J. Biochem. **139**, 753–759 (2006)
doi:10.1093/jb/mvj085

Structural Analysis of a Mutant of the HIV-1 Integrase Zinc Finger Domain That Forms a Single Conformation

Yusuke Nomura[1], Takao Masuda[2] and Gota Kawai[1,*]

[1]Department of Life and Environmental Sciences, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino, Chiba 275-0016; and [2]Department of Immunotherapeutics, Medical Research Division, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519

Received December 3, 2005; accepted February 20, 2006

HIV-1 integrase consists of three functional domains, an N-terminal zinc finger domain, a catalytic core domain and a C-terminal DNA binding domain. NMR analysis of an isolated N-terminal domain ($IN^{1–55}$) has shown that $IN^{1–55}$ exists in two conformational states [E and D forms; Cai *et al.* (1997) *Nat. Struct. Biol.* 4, 567–577]. The two forms differ in the coordination of the zinc ion by two histidine residues. In the present study, structural analysis of a mutant of $IN^{1–55}$, Y15A, by NMR spectroscopy indicated that the mutant protein folds correctly but takes only the E form. Since the Y15A mutation abrogates the HIV-1 infectivity, Y15 might have some important role in the full-length integrase activity during the virus infection cycle. Our results suggest a possible role of Y15 in structural transition between the E and D forms of HIV-1 integrase to allow the optimal tetramerization.

Key words: IN, HIV-1, NMR, structure, Zn finger.

Human immunodeficiency virus type 1 (HIV-1) integrase mediates the insertion of the newly synthesized proviral DNA intermediate into the host chromosomal DNA as the last event of the provirus establishment (*1*). HIV-1 integrase consists of three functional domains: a central catalytic core domain, an N-terminal zinc binding domain, and a C-terminal nonspecific DNA binding domain. The central core domain contains the highly conserved D,D(35)E motif, which is directly involved in the catalytic activity of integrase. The C-terminus, having a structure that closely resembles Src homology 3 domains, possesses sequence– and metal ion–independent DNA binding activity. The N-terminal domain contains a highly conserved zinc-binding HHCC motif consisting of two His and two Cys residues. Although the catalytic core is not located in the N-terminal domain, this domain at least influences the catalytic activity.

The solution structure of the isolated N-terminal domain ($IN^{1–55}$) has been determined by NMR (*2*). In the solution structure, the $IN^{1–55}$ monomer consists of four helices with a zinc ion tetrahedrally coordinated to the HHCC residues, His12, His16, Cys40 and Cys43, and helices 2 and 3 of $IN^{1–55}$ form a helix-turn-helix motif. $IN^{1–55}$ is dimeric and, interestingly, exists in two interconverting conformational states (E and D forms) that differ with regard to the coordination of the two histidine side chains to zinc. Helices 2, 3 and 4 are the same in the E and D forms of $IN^{1–55}$, and helix 1 extends from residues 2–14 in the E form and from 2–8 in the D form. The loop connecting helices 1 and 2 in the E form is changed into a helical turn (residues 14–17) in the D form. Thus, there is a large difference in the conformation of the polypeptide chain comprising residues 9–18. In the E form, His12 is located in helix 1 (residues 2–14), whereas, in the D form, His12 locating between the shortened helix 1 and the helical turn is exposed to the solvent. As a consequence, the coordination of His12 to zinc is different for the E and D forms (Fig. 1); the N_{ε_2} and N_{δ_1} atoms are coordinated to zinc, respectively. On the other hand, the crystal structure of the N-terminal and catalytic core domains ($IN^{1–212}$) of HIV-1 integrase has been determined (*3*). At the monomeric level, the structure of the N-terminal domain in the crystal structure is similar to the $IN^{1–55}$ E form in the solution structure, but the dimer structure is entirely different between the two structures. The dimer structure of the core domain in the crystal structure of HIV-1 $IN^{1–212}$ is essentially identical to the crystal structures of the core domain (*4*), and the core and C-terminal domains (*5*). It should be noted that, in the crystal structure of HIV-1 $IN^{1–212}$, two dimers form a tetramer.

