Protein & Cell

COMMUNICATION

Crystal structures of catalytic core domain of BIV integrase: implications for the interaction between integrase and target DNA

Xue Yao[1,2*], Shasha Fang[1*], Wentao Qiao[2], Yunqi Geng[2 ✉], Yuequan Shen[1 ✉]

[1] Tianjin Key Laboratory of Protein Science, College of Life Sciences, Nankai University, 94 Weijin Road, Tianjin 300071, China
[2] Key Laboratory of Molecular Microbiology and Biotechnology (Ministry of Education), the College of Life Science, Nankai University, 94 Weijin Road, Tianjin 300071, China
✉ Correspondence: gengyq@nankai.edu.cn (Y. Geng), yshen@nankai.edu.cn (Y. Shen)

ABSTRACT

Integrase plays a critical role in the recombination of viral DNA into the host genome. Therefore, over the past decade, it has been a hot target of drug design in the fight against type 1 human immunodeficiency virus (HIV-1). Bovine immunodeficiency virus (BIV) integrase has the same function as HIV-1 integrase. We have determined crystal structures of the BIV integrase catalytic core domain (CCD) in two different crystal forms at a resolution of 2.45 Å and 2.2 Å, respectively. In crystal form I, BIV integrase CCD forms a back-to-back dimer, in which the two active sites are on opposite sides. This has also been seen in many of the CCD structures of HIV-1 integrase that were determined previously. However, in crystal form II, BIV integrase CCD forms a novel face-to-face dimer in which the two active sites are close to each other. Strikingly, the distance separating the two active sites is approximately 20 Å, a distance that perfectly matches a 5-base pair interval. Based on these data, we propose a model for the interaction of integrase with its target DNA, which is also supported by many published biochemical data. Our results provide important clues for designing new inhibitors against HIV-1.

KEYWORDS   bovine immunodeficiency virus, integrase, catalytic core domain, crystal structure, dimerization

INTRODUCTION

Integration is an obligate step in productive retrovirus infection (Weidhaas et al., 2000; Maxfield et al., 2005). Integration of the retroviral DNA into the host cell genome requires the interaction of the retroviral integrase protein with the outer ends of both viral long-terminal repeats (LTRs), thereby removing two nucleotides from the 3′ ends (3′ processing), and the newly created 3′ OH attacking a phosphate of the target DNA (strand transfer) (Engelman et al., 1991; Craigie, 2001; Vandegraaff and Engelman, 2007; Delelis et al., 2008). Integrase usually consists of a catalytic core domain (CCD) that is responsible for catalytic activities and N- and C-terminal domains that play roles mainly in the binding of substrate to DNA (Ceccherini-Silberstein et al., 2009). Several studies have shown that the entire integrase protein can form a tetramer that is required to carry out the full biologic function, whereas the CCD domains or other oligomeric forms of integrases retain only partial function (Cherepanov et al., 2003; Hayouka et al., 2007; Ren et al., 2007).

To uncover the mechanism of integration at the molecular level, many structural studies of integrase have been carried out. Thus far, crystal structures of partial integrase structures, including the CCD domain, the CCD domain plus the N-terminal region and the CCD domain plus the C-terminal region, have been solved (Dyda et al., 1994; Goldgur et al., 1998; Yang et al., 2000; Wang et al., 2001; Cherepanov et al., 2005). Based on the above information, a model of the

---

structure of full-length integrase was deduced (Wang et al., 2001). Interestingly, all of the integrase structures show that the CCDs form a very stable dimer, rather than a tetramer. Moreover, the active site of each CCD is on the opposite side (i.e., the back-to-back mode) (Goldgur et al., 1998; Yang et al., 2000; Wang et al., 2001; Cherepanov et al., 2005). This result was able to explain the 3′ processing function of integrase, but did not give any information about the interaction of integrase with its target DNA. DNA transposase is the recombinase that is most related to integrase (Czyz et al., 2007; Reznikoff, 2008). The crystal structure of prokaryotic Tn5 transposase complexed with Tn5 transposon end DNA has been solved, and it revealed that the active site of the dimer was arranged in the face-to-face mode and that the target DNA presumably bound in the middle of the dimer (Davies et al., 2000). The most recent crystal structure of eukaryotic Mos1 transposase in complex with transposon end DNA also showed face-to-face dimeric active sites. Once again, the target DNA was proposed to bind in the middle of the dimer (Richardson et al., 2009). Based on the homology between transposase and integrase, it is reasonable to assume that two integrase dimers may assemble using the face-to-face mode, as in the tetramer. However, such a face-to-face arrangement has not been truly observed in the integrase dimer and therefore the model of the fully functional integrase tetramer and its interaction with target DNA remain controversial. Based on crystal packing data from the crystal structure of type-1 human immunodeficiency virus (HIV-1) CCD plus N-terminal domain, an integrase tetramer model has been proposed (Wang et al., 2001). However, the two active sites in the face-to-face dimer were separated by around 40 Å, which was much longer than the five-base pair spacing of the insertion sites into the target DNA (Wang et al., 2001). The latest cryo-electron microscopic studies of full-length integrase with its cellular cofactor LEDGF/p75 plus viral DNAs showed another tetramer model which was very different from others that have been proposed previously (Michel et al., 2009).

