded. That caveat aside, the results are most consistent with a cross-linked population comprising alternate reaching dimer configurations as well as some core-core dimer forms. With respect to the wild type reaching dimer architecture, the NTD-core cross-link between Glu-35 and Lys-160 is likely to represent a biologically relevant interface, as a careful analysis using PROTCID (21) reveals that a common interface with these residues in close proximity exists in all two-domain (NTD + core) IN proteins that have been crystallized to date, including those of HIV-1, HIV-2, PFV, and maedi visna virus. The interface (supplemental Fig. S1) is observed in both intra- and intermolecular interactions. In contrast to the wild type, there is no evidence for this common interface in the IN F181T reaching dimer. It seems probable that the non-conservative substitution of Phe-181, which is located in this interface (supplemental Fig. S1), disrupts this common NTD-core interaction leading to a shifted reaching dimer conformation that is manifested most clearly in the adjacent arrangement of NTDs in the IN F181T dimer. Such a shift could account for the increased stability of the F181T dimer noted above.

Data-driven Docking and Model Fit into SAXS Envelopes—To obtain a more detailed model of the HIV IN F181T dimer, we used the mass spectrometric data obtained from the cross-linking experiment summarized in Fig. 7B. The starting template for a monomer of F181T IN was a homology model derived from the ASV IN reaching dimer (3), with data-driven docking using the HADDOCK module. We noted that the presence of mutually exclusive β-strand cross-links at the CTD-CTD interface (Fig. 7, B and C) is likely to reflect dynamic movement at this interface. For example, NTD residues are observed with cross-links to either core or CTD, and further cross-links between CTD domains engage different β-strands (Fig. 7B). Therefore, to characterize the predominant conformational state of the IN F181T dimer, cross-link contacts were parsed into three groups for which the distance constraints are compatible and satisfy the maximum distance (Dmax) from SAXS data (Fig. 8).

In the first group dock model A (Fig. 8A, top) the distance constraints of the Glu-246—Lys-240 were maintained in addition to NTD-NTD interactions with cross-links of G1 with Glu-10, Glu-11, Glu-13, and Glu-35; NTD cross-links to CTD at G1—Glu-212 and Glu-11—Lys-215 are also included (Fig. 7B). In the second group dock model B, the salt bridge distances of Glu-246—Lys-264 were maintained in addition to the links between Glu-11—Lys-215, Glu-35—Lys-264, and G1—Asp-116. In this conformation, the interactions between the Trp-243 residues from each monomer are disrupted. In the third group dock model C interactions are similar to dock model A, except NTD-NTD proximities were relaxed. NTD-to-core distances were enforced at G1—Asp-116, G1—Glu-157, Asp-6—Lys-159, and Asp-6—Lys-188. In this model, NTD cross-links to CTD at G1—Glu-212 and G1—Lys-215 are not taken into consideration. During successive cycles of the optimization for
each group, we included newly observed features that increased the hydrophobic packing in the interacting domains at the interface and that satisfied the SAXS determined maximum distance $D_{\text{max}}$.

Because of the dynamic nature of movements of the NTD and CTD with respect to the core domain, in all three dock models we endeavored to determine which configuration might represent the closest parity with the actual experimental data and, therefore, the predominant dimer architecture. Although independently obtained, comparison of the CRYSOL-derived $P(r)$ function for each of the models with the experimental data and analysis of the predicted scattering profiles (Fig. 8, B and C) indicated that the closest identity was obtained with model A. Furthermore, a predicted inner dimer model based on the PFV intasome crystal structure (22) with viral DNA removed (i.e. intasome inner dimer sans DNA; Fig. 8A) showed a large difference of 30 Å in $D_{\text{max}}$ and scattering when compared with experimental data. These analyses showed good correlation between the data-driven docking reaching dimer model A and the observed SAXS parameters.