Recently, it was found that mutation of a conserved amino acid Tyr15 replaced by Ala in the zinc finger domain (Fig. 2) resulted in loss of the infectivity (Masuda *et al.*, unpublished data). In the present study, structure of the mutant of $IN^{1–55}$, Y15A, was analyzed by NMR and it was found that Y15A takes only the E form, whereas the wild type ($IN^{1–55}$) can take on both the E and D forms.

MATERIALS AND METHODS

Cloning and Preparation of $IN^{1–55}$ or Y15A—A plasmid that encodes the gene for glutathione-*S*-transferase (GST) fused to $IN^{1–55}$ or Y15A was prepared. The integrase gene fragment was created by amplification of the appropriate region of HIV-1 pNL43 (Genbank accession number, M19921) with primers introducing a *Bam*HI site at the

*To whom correspondence should be addressed. Phone/Fax: +81-47-478-0425, E-mail: gkawai@sea.it-chiba.ac.jp

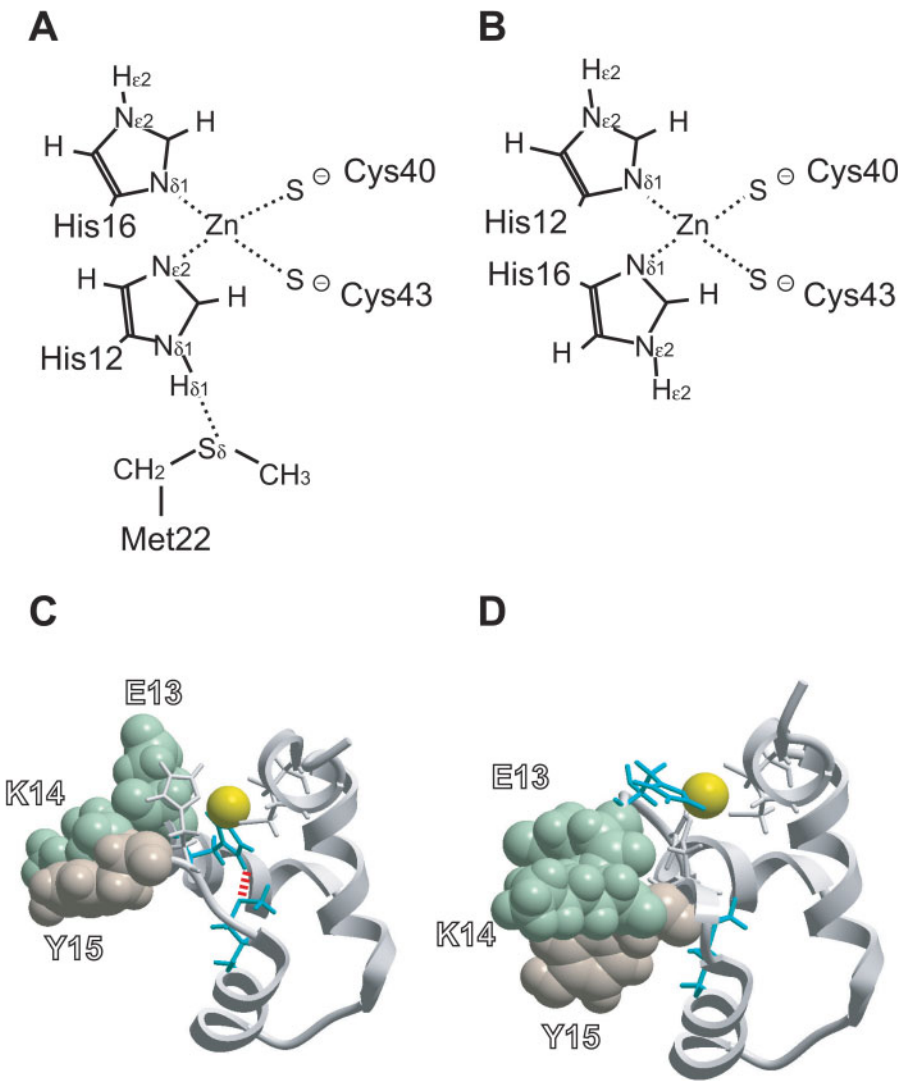

Fig. 1. **Schematic drawings and structures of the zinc coordination for the E (A, C) and D (B, D) forms of the solution structure of IN$^{1–55}$ (*2*).** For panels C and D, Zn^{2+} is shown as a space filling model in yellow. His12, His16, Cys40, Cys43 and Met22 are shown as a stick model, and His12 and Met22 are colored light blue. Tyr15 is shown as a space filling model in pink, and Glu13 and Lys14 as a space filling model in green. The dashed red line indicates the hydrogen bond between the His12 H_{δ_1} proton and Met22 S_δ sulfur. The diagrams were created with the program ICM (MolSoft).