Bovine immunodeficiency virus (BIV) integrase shares a 35% sequence identity with and has similar enzymatic activities to HIV-1 integrase (3′ end processing, strand transfer and disintegration of the proviral DNA) (Cherepanov, 2007; Avidan and Hizi, 2008). The catalytic reaction of BIV integrase is absolutely dependent on the presence of the divalent cations, Mn^{2+} or Mg^{2+}. In addition, the protein is found to be active toward oligonucleotides containing sequences that are derived from the U5 end of HIV-1 (Avidan and Hizi, 2008). BIV integrase also interacts with the host factor LEDGF/p75, as do HIV and other lentiviruses (Cherepanov, 2007).

Herein, we present two different structures of dimeric CCDs (back-to-back and face-to-face) in two different crystals. Moreover, the active site separation in the face-to-face dimer is approximately 20 Å, which fits perfectly with

the five-base pair interval of the target DNA. Based on these results, we propose and discuss a model of the interaction between integrase and its target DNA.

RESULTS AND DISCUSSION

The overall structure of BIV integrase CCD

BIV integrase CCD proteins were crystallized in two crystal forms. Crystal form I had one molecule per asymmetric unit and crystal form II had two molecules per asymmetric unit. The overall structure of BIV integrase CCD has five β strands flanked by six α-helices (Fig. 1A). The structure is quite similar to many integrase CCD structures that have been reported previously from HIV-1 or other species (Fig. 1B and 1C). The superimposition of integrase CCD of BIV onto that of HIV gives root-mean-square deviation (rmsd) value of 1.4 Å for 147 Cα atoms. The most important conformational difference is in the active site loop (Fig. 1C), reflecting its conformational flexibility (Greenwald et al., 1999). This loop may be stabilized upon substrate DNA binding. High sequence similarity and three-dimensional structural conservation indicate that BIV integrase probably has a catalytic mechanism similar to that of its HIV counterpart.

The final refined structures show that these three molecules are very similar to each other. Superimposing crystal form I onto two crystal form II molecules gives rmsd values of 0.56 Å and 0.65 Å. The main difference between the structures arises from the variable conformation of loop α5–α6. All except two c-terminal amino acids in one of the asymmetric molecules of crystal form II are clearly visible in the final density map. The catalytic center consists of three conserved residues, Asp64, Asp121 and Glu157. No electron density corresponding with metal ion is observed in the catalytic center.