The unique features of the HIV IN F181T model A reaching dimer are a CTD-CTD interface with prominent Trp-243 hydrophobic interactions (yellow spheres, Fig. 8A) and stacking of both NTD domains above the CTD domains. The final, most accurate model from the HADDOCK docking was fitted into the experimental SAXS envelope, and its occupancy inside the

---

**FIGURE 8. Models of the HIV IN F181T reaching dimer and selection of the best fit.** A, shown is a depiction of the three HADDOCK-derived models, generated by docking with distance constraints from EDC cross-links, and published intasome models sans DNA. The best-fit Model A, which includes NTD and CTD stacking interactions at the interface, is shown in two orthogonal views. The interface of this reaching dimer includes interactions between the NTD of one monomer (red) to the NTD of the second monomer (pink) and also Trp-243:Trp-243 stacking between the subunit CTDs (yellow spheres). Model B was derived without Trp-243 stacking, and model C was derived with inclusion of NTD to core interactions. The intasome and intasome inner dimer models are derived from the published HIV intasome model (22) by omitting the viral DNAs. IN domains are colored as in Fig. 1, and the second monomer is shown in faded colors. B and C shown a comparison of the theoretical $P(r)$ functions and scattering profiles (generated with CRYSOL) of the HADDOCK models and the intasome inner dimer with experimental data.
The envelope was 95% (Fig. 9, left). Although model A in Fig. 8A includes only residues 1–270, our analyses indicated that the addition of the 18-amino acid tail and disordered N-terminal His6 into the final model had insignificant effects on the CRYSOL profiles (data not shown).

As noted above, the SAXS-derived envelope of the E11K IN dimer is longer than either the F181T dimer or the wild type tetramer; its unique shape includes narrow extremities and bulging occupancy at the center (Fig. 4C). As the core interface residues are not changed in the E11K substitution, this dimer should include a core-core domain interaction. To model an architecture that would satisfy these features, residues 60–186 of the core domain crystal structure (PDB ID 1BIS) was selected to represent the core domain dimer, which could only be accommodated in the central bulge of the E11K envelope. Attempts to place NTD and CTD dimer pairs at the narrow extremities of the envelope (PDB IDs 1E0E and 1QMC, respectively) resulted in a model that did not fit the experimental data (not shown). The alternative arrangement, with an NTD and CTD at both ends, was a better fit, and this model was further refined with SWISS MODEL by placing one CTD domain at either extremity with the NTDs nested between the CTD and core domains; the missing linkers were built between the domains as shown in Fig. 9A, right. The SAXS scattering data and P(r) function derived for this E11K dimer model compared well with the experimental data (Fig. 9, B and C) and, as expected, was distinct from F181T data.

**Building an HIV IN Tetramer**—As described above, disruption of core-core interactions at Phe-181 of HIV IN revealed a unique homo-dimer whose SAXS analyses yielded very similar $D_{\text{max}}$ to that of the homo-tetramer of wild type IN (Fig. 4B). The volumes from the SAXS-derived envelopes of the wild type IN tetramer and F181T IN dimer, 260 and 130 Å³, respectively (Table 2), indicated that the homo-tetramer has twice the volume of a homo-dimer with similar $D_{\text{max}}$. Comparison of these parameters and the SAXS envelope shapes suggests that IN tetramers assemble by the stacking of two reaching dimers through core-core interactions of both dimers. Fig. 10A shows a plausible model for the formation of a wild type tetramer by such stacking. Coalescence of the two reaching dimers in this manner results in a tetramer that has same maximum dimension as a reaching dimer. The corresponding wild type tetramer model, which places the NTDs in close proximity to the core interfaces, fits into the SAXS envelope as shown in the Fig. 10B.