5′-end of the product, and a stop codon and an *Eco*RI site at the 3′-end of the product. The fragment was inserted into the *Bam*HI and *Eco*RI sites of the multiple cloning site of pGEX-2T.

^{15}N-labeled GST-IN$^{1–55}$ or GST-Y15A was expressed in the *E. coli* BL21 strain using the plasmid vector. Cells were grown in 5 mL LB medium at 37°C. Upon reaching an optical cell density at 600 nm of 0.6, the cells were subcultured in 300–1,000 ml isotopically-labeled minimal medium including 1 g/liter [^{15}N]ammonium chloride (Taiyo Nippon Sanso, Japan). When the optical cell density at 600 nm reached 0.8, isopropylthio-β-D-galactoside (IPTG) was added to a concentration of 1 mM. After incubation for 6 h, the cells were harvested by centrifugation at 4°C and 7,000 rpm for 10 min. Wet cells containing ^{15}N-labeled GST-IN$^{1–55}$ or GST-Y15A of 2.0–6.0 g were obtained.

Expression of ^{13}C/^{15}N-labeled GST-Y15A was preformed as described by Marley *et al.* (*6*). Cells were grown in 3 L LB medium at 37°C. When the optical cell density at 600 nm reached 0.6, the cells were pelleted by 10 min centrifugation at 7,000 rpm. The cells were then washed and pelleted two times using an M9 salt solution to exclude all nitrogen and carbon sources. Cell pellets were resuspended in 1 L isotopically-labeled minimal medium including 1 g/liter [^{15}N] ammonium chloride and 2 g/liter [^{13}C]glucose (Taiyo Nippon Sanso, Japan), and then incubated until the recovery of growth. Protein expression was induced after 15 min by the addition of IPTG to a concentration of 1 mM. After incubation for 6 h, the cells were harvested by centrifugation at 4°C and 7,000 rpm for 10 min. Wet cells containing ^{13}C/^{15}N-labeled GST-Y15A of 4.0 g were obtained.

Wet cells containing ^{13}C/^{15}N or ^{15}N-labeled GST-IN$^{1–55}$, or GST-Y15A were resuspended in sonication buffer (20 mM HEPES, 500 mM NaCl, 2 mM β-mercaptoethanol (BME), 0.4 g/liter lysozyme, 0.1 mM $ZnCl_2$, pH 7.5). Subsequently, the cells were sonicated and the cell walls were removed by centrifugation at 4°C and 15,000 × *g* for 60 min. The supernatant containing soluble proteins were filtered through a 0.2 μm filter, and then loaded onto a GST affinity column equilibrated with binding buffer (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 140 mM NaCl, 2.7 mM KCl, 2 mM BME, 0.1 mM $ZnCl_2$, pH 7.3). The column was washed with 5

```
                          ▼
HIV-2ROD    FLEKIEPAQE EHEKYHSNVK ELSHKFGIPN LVARQIVNSC AQCQQKGEAI HGQVN
            **# *# *** ******** #  ##  * #*  #*####* **  #** ****# ****#
HIV-1NL43   FLDGIDKAQE EHEKYHSNWR AMASDFNLPP VVAKEIVASC DKCQLKGEAM HGQVD
            *** *#####  #* ***#*** #*  #* **  #********#*  ***##**   *****
SIVAGM      FLDRIEEAQD DHAKYHNNWR SMVQEFGLPN IVAKEIVAAC PKCQIRGEPK HGQVD
```

Fig. 2. **Amino acid sequences of the N-terminal zinc finger domains of integrases from HIV-1_{NL43}, HIV-2_{ROD} and SIV_{AGM}.** The accession numbers for HIV-1_{NL43}, HIV-2_{ROD} and SIV_{AGM} are M19921, M15390 and M66437, respectively. The amino acid sequence alignment was performed using BLAST. The location of Tyr15 is indicated by a triangle. The asterisk and sharp indicate homology and analogy in amino acids, respectively.