Comparison of the two different dimers

Crystal form I of the BIV integrase CCD molecule forms a back-to-back dimer (Fig. 2A) produced with a crystallographic 2-fold axis, in which the two catalytic loops face away from the dimeric interface. However, in crystal form II, two non-crystallographic 2-fold axis-related BIV integrase CCD molecules form a face-to-face dimer in one asymmetric unit (Fig. 2B), in which the two catalytic loops are close to each other. We note that each monomer in the same asymmetric unit of crystal form II can also generate back-to-back dimers using a crystallographic 2-fold axis. The buried surface area (Lee and Richards, 1971) is 2648.3 Å² for the back-to-back dimer and 1,168.7 Å² for the face-to-face dimer. Within the back-to-back dimer interface, the interaction residues are contributed mainly by helices α1, α3, α4 and α5 of each monomer, and many hydrogen bonds are formed (Fig. 2C). In sharp contrast, the interactions within the interface of the
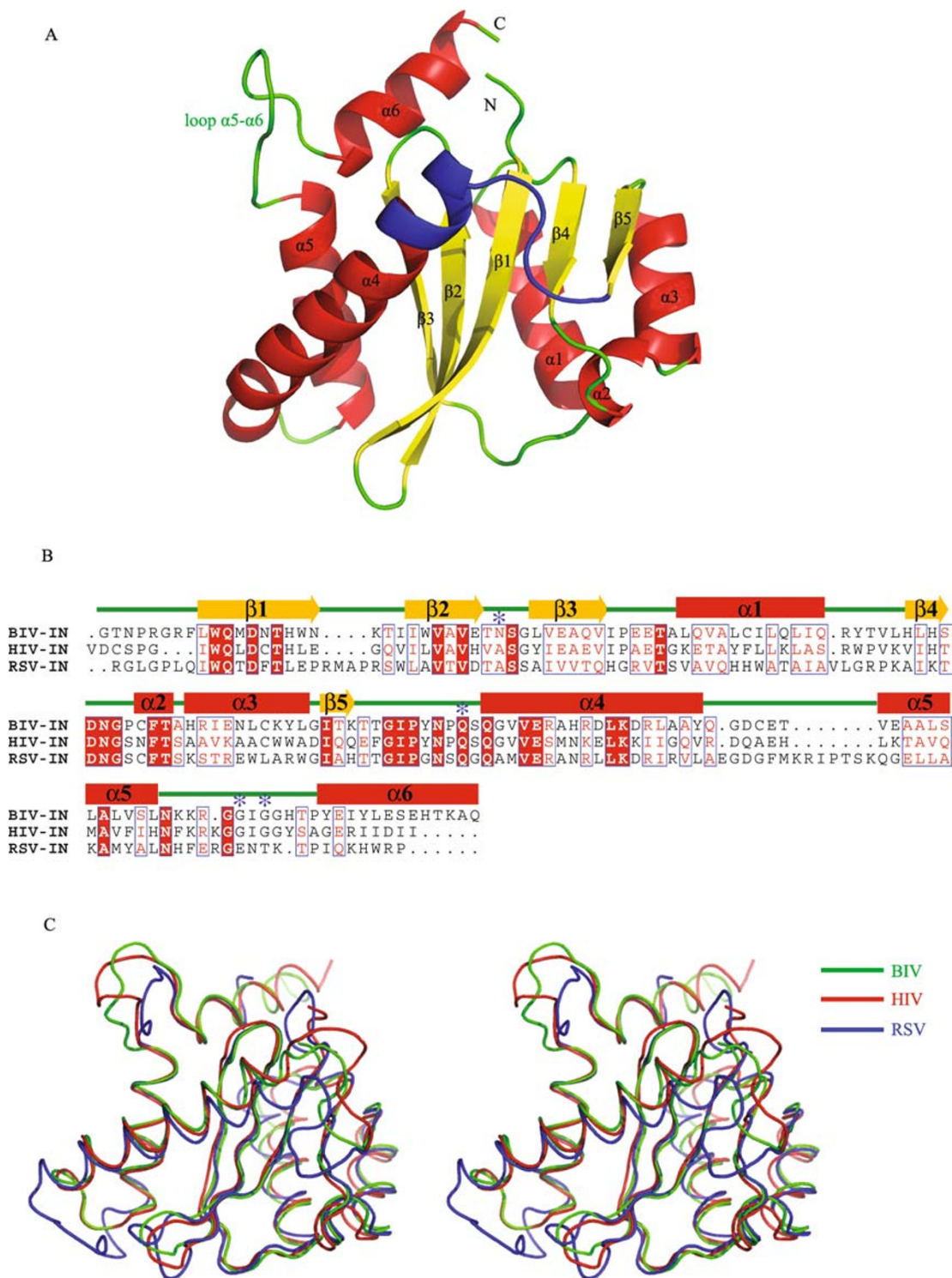
Figure 1.   The crystal structure of the BIV integrase CCD monomer. (A) Cartoon representation of the BIV integrase CCD monomer. The α-helices, β-strands and loops are colored red, yellow and green, respectively. The activation loop (i.e., amino acids 140–149) is colored blue. (B) The structure-based sequence alignment of the CCD of BIV (UniProtKB