CRYSOL was applied to gauge the correctness of this tetramer model without any bias with respect to the experimental data. The intasome sans DNA model (22) was also included in this comparison. The results indicate that the apo-form of the HIV IN tetramer comprising stacked reaching

![Figure 9.](image-url)
dimers is most compatible with the experimental data (Fig. 10C). Furthermore our SAXS data for apoIN are inconsistent with the elongated model of the HIV intasome sans DNA tetramer derived from the crystal structures (2, 22) or solution structures (23) of PFV.

**DISCUSSION**

**Elucidation of the Organization of Two Dimer Forms of HIV IN**—In this study we applied SAXS, protein cross-linking coupled with mass spectrometry, and molecular modeling to reveal the architectures of full-length HIV IN dimers in the absence of DNA substrates (apoIN). Our analyses have allowed us to distinguish two dimer forms of the protein. One form, stabilized by core-core interactions, is observed when a charge substitution, E11K, is introduced into the NTD (Table 1; Fig. 4) or the NTD is destabilized through removal of an essential Zn$^{2+}$ ion from the wild type protein (Table 2; Fig. 6). Previous investigations have shown that that wild type EDTA-treated IN formed monomers but at much lower concentrations (19) that may be below that required for stable core-core interactions and are not suitable for SAXS analysis. Our preliminary modeling experiments indicate that the envelopes of the E11K derivative (Fig. 9) and EDTA-derived wild type dimer (not shown) can accommodate a centrally located core-core interface, with the NTD and CTD domains extended to either side. It is noteworthy that neither of these core-core-stabilized dimer envelopes resembles that predicted for the HIV IN dimer model derived from consolidation of the crystal structures of two-domain fragments (3).

The second dimer form has a reaching dimer architecture similar to that derived for wild type ASV IN (3). In this dimer, the core domains lie at opposite poles, and the structure is stabilized by interactions of the NTD of one monomer with the core and CTD of the second monomer as well as CTD-CTD interactions, mediated by Trp-243 stacking (Table 2 and Fig. 8). We showed that interruption of the 4-tiered interaction between conserved hydrophobic amino acids at the HIV IN core-core interface resulted in formation of a homogeneous reaching dimer. Although several residues were proposed to participate in the 4-tiered interaction, Trp-132, Met-178, Phe-181, and Phe-185 (13), in our studies only substitution of Phe-181 to either T or A produced monodispersed dimer preparations. Proteins with substitutions at the other positions formed either aggregates or equilibrium mixtures. As expected, when the core-core interface is compromised by the F181T substitution, disruption of NTD interactions by EDTA treatment leads to production of IN monomers (Fig. 6).

The HIV IN F181T reaching dimer is highly stable at high protein concentration and suitable for SAXS analysis (Table 2). Our cross-linking analyses showed that the positions of the terminal domains in the F181T reaching dimer varied somewhat from those in the reaching dimers of wild type HIV IN and ASV dimer. B, shown are two orthogonal views of the tetramer model arranged inside the SAXS determined envelope of the wild type HIV IN tetramer; for clarity, each of the monomers are colored individually. C, shown is a comparison of experimental data for the wild type HIV IN tetramer with the proposed tetramer model and an intasome tetramer model without the viral DNA substrate.
HIV-1 Integrase (Dimer and Tetramer) Solution Structures

IN (3). Although the reason for this difference is unknown, we note that in the HIV IN F181T core domain, a critical hydrophobic amino acid that could interact with residues in the NTD (supplemental Fig. S1) has been replaced with a non-aromatic amino acid. Furthermore, the HIV-1 NTD has four surface-exposed aromatic residues (Phe-1, Tyr-15, Trp-19, and Phe-26) that are available for interaction with another NTD as well as the core and CTD. In contrast, ASV IN lacks aromatic residues in its NTD and may have less propensity to make such contacts.

We speculate, therefore, that the NTD-NTD interaction in HIV IN F181T may reflect a “domain swapping” phenomenon (24) in which an aromatic interface-involving interaction between the NTD with the core domain is substituted by an alternative interaction with the other NTD to obtain maximum hydrophobic packing. It is possible that such swapping may also occur transiently in a wild type reaching dimer.