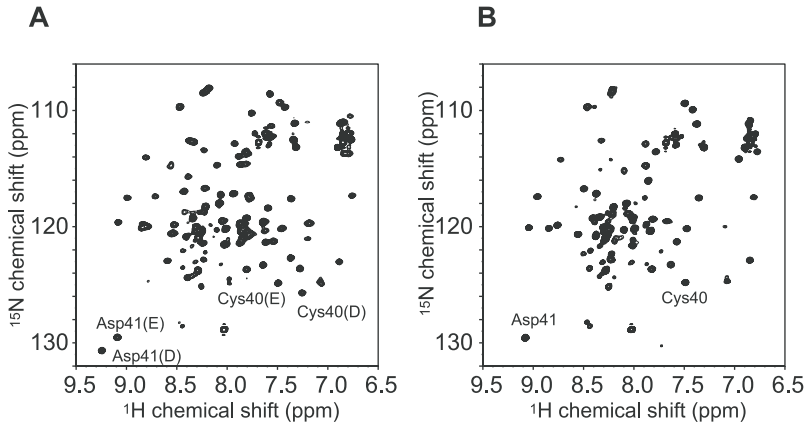

Fig. 3. **^{15}N-^{1}H HSQC spectra of $IN^{1–55}$ (A) and Y15A (B) recorded at pH 6.5 and 293 K.** "E" and "D" indicate the E and D forms of $IN^{1–55}$, respectively.

column volumes of the binding buffer. GST-$IN^{1–55}$ or GST-Y15A was eluted with an elution buffer (50 mM Tris-HCl, 10 mM glutathione in the reduced form, 2 mM BME, 0.1 mM $ZnCl_2$). GST was removed by incubating the purified GST-$IN^{1–55}$ or GST-Y15A with thrombin at 22°C for 16 h.

NMR Spectroscopy—The purified stable-isotopic labeled $IN^{1–55}$ or Y15A was concentrated and then dissolved in buffer for the NMR experiment (50 mM sodium phosphate, 150 mM NaCl, 2 mM BME, 0.1 mM $ZnCl_2$, pH 6.5) by using Centricon-3 (Millipore Inc.). $^{13}C/^{15}N$-labeled Y15A was concentrated to 1.8 mM, and ^{15}N-labeled $IN^{1–55}$ and Y15A were concentrated to 1.0 mM. All NMR experiments were carried out at 293 K on Bruker DRX500 and DRX600 spectrometers. ^{15}N-^{1}H HSQC (*7*) spectra were measured for ^{15}N-labeled $IN^{1–55}$ and Y15A. The main chain signals for Y15A were assigned using 3D ^{15}N-edited NOESY (mixing times: 200 ms, 75 ms) (*8*), 3D HBHACONH (*9*), 3D HNCA (*10*), 3D HN(CO)CA (*11*), 3D CBCANH (*12*), 3D CBCA(CO)NH (*13*), 3D HNCO (*10*), and 3D HNCACO (*14*). Side chain signals were assigned using 3D ^{13}C-edited NOESY (mixing times: 150 ms, 75 ms) and 3D HC(C)H-TOCSY (*15*). Data were processed with the program X-WinNMR (Bruker Biospin) and spectra were analyzed using the program Felix (Accelrys Software Inc.). Secondary structures were predicted with the program TALOS (*16*) using N, C_β, C_α, and H_α chemical shifts.

Computational Analysis—The coordinates of the $IN^{1–55}$ E form (PDB code 1WJC) and the D form (PDB code 1WJA) were obtained from the Brookhaven Protein Data Bank (PDB) (*2*). All free energy calculations were performed with the GIBBS module of the program AMBER version 7 (*17*) to calculate the mutation energy, $\Delta G_{WT\text{-}Y15A} = G_{WT} - G_{Y15A}$, for the E or D form. For each of the E and D forms of $IN^{1–55}$, a model system was built in a periodic box of the TIP3P solvent molecules with a minimum distance of 10 Å from the protein molecule. All simulations were prepared with 100 steps of minimization and 130 ps of solvent equilibration, with a slow warm up from 0 to 298 K over 20 ps. The temperature was set at 298 K for the molecular dynamics simulation. The simulations were performed at constant pressure and with periodic boundary conditions with the slow growth method, and SHAKE was used for bonds involving hydrogen. The total length of the simulation was set at 20 ps. For each of the E and D forms of $IN^{1–55}$, the side chain solvent accessible surface area (ASA) for Tyr15 was calculated with the program GETAREA 1.1 (*18*).