P19560), HIV (UniProtKB P12497) and RSV (UniProtKB P03354). The secondary structure elements for BIV integrase CCD are shown as cylinders (α-helices) and arrows (β-strands). Amino acids that are conserved across three proteins are highlighted in red. Four residues making hydrogen bonds within the interface of face-to-face dimer are asterisked. (C) Stereoview of the superimposition of BIV integrase CCD (green) onto that of HIV (red, PDB ID 1BIS) and RSV (blue, PDB ID 1C1A).
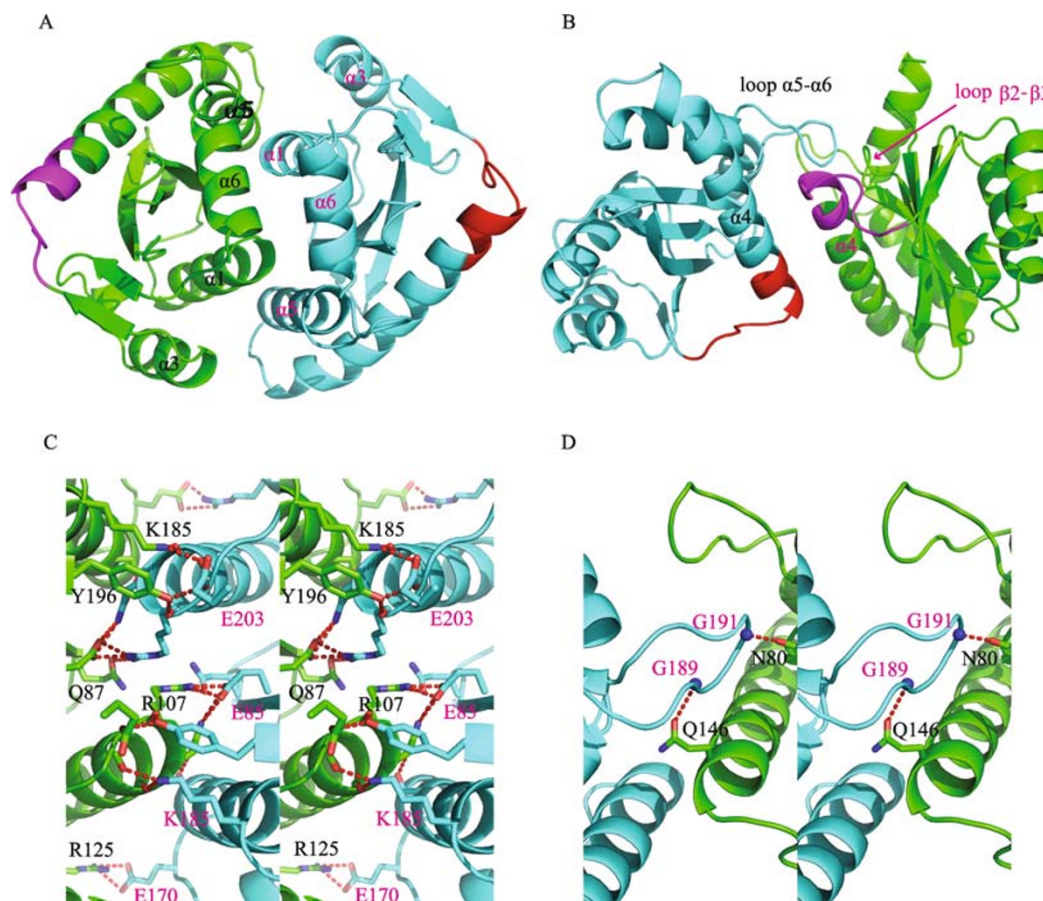
Figure 2. Comparison of the two different dimers. Cartoon representations of the back-to-back (A) and face-to-face (B) BIV integrase CCD dimers are shown. Monomers A and B are colored green and cyan, respectively, and the active sites in monomers A and B are colored magenta and red, respectively. We also show a stereoview of the detailed interactions at the interface of the back-to-back dimer (C) and the face-to-face dimer (D). Each monomer is shown as a cartoon representation and monomers A and B are colored green and cyan, respectively. The amino acid side chains are shown as stick representations. The oxygen and nitrogen atoms are colored red and blue, respectively. The carbon atoms are colored following the color of the monomer. Hydrogen bonds are shown as red dashed lines. The labels are colored black and magenta to distinguish them from the two monomers.