Modeling an HIV ApoIN Tetramer—Stability of the HIV IN tetramer has been investigated extensively by various groups (13, 20, 25–28). Single amino acid substitutions of particular residues were observed to result in either equilibrium mixtures of dimers with tetramers or pure homo-dimers. Although usually assumed to be core-core dimers (29, 30), the structure of these HIV dimers or the tetramer was unknown. Our SAXS analysis indicates that the wild type IN tetramer has a $D_{\text{max}}$ similar to that of our F181T dimer but that the envelope volume of this tetramer and the F181T IN dimer are 260 and 130 Å³, respectively (Table 2). With the same $D_{\text{max}}$ but twice the volume, the simplest interpretation is that the tetramer comprises two stacked reaching dimers, likely stabilized by core-core interactions at both ends (Fig. 10A). Interestingly our SAXS-derived tetramer model resembles a combined model of Cai et al. (25) that was proposed solely on consideration of structures and interactions of the isolated domains of HIV IN. Our tetramer model is also consistent with the SAXS parameters in the preferred model 1 of a LEDGF(IDB) bound HIV IN tetramer of Gupta et al. (31), as IBD attachment to distally placed catalytic cores can account for the increased $D_{\text{max}}$. Our results also showed that the HIV IN tetramer conformations are only slightly affected by the presence of the metal cofactor Mg$^{2+}$ without any gross change in the overall architecture. The minor changes observed in the $D_{\text{max}}$ values of these two HIV IN proteins (wild type and D64N) in the presence of metal (Fig. 5) might reflect some of the conformational alterations detected in earlier studies (32, 33).

After an analysis of many homo-multimeric enzymes, Powers and Powers (34) have observed that although two alternate dimer assemblies could be envisioned, one dimer form is utilized predominantly in pathways of homo-tetramerization. They also present theoretical reasons for why interconversion of dimeric forms is not permitted without dissociation to monomers and why formation of trimers by monomer addition to a dimer is not observed. The dissociative requirement in the pathway of monomer-dimer-tetramer formation is interesting in light of a previously reported observation that intasome formation is enhanced by IN monomerization before DNA addition (19).

Application of PISA (Proteins Interfaces Surfaces and Assemblies) analysis (35) to our F181T reaching dimer (model A) provides a value for the free energy of formation ($\Delta G$) of this dimer interface of $-15.8$ kcal/mole. This number is similar to the $-14.0$ kcal/mole that can be obtained from the outer core-core dimer interface in a model of the HIV IN intasome (22). Comparable values were obtained from the PFV intasome crystal structure: $\Delta G$ for the NTD-core interface of the PFV inner dimer is $-14.5$ kcal/mol and for the core-core interface in the outer dimer the value is $-15.3$ kcal/mol. Although this calculation only relates to protein-protein interactions in the reaching dimer interface of the PFV NTD + NED with the core residues, it is likely that bound viral DNA contributes to complex formation and stability of the intasome. Overall, the HIV and PFV examples suggest that both dimer interfaces have similar stabilities. These estimates imply that wild type HIV IN can exist in two dimer forms in solution. If one interface has been compromised, the predominant form will be the alternate dimer.

Possible Modes of Intasome Assembly—An HIV IN tetramer model with longer $D_{\text{max}}$ and a smaller $R_g$ has been derived by removal of the DNA substrate from the intasome model (22). Given the existence of flexible linkers from the core to the NTD and CTD and the potential for dynamic motion of these domains, as suggested earlier (3) it is possible that a transient tetramer conformation can favor viral DNA capture by one of the stacked reaching dimers, which then becomes the inner dimer of an intasome that performs catalysis. The terminal domains of the other stacked reaching dimer might simultaneously disengage to assume auxiliary functions. Others have proposed scenarios in which HIV IN subunits exchange and form core-core dimers, which bind DNA and come together to form a tetramer (29), or in which formation of functional intasomes involves the association of free monomers (19), as appears to be the case for PFV IN (2). Further experiments are needed to distinguish among these differing modes of intasome assembly, and the extent to which high salt concentration, a necessary condition of our biophysical analyses, affects the dimer-tetramer equilibrium or the disposition of terminal domains.