RESULTS

Spectral Comparison—The structures of purified $IN^{1–55}$ and Y15A were analyzed by NMR spectroscopy. The ^{15}N-^{1}H HSQC spectra of $IN^{1–55}$ and Y15A are shown in Fig. 3. The well dispersed resonances in the spectra indicated that $IN^{1–55}$ and Y15A are in correctly folded conformations. We observed resonances corresponding to the E and D forms of $IN^{1–55}$, whereas mutant Y15A was found to exist in a single conformation under the same condition based on the fact that Cys40 and Asp41 exhibited only a single set of resonances per residue in the HSQC spectrum of Y15A. Their HN chemical shifts suggested that Y15A is in the E form.

Structural Analysis of Y15A—The structure of Y15A was further analyzed. Triple resonance experiments were performed by with $^{13}C/^{15}N$ Y15A. N, HN, C_α and C_β chemical shifts were derived from the following spectra, 3D HNCA, 3D HN(CO)CA, 3D CBCANH and

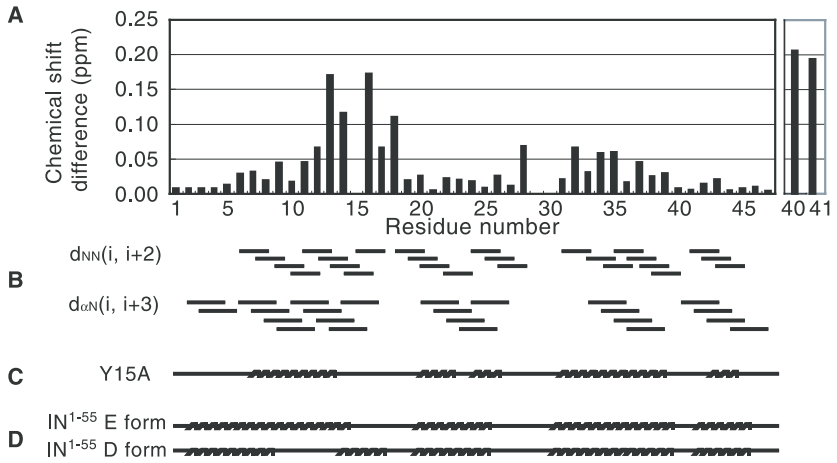

Fig. 4. **Structural analysis of Y15A.** (A) Chemical shift differences in HN and N signals between the $IN^{1–55}$ E form and Y15A. The mutated residue (residue 15) and proline residues at positions 29 and 30 are not shown. Chemical shift differences were measured by comparison of the cross peaks in the ^{15}N-^{1}H HSQC spectra of the $IN^{1–55}$ E form and Y15A. Δav is a weighted average of the ^{1}H and ^{15}N chemical shift differences; $\Delta av = [((\Delta HN)^2 + (\Delta N/5)^2)/2]^{1/2}$, where ΔHN and ΔN are the ^{1}H and ^{15}N chemical shift differences in ppm, respectively (*19*). Chemical shift differences between the E and D forms of $IN^{1–55}$ for Cys40 and Asp41 are shown in the right panel for reference. (B) $H_\alpha(i)$–HN(i+3) and HN(i)–HN(i+2) NOEs indicating α helices. (C) α helices of Y15A analyzed by TALOS (*16*) with N, C_β, C_α and H_α chemical shifts. (D) Arrangements of α helices for the E and D forms determined by Cai, M. *et al.* (*2*).