face-to-face dimer occur between helix α4 and loop α5–α6 of one monomer and between helix α4 and loop β2–β3 of another monomer. Two hydrogen bonds are found at the interface of the face-to-face dimer (Fig. 2D). It appears that the back-to-back dimer has stronger interactions than the face-to-face dimer. Consequently, the back-to-back dimer has been observed in many integrase crystal structures reported previously from various species. The face-to-face dimer has not been reported previously.

A model of the interaction of BIV integrase CCD with target DNA

The new face-to-face BIV integrase CCD dimer provides a possibility that the target DNA could bind in the middle of the dimer. We first superimposed molecule A of the crystal structure of the HIV-1 CCD domain (PDB code 1K6Y) onto

each monomer of our face-to-face BIV integrase CCD dimer. At this point, we included the phosphate ion and two water molecules from the active site of the HIV-1 structure to mimic the phosphates of the target DNA backbone and two metal ions in each monomer of BIV integrase CCD dimer (Wang et al., 2001). Two phosphate ions are separated by 28.9 Å in the BIV integrase CCD face-to-face dimer, which corresponds well with a seven-base pair interval (29.5 Å). Finally, the DNA template model was taken from the structure of the Tn5 transposon in complex with DNA (PDB code 1F3I) and manually docked into position using the position of two phosphate ions as a guide (Fig. 3A). Since no close contact was found in our model of the complex, we ignored the minimization refinement. This led to our final model of the interaction between the BIV integrase CCD face-to-face dimer and its target DNA (Fig. 3A and 3B). In this model, one magnesium ion stabilizes the neighboring phosphate,
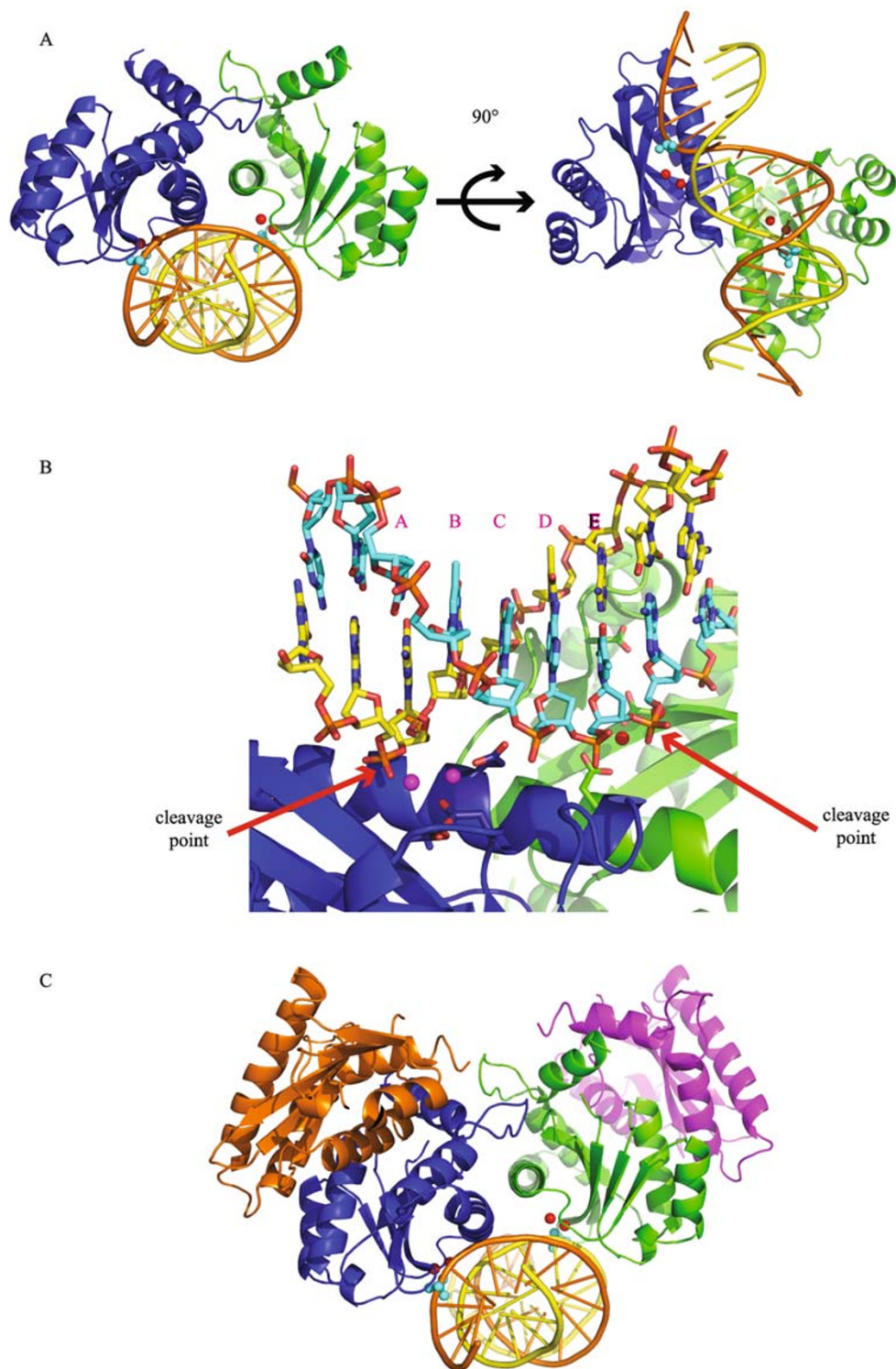
Figure 3.   Our proposed model of the interaction of integrase with target DNA. (A) Cartoon representation of the face-to-face dimer binding with target DNA. The two monomers are colored blue and green. Two magnesium ions in each monomer are shown as red spheres. Each phosphate molecule is shown as cyan sphere. The double helices of the DNA template are colored orange and yellow, respectively. The base pairs are represented by sticks. (B) Details of the binding interactions of the target DNA in the active sites of the dimer. The DNA template is represented using sticks. The carbon atoms are colored cyan and yellow in the different chains. The oxygen, nitrogen and phosphorus atoms are colored red, blue and orange, respectively. Magnesium ions are shown as red and magenta spheres. (C) Cartoon representation of BIV integrase CCD tetramer binding with target DNA. Four CCDs are colored as orange, blue, green and magenta, respectively.

and a second magnesium ion can potentially activate a water molecule to hydrolyze the DNA backbone. Therefore, two active sites can accommodate and cleave the DNA at a 5-base pair interval (Fig. 3B). This result appears to be consistent with the experimental observation that BIV or HIV-1 integrase inserted two viral DNA ends into the target DNA at a spacing of 5 base pairs (Ellison et al., 1990). The final model of BIV integrase CCD tetramer interacted with target DNA is proposed in Fig. 3C.