ApoHIV IN Is a Viable Drug Target—The lack of the full-length HIV apoIN structures and knowledge of their multimter architectures have been an impediment in the development of non-active site inhibitors. Previous studies with single chain variable fragments derived from NTD- and CTD-specific anti-HIV IN monoclonal antibodies revealed that their intracellular expression can protect human T cells from HIV-1 infection at a stage before integration, implying that the apoIN form is susceptible to drug intervention (36). As IN does not engage its viral DNA substrate productively until reverse transcription completes the viral DNA ends, there is a substantial window of opportunity after infection for compounds that target the apoIN multimers specifically.

Hayouka et al. (37) have demonstrated that LEDGF-derived peptides that bind to the core-core interface of HIV IN inhibit binding of the viral DNA substrate by stabilizing an inactive tetramer conformation. They called such allosteric inhibitors “shiftides” to highlight their ability to shift the equilibrium between active and inactive multimers. More recent studies have shown that small molecules that bind to the LEDGF binding site can also inhibit conformational changes that are
required for DNA engagement at the active site (29, 38–40). We speculate that small molecules that attach to other sites on the apoIN multimer and either increase the stability of an inactive conformation or prevent establishment of the active conformation could also effectively block IN activity. The presence of alternate multimeric IN conformations is consistent with the inclusion of this protein in the morpheme paradigm, which describes a mechanism of enzyme regulation via modulation between alternate, functionally distinct oligomeric assemblies (41, 42). Proteins that assume such alternate conformations represent attractive targets for small molecule allosteric inhibitors (43), which in the case of HIV IN would be effective against emerging, active site inhibitor resistant viruses.

In summary, solution studies including SAXS have proven to be a powerful approach for investigating the flexible 3-domain IN proteins. In this study SAXS analysis in combination with chemical cross-linking and modeling has revealed the modular architecture of HIV IN multimers and suggested the most likely mode of tetramer assembly. The findings have provided new insight into IN protein structure and function and suggested valuable new strategies for inhibition.

Acknowledgments—We thank our colleagues Drs. Richard A. Katz, Eileen Jaffe, and Heinrich Röder for helpful comments regarding the experiments reported in this manuscript and Marie Estes for assistance in its preparation. Use of Advanced Photon Source was supported by the United States Department of Energy, Office of Science, Office of Basic Energy Sciences (DE-AC02–06CH11357) under general user proposal access. Our studies also benefited from access to the Office of Basic Energy Sciences (DE-AC02–06CH11357) under general user proposal access. Our studies also benefited from access to the Office of Basic Energy Sciences (DE-AC02–06CH11357) under general user proposal access. Our studies also benefited from access to the Office of Basic Energy Sciences (DE-AC02–06CH11357) under general user proposal access.