3D CBCA(CO)NH. H_α and H_β chemical shifts were derived from the following spectra, 3D ^{15}N-edited NOESY (mixing times: 200 ms, 75 ms) and 3D HBHA(CO)NH. Side chain chemical shifts were derived from the following spectra, 3D ^{13}C-edited NOESY (mixing times: 150 ms, 75 ms) and 3D HC(C)H-TOCSY. NOEs were derived from 3D ^{15}N-edited NOESY (mixing times: 200 ms, 75 ms) and 3D ^{13}C-edited NOESY (mixing times: 150 ms, 75 ms). The chemical shift differences between $IN^{1–55}$ in the E form and Y15A are shown in Fig. 4A together with the chemical shift differences of Cys40 and Asp41 between $IN^{1–55}$ in the E and D forms in the right panel. Although differences of more than 0.1 ppm were observed for the residues flanking the mutation site, the chemical shift differences between $IN^{1–55}$ in the E form and Y15A were small in general. Moreover, the NOEs for $H_\alpha(i)$-HN(i+3) and HN(i)-HN(i+2) indicated that the location of the α helices (Fig. 4B) is also consistent with those for $IN^{1–55}$ in the E form (Fig. 4D). The secondary structures predicted with the program TALOS with the chemical shifts of C_β, C_α, H_α and N were also consistent with the E form (Fig. 4, C and D). In the solution structure of the $IN^{1–55}$ E form, the $H_{\delta1}$ proton of His12 donates a hydrogen bond to the S_δ atom of Met22 (*2*), as shown in Fig. 1, and thus the H_{δ_1} proton signal of His12 can be detected in the ^{15}N-^{1}H HSQC spectrum. In fact, an H_{δ_1} proton signal was observed for Y15A suggesting that the zinc coordination also corresponds to that in the E form (Fig. 5). Table 1 shows some peak intensities in the ^{15}N-^{1}H HSQC spectra referred to the peak intensity of Asp55 at the carboxyl terminal. For the split peaks of $IN^{1–55}$, the intensity ratio for the E and D forms is almost 1.0, indicating that the molar ratio of the E and D forms is about 1.0. In the case of Y15A, the peaks corresponding to the E form of $IN^{1–55}$ show almost twice the intensity of the E form of $IN^{1–55}$, clearly indicating that Y15A exists in a single conformation. It should be noted that the peak intensity for His12 H_{δ_1}–$^{15}N_{\delta_1}$ of Y15A is also roughly twice that of the E form of $IN^{1–55}$. These facts indicated that Y15A takes on the E form exclusively under the present experimental conditions.

Computational Analysis of the $IN^{1–55}$ Mutation—The free energy differences between $IN^{1–55}$ and Y15A ($\Delta G_{WT\text{-}Y15A}$) were calculated by using AMBER7 GIBBS module (*17*).

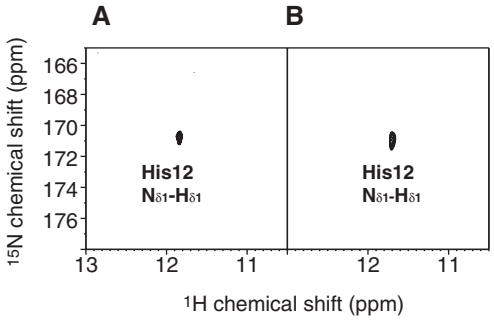

Fig. 5. **^{15}N-^{1}H HSQC spectra of $IN^{1–55}$ (A) and Y15A (B) showing the $N_{\delta1}$-$H_{\delta1}$ cross peak for His12.** The H_{δ_1} proton signal of His12 is observed when its H_{δ_1} proton donates a hydrogen bond to the S_δ atom of Met22 in the E form type zinc finger in solution (*2*). The N_{ε_2}-H_{ε_2} cross peaks of the other His can not be observed due to the rapid exchange of the H_{ε_2} proton with the bulk solvent.

Table 1. **Relative peak intensities of $IN^{1–55}$ and Y15A determined from ^{15}N-^{1}H HSQC spectra.**

Atom	Intensity			Ratio	
	E[1]	D[1]	Y15A	E/D	Y15A/E
Split peaks					
$Q9H_N$	0.08	0.08	0.20	1.07	2.41
$A21H_N$	0.10	0.09	0.21	1.08	2.15
$S24H_N$	0.10	0.10	0.15	1.00	1.60
$N27H_N$	0.08	0.08	0.13	1.03	1.63
$K34H_N$	0.04	0.04	0.09	1.03	2.21
$C40H_N$	0.09	0.06	0.16	1.42	1.80
$D41H_N$	0.11	0.08	0.21	1.24	2.00
$H12H_{\delta1}$	0.07	N.D.	0.10	N.D.	1.45
Atom	$IN^{1–55}$	Y15A		$Y15A/IN^{1–55}$	
Single peaks					
$A49H_N$	0.47	0.44		0.94	
$G52H_N$	0.19	0.15		0.77	
$V54H_N$	0.78	0.66		0.85	
$D55H_N$	1.00	1.00		1.00	