Based on the proposed structure, we speculate that the two viral DNA ends, when bound to integrase, assemble a tetramer that interacts with the target DNA. This model is quite similar to the model that had been proposed for the HIV-1 integrase tetramer in complex with viral and target DNAs (Gao et al., 2001; Zhao et al., 2008). The interaction of integrase with target DNA that we have proposed is also similar to the model that had been proposed for the interaction of eukaryotic Mos1 transposase with its target DNA (Richardson et al., 2009). Our proposed model also fully agrees with previously published biochemical data. In our model, loop α5–α6 (residues 186–194) is involved in integrase multimerization (Fig. 2B and 2D), which is consistent with the biochemical data that residues 186–188 played important role in HIV-1 intergrase oligomerization (Berthoux et al., 2007) and the data that the deletion of residues 190–192 abrogated multimerization (Hare et al., 2009). In our model, the active site loop (residues 140–149) directly interacts with target DNA (Fig. 2B and 3A). The molecule dynamics simulation results of full-length integrase-DNA complex suggested that the catalytic loop played crucial role during integration process (De Luca et al., 2005). It had been reported that the residue Gln148 of HIV-1 integrase was critical for strand transfer (Johnson et al., 2006).

Our model may not accurately reflect the interaction of the full-length integrase with target DNA since only CCD dimer is used in our studies and the conformation changes upon binding are hard to predict. Nevertheless, our results may represent a starting point toward the identification of the bona fide integrase oligomeric state. This may provide important clues in the search for more efficient inhibitors in the fight against HIV.

MATERIALS AND METHODS

Construction of the BIV integrase-encoding plasmid

A DNA fragment containing BIV integrase CCD (60–208) was amplified by PCR, using the BIV proclone, BIV127 (Avidan and Hizi, 2008), as template, and the primers 5′-GGTGAATTCTTATGG-CAGATGGACAATAC (forward) and 5′-GTTGTCGACTTATTGG-TATTTGGTATGTTC (reverse). The desired PCR fragment was then digested using *Eco*RI and *Sal*I restriction enzymes and cloned into the pETH vector (a derivative of pET32a). The correctness of the BIV integrase CCD construct was confirmed by DNA sequencing.

Protein expression and purification

Plasmids carrying the BIV integrase CCD gene were transformed into *Escherichia coli* BL21 (DE3) cells harboring pUBS520. The resulting cells were grown in Luria broth medium containing 50 μg/mL ampicillin and 30 μg/mL kanamycin at 37°C until A_{600} = 1.8, when the cells were induced by adding isopropyl-1-thiogalactopyranoside to a final concentration of 20 μM and harvested after an overnight post-induction incubation at 18°C. The *E. coli* cells were lysed in $T_{20}\beta_2 N_{100} P_{0.1}$ (20 mM Tris-HCl, pH 7.5, 2 mM β-ME, 100 mM NaCl and 0.1 mM PMSF) using lysozyme (0.1 mg/mL) and sonication. After the lysate was centrifuged at 100,000 g for 60 min at 4°C, the supernatant from the *E. coli* lysate was loaded directly onto a Ni^{2+}-NTA column that had been equilibrated with $T_{20}\beta_2 N_{100} P_{0.1}$. The Ni^{2+}-NTA column was then washed with five column volumes of $T_{20}\beta_2 N_{100} P_{0.1}$. The His_6-tagged BIV integrase CCD was eluted with $T_{20}\beta_2 N_{100} P_{0.1}$ buffer containing 200 mM Imidazole. The eluate was concentrated and loaded onto a gel filtration column with $T_{20} D_2 E_1 N_{100}$ buffer (20 mM Tris-HCl, pH 7.5, 2 mM DTT, 1 mM EDTA, 100 mM NaCl). The elution profile indicates that BIV integrase CCD in solution is a dimer predominantly. Using this method, we achieved a purity of greater than 95%.