REFERENCES

1. Craigie, R. (2012) The molecular biology of HIV integrase. *Future Virol.* 7, 679–686
2. 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
3. Bojja, R. S., Andrake, M. D., Weigand, S., Merkel, G., Yarychkivska, O., Henderson, A., Kummerling, M., and Skalka, A. M. (2011) Architecture of a full-length retroviral intasome monomer and dimer, revealed by small angle x-ray scattering and chemical cross-linking. *J. Biol. Chem.* 286, 17047–17059
4. Grabarek, Z., and Gergely, J. (1990) Zero-length cross-linking procedure with the use of active esters. *Anal. Biochem.* 185, 131–135
5. Kessl, J. J., Eidahl, J. O., Shkriabai, N., Zhao, Z., McKee, C. J., Hess, S., Burke, T. R., Jr., and Kvaratskhelia, M. (2009) An allosteric mechanism for inhibiting HIV-1 integrase with a small molecule. *Mol. Pharmacol.* 76, 824–832
6. de Vries, S. J., van Dijk, M., and Bonvin, A. M. (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* 5, 883–897
7. 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
8. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL repository and associated resources. *Nucleic Acids Res.* 37, D387–D392
9. Peitsch, M. C. (1995) Protein modeling by e-mail. *Nat. Biotechnol.* 13, 658–660
10. Svergun, D. I., Petaukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from x-ray solution scattering. *Biophys. J.* 80, 2946–2953
11. Ilavsky, J., and Jemian, P. R. (2009) Irena: tool suite for modeling and analysis of small-angle scattering. *J. Appl. Crystallogr.* 42, 347–353
12. Craigie, R., Mizuuchi, K., Bushman, F. D., and Engelman, A. (1991) A rapid in vitro assay for HIV DNA integration. *Nucleic Acids Res.* 19, 2729–2734
13. Al-Mawusi, L. Q., Hombrouck, A., Dayam, R., Debyser, Z., and Neamati, N. (2008) Four-tiered $\pi$ interaction at the dimeric interface of HIV-1 integrase critical for DNA integration and viral infectivity. *Virology* 377, 355–363
14. Serrao, E., Thys, W., Demeulemeester, J., Al-Mawusi, L. Q., Christ, F., Debyser, Z., and Neamati, N. (2012) A symmetric region of the HIV-1 integrase dimerization interface is essential for viral replication. *PLoS One* 7, e45177
15. Faure, A., Calmels, C., Desjobert, C., Castroviejo, M., Caumont-Sarcos, A., Tarrago-Litvak, L., Litvak, S., and Parissi, V. (2005) HIV-1 integrase cross-linked oligomers are active in vitro. *Nucleic Acids Res.* 33, 977–986
16. Guiot, E., Carayon, K., Deléis, O., Simon, F., Tauc, P., Zubin, E., Gottikh, M., Mouscadet, J. F., Brochon, J. C., and Dexpert, P. (2006) Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity. *J. Biol. Chem.* 281, 22707–22719
17. Bushman, F. D., Engelman, A., Palmer, L., Wingfield, P., and Craigie, R. (1993) Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotid transfer and zinc binding. *Proc. Natl. Acad. Sci. U.S.A.* 90, 3428–3432
18. Lee, S. P., Xiao, J., Knutson, J. R., Lewis, M. S., and Han, M. K. (1997) Zn$^{2+}$ promotes the self-association of human immunodeficiency virus type 1 integrase in vitro. *Biochemistry* 36, 173–180
19. Pandey, K. K., Bera, S., and Grandgenett, D. P. (2011) The HIV-1 integrase monomer induces a specific interaction with LTR DNA for concerted integration. *Biochemistry* 50, 9788–9796
20. 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
21. Xu, Q., and Dunbrack, R. L., Jr. (2011) The protein common interface database (ProtCID). A comprehensive database of interactions of homologous proteins in multiple crystal forms. *Nucleic Acids Res.* 39, D761–D770
22. 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. U.S.A.* 107, 15910–15915
23. Gupta, K., Curtis, J. E., Krueger, S., Hwang, Y., Cherepanov, P., Bushman, F. D., and Van Duyne, G. D. (2012) Solution conformations of prototype foamy virus integrase and its stable synaptic complex with U5 viral DNA. *Structure* 20, 1918–1928
24. Gronenborn, A. M. (2009) Protein acrobatics in pairs. Dimerization via domain swapping. *Curr. Opin. Struct. Biol.* 19, 39–49
25. 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
26. Chen, J. C., Krucinski, J., Mierie, L. J., Finer-Moore, J. S., Tang, A. H., Leavitt, A. D., and Stroud, R. M. (2012) Solution structure of prototype foamy virus integrase and its stable synaptic complex with U5 viral DNA. *Structure* 20, 1918–1928
27. Lutzke, R. A., and Plasterk, R. H. (1998) Solution-based mutational analysis of the C-terminal DNA binding domain of human immunodeficiency virus type 1 integrase. Critical residues for protein oligomerization and DNA binding. *J. Biol. Chem.* 273, 3953–3963
28. 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
29. Kessl, J. J., Li, M., Ignatov, M., Shkriabai, N., Eidahl, J. O., Feng, L., Musier-Forsyth, K., Craigie, R., and Kvaratskhelia, M. (2011) FRET analysis reveals distinct conformations of IN tetramers in the presence of viral DNA or LEDGF/p75. *Nucleic Acids Res.* 39, 9009–9022
30. 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