[1]"E" and "D" refer to the "$IN^{1–55}$ E form" and "$IN^{1–55}$ D form," respectively. NMR data were processed using π/2 shifted squared sine bell (F2) and π/2 shifted sine bell (F1) window functions.

$\Delta G_{WT\text{-}Y15A}$ for the E and D forms were –5.33 and –42.5 kJ/mol, respectively. Based on the fact that $IN^{1–55}$ takes on the E and D forms equally at room temperature, it was suggested that the E form is more stable than the D form in Y15A, which is consistent with the above conclusion. The ASA values of the side chain of Tyr15 for the E and D forms were calculated using the program GETAREA 1.1 (*18*), it being found that the ASA values are 109.50 and 122.73 $Å^2$ for the E and D forms, respectively. The greater exposure of the hydrophobic side chain of Ala15 may make the D form less stable in Y15A.

DISCUSSION

Comparison Between Y15A and Related Zinc Fingers—Proteins containing the classical CCHH zinc fingers are found in many nucleic acid binding proteins. In contrast, HHCC zinc finger is peculiar to integrase. The HIV-1 integrase N-terminal zinc finger domain, $IN^{1–55}$, is unstructured in the absence of zinc, but in the presence of zinc folds into a well-defined dimeric structure comprising four helices per monomer (*2*), and the zinc ion coordinated by His12, His16, Cys40 and Cys43. Interestingly, $IN^{1–55}$ exists in two interconverting conformational states (E and D forms) in solution that differ with regard to the coordination of the two histidine side chains to zinc, as shown in Fig. 1 (*2*). The solution structure of $IN^{1–55}$ mutant H12C has also been determined by NMR (*20*), it being found that H12C forms multiple conformations when in a complex with zinc, whereas the cadmium substituted protein takes on a single conformation. The conformation of H12C in the presence of cadmium is intermediate between the two forms of $IN^{1–55}$, and the cadmium is coordinated by Cys12, His16, Cys40 and Cys43. The presence of two interconverting structures (E and D forms) was previously observed for a mutant nucleocapsid protein, NCp7, of HIV-1, in which the CCHC motif was replaced by a CCHH motif (*21*). The structure of the CCHH motif in the zinc finger domain of NCp7 is different to that of the native CCHC motif. In the present study, the structure of Y15A was analyzed by NMR, it being found that Y15A takes only the E form. Although the Y15A and H12C mutations in $IN^{1–55}$ as well as the CCHC/CCHH mutation in NCp7 affect the equilibrium between the two conformations, the interesting feature of Y15A is that a mutation introduced at a residue other than the zinc coordinating His or Cys residue contributes to the stability of the protein structure.

Why Does Y15A Take Only the E Form?—Tyr15 in $IN^{1–55}$ is not involved in the hydrophobic core which contributes to the formation of the helix-turn-helix motif and the zinc finger. Nevertheless, Y15A was found to form only the E form. The side chain of Tyr15 is exposed to the solvent (Fig. 6). The ASA values of Tyr15 in the E and D forms are 109.55 and 122.73 $Å^2$, respectively, indicating that the side chain of Tyr15 in the D form is exposed to the solvent more than that in the E form. Because the hydrophobicity of Ala with the value of 1.6 is higher than that of Tyr, –0.7 (*22*), the Y15A mutation may destabilize the D form structure more than the E form structure. The interaction between His12 and Met22 is one of the possible underlying factors for the stability of the E form (Fig. 1, A and C). This interaction does not occur in the D form and this energy loss may be compensated for by the amphiphilic interaction among E13, K14 and Y15, as shown in Fig. 1D; these amino acid residues are classified as being highly amphiphilic (*23*). The mutation of Tyr15 to Ala may disrupt this amphiphilic cluster. Further analysis will lead to a deeper understanding of this interesting phenomenon.