Crystallization and data collection

BIV integrase CCD crystals were grown using vapor diffusion at 20°C in a hanging drop format. BIV integrase CCD (1 μL of 10 mg/mL protein in 20 mM Tris-HCl, pH 7.5, 2 mM DTT, 1 mM EDTA, 100 mM NaCl) was mixed with 1 μL of a well solution. Crystal form I was obtained using a well solution containing 0.9 M ammonium sulfate, 0.02 M HEPES, pH 7.5, and 0.7 M $NaNO_3$. Crystal form II was obtained using another well solution containing 17% PEG 10,000, 0.1 M sodium cacodylate, pH 6.5, and 0.1 M sodium acetate. Two types of crystals grew over one week. These were frozen in a cryoprotectant that consisted of the reservoir solution supplemented with 25% (*v/v*) glycerol. Diffraction data were collected in-house at 100K using an R-axis HTC detector (Rigaku) and CuKα radiation, and the data were processed using HKL2000 (Otwinowski and Minor, 1997).

Structural determination and refinement

Crystal form I of the BIV integrase CCD diffracted to 2.4 Å and belonged to the space group $P4_1 2_1 2$ with unit cell parameters $a = b$ = 82.32 Å and c = 51.06 Å. Crystal form II of BIV integrase CCD diffracted to 2.2 Å and belonged to the space group $P4_1 2_1 2$ with unit cell parameters $a = b$ = 88.05 Å and c = 81.78 Å. The initial phase was obtained by molecular replacement using the structure of HIV-1 integrase CCD (PDB code 1BIS) as a template. The program PHASER (McCoy, 2007) was able to locate one molecule in the asymmetric unit from the data of crystal form I and two molecules from the data of crystal form II. The models were built using the program COOT (Emsley and Cowtan, 2004) and refined using the programs CNS (Brünger et al., 1998) and PHENIX (Terwilliger et al., 2008). The final structure had an $R_{crystal}$ value of 19.2% and an R_{free} value of 24.2% for crystal form I and an $R_{crystal}$ value of 23.9% and an R_{free} value of 27.2% for crystal form II. Detailed data collection and refinement statistics are summarized in Table 1.

Table 1 Crystallographic statistics for the BIV integrase CCD structures

crystal name	form 1	form 2
space group	$P4_12_12$	$P4_12_12$
unit cell (Å)	$a = b = 82.32$, $c = 51.06$	$a = b = 88.05$, $c = 81.78$
beamline	in-house	in-house
wavelength (Å)	1.5418	1.5418
resolution range (Å)	50~2.45 (2.54−2.45)[b]	50~2.2 (2.28−2.2)[b]
No. of unique reflections	6,827	16,868
redundancy	13.5 (14.0)[b]	12.7 (13.0)[b]
R_{sym} (%)[a]	6.3 (23.3)[b]	10.3 (50.0)[b]
I/σ	40.8 (14.0)[b]	26.5 (6.6)[b]
completeness (%)	99.8 (100.0)[b]	99.8 (99.9)[b]
refinement		
$R_{crystal}$ (%)[c]	19.2	23.9
R_{free} (%)[d]	24.2	27.2
RMSD$_{bond}$ (Å)	0.006	0.006
RMSD$_{angle}$(°)	0.9	1.3
number of		
protein atoms	1,195	2,366
ligand atoms[e]	19	20
solvent atoms	123	93
residues in (%)		
most favored	89.5	94.8
additionally allowed	9.8	4.5
generously allowed	0.8	0.7
disallowed	0	0
average B factor (Å^2) of		
chain A	39.1	33.6
chain B	N/A	33.4

[a] $R_{sym} = \sum_j |\langle I \rangle - I_j| / \sum \langle I \rangle$

[b] The highest resolution shell.

[c] $R_{crystal} = \sum_{hkl} |F_{obs} - F_{calc}| / \sum_{hkl} F_{obs}$

[d] R_{free} was calculated in the same way as $R_{crystal}$, but using a test set containing 5% of data that had been excluded from the refinement calculations.

[e] Nitrate ions in crystal form 1, and acetate ions and glycerol molecules in crystal from 2 were included into the final refined structures.

PROTEIN DATA BANK ACCESSION CODES

The atomic coordinates and structure factors for the structure of BIV integrase CCD have been deposited in the PDB. The accession code for crystal form I is 3KKR, and the code for crystal form II is 3KKS.

REFERENCES

Avidan, O., and Hizi, A. (2008). Expression and characterization of the integrase of bovine immunodeficiency virus. Virology 371, 309–321.

Berthoux, L., Sebastian, S., Muesing, M.A., and Luban, J. (2007). The role of lysine 186 in HIV-1 integrase multimerization. Virology 364, 227–236.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macro-molecular structure determination. Acta Crystallogr. Sect D: Biol Crystallogr 54, 905–921.