31. 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

32. Asante-Appiah, E., Seeholzer, S. H., and Skalka, A. M. (1998) Structural determinants of metal-induced conformational changes in HIV-1 integrase. *J. Biol. Chem.* **273**, 35078–35087

33. Asante-Appiah, E., and Skalka, A. M. (1997) A metal-induced conformational change and activation of HIV-1 integrase. *J. Biol. Chem.* **272**, 16196–16205

34. Powers, E. T., and Powers, D. L. (2003) A perspective on mechanisms of protein tetramer formation. *Biophys. J.* **85**, 3587–3599

35. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797

36. Levy-Mintz, P., Duan, L., Zhang, H., Hu, B., Dornadula, G., Zhu, M., Kulkosky, J., Bizub-Bender, D., Skalka, A. M., and Pomerantz, R. J. (1996) Intracellular expression of single-chain variable fragments to inhibit early stages of the viral life cycle by targeting human immunodeficiency virus type 1 integrase. *J. Virol.* **70**, 8821–8832

37. Hayouka, Z., Rosenbluh, J., Levin, A., Loya, S., Lebendiker, M., Veprintsev, D., Kotler, M., Hizi, A., Loyter, A., and Friedler, A. (2007) Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 8316–8321

38. Christ, F., Voet, A., Marchand, A., Nicolet, S., Desimmie, B. A., Marchand, D., Bardiott, D., Van der Veken, N. J., Van Remoortel, B., Strelkov, S. V., De Maeyer, M., Chaltin, P., and Debyser, Z. (2010) Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat. Chem. Biol.* **6**, 442–448

39. Taltynov, O., Desimmie, B. A., Demeulemeester, J., Christ, F., and Debyser, Z. (2012) Cellular cofactors of lentiviral integrase. From target validation to drug discovery. *Mol. Biol. Int.* 863405

40. Tintori, C., Demeulemeester, J., Franchi, L., Massa, S., Debyser, Z., Christ, F., and Botta, M. (2012) Discovery of small molecule HIV-1 integrase dimerization inhibitors. *Bioorg. Med. Chem. Lett.* **22**, 3109–3114

41. Jaffe, E. K. (2005) Morpheeins. A new structural paradigm for allosteric regulation. *Trends Biochem. Sci.* **30**, 490–497

42. Selwood, T., and Jaffe, E. K. (2012) Dynamic dissociating homo-oligomers and the control of protein function. *Arch. Biochem. Biophys.* **519**, 131–143

43. Lawrence, S. H., Ramirez, U. D., Tang, L., Fazliyez, F., Kundrat, L., Markham, G. D., and Jaffe, E. K. (2008) Shape shifting leads to small-molecule allosteric drug discovery. *Chem. Biol.* **15**, 586–596

44. Merkel, G., Andrake, M. D., Ramcharan, J., and Skalka, A. M. (2009) Oligonucleotide-based assays for integrase activity. *Methods* **47**, 243–248

45. Li, M., and Craigie, R. (2005) Processing of viral DNA ends channels the HIV-1 integration reaction to concerted integration. *J. Biol. Chem.* **280**, 29334–29339