Ceccherini-Silberstein, F., Malet, I., D'Arrigo, R., Antinori, A., Marcelin, A.G., and Perno, C.F. (2009). Characterization and structural analysis of HIV-1 integrase conservation. AIDS Rev 11, 17–29.

Cherepanov, P. (2007). LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. Nucleic Acids Res 35, 113–124.

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.

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 U S A 102, 17308–17313.

Craigie, R. (2001). HIV integrase, a brief overview from chemistry to therapeutics. J Biol Chem 276, 23213–23216.

Czyz, A., Stillmock, K.A., Hazuda, D.J., and Reznikoff, W.S. (2007). Dissecting Tn5 transposition using HIV-1 integrase diketoacid inhibitors. Biochemistry 46, 10776–10789.

Davies, D.R., Goryshin, I.Y., Reznikoff, W.S., and Rayment, I. (2000). Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. Science 289, 77–85.

De Luca, L., Vistoli, G., Pedretti, A., Barreca, M.L., and Chimirri, A. (2005). Molecular dynamics studies of the full-length integrase-DNA complex. Biochem Biophys Res Commun 336, 1010–1016.

Delelis, O., Carayon, K., Saïb, A., Deprez, E., and Mouscadet, J.F. (2008). Integrase and integration: biochemical activities of HIV-1 integrase. Retrovirology 5, 114.

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.

Ellison, V., Abrams, H., Roe, T., Lifson, J., and Brown, P. (1990). Human immunodeficiency virus integration in a cell-free system. J Virol 64, 2711–2715.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. . Acta Crystallogr, Sect D: Biol Crystallogr 60, 2126–2132.

Engelman, A., Mizuuchi, K., and Craigie, R. (1991). HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67, 1211–1221.

Gao, K., Butler, S.L., and Bushman, F. (2001). Human immunode-ficiency virus type 1 integrase: arrangement of protein domains in active cDNA complexes. EMBO J 20, 3565–3576.

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 U S A 95, 9150–9154.

Greenwald, J., Le, V., Butler, S.L., Bushman, F.D., and Choe, S. (1999). The mobility of an HIV-1 integrase active site loop is correlated with catalytic activity. Biochemistry 38, 8892–8898.

Hare, S., Di Nunzio, F., Labeja, A., Wang, J., Engelman, A., Cherepanov, P., and Luban, J. (2009). Structural basis for functional tetramerization of lentiviral integrase. PLoS Pathog 5, e1000515.

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.

Johnson, A.A., Santos, W., Pais, G.C., Marchand, C., Amin, R., Burke, T.R. Jr, Verdine, G., and Pommier, Y. (2006). Integration requires a specific interaction of the donor DNA terminal 5′-cytosine with glutamine 148 of the HIV-1 integrase flexible loop. J Biol Chem 281, 461–467.

Lee, B., and Richards, F.M. (1971). The interpretation of protein structures: estimation of static accessibility. J Mol Biol 55, 379–400.

Maxfield, L.F., Fraize, C.D., and Coffin, J.M. (2005). Relationship between retroviral DNA-integration-site selection and host cell transcription. Proc Natl Acad Sci U S A 102, 1436–1441.

McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with phaser. Acta Crystallogr D Biol Crystallogr 63, 32–41.

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.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 276, 307–326.

Ren, G., Gao, K., Bushman, F.D., and Yeager, M. (2007). Single-particle image reconstruction of a tetramer of HIV integrase bound to DNA. J Mol Biol 366, 286–294.

Reznikoff, W.S. (2008). Transposon Tn5. Annu Rev Genet 42, 269–286.

Richardson, J.M., Colloms, S.D., Finnegan, D.J., and Walkinshaw, M. D. (2009). Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. Cell 138, 1096–1108.

Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N. W., Zwart, P.H., Hung, L.W., Read, R.J., and Adams, P.D. (2008). Iterative model building, structure refinement and density mod-ification with the PHENIX AutoBuild wizard. Acta Crystallogr 64, 61–69.

Vandegraaff, N., and Engelman, A. (2007). Molecular mechanisms of HIV integration and therapeutic intervention. Expert Rev Mol Med 9, 1–19.

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.

Weidhaas, J.B., Angelichio, E.L., Fenner, S., and Coffin, J.M. (2000). Relationship between retroviral DNA integration and gene expres-sion. J Virol 74, 8382–8389.

Yang, Z.N., Mueser, T.C., Bushman, F.D., and Hyde, C.C. (2000). Crystal structure of an active two-domain derivative of Rous sarcoma virus integrase. J Mol Biol 296, 535–548.

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 rearran-gements and a role for N-terminal Lys-14 of HIV-1 Integrase during viral DNA binding. J Biol Chem 283, 5632–5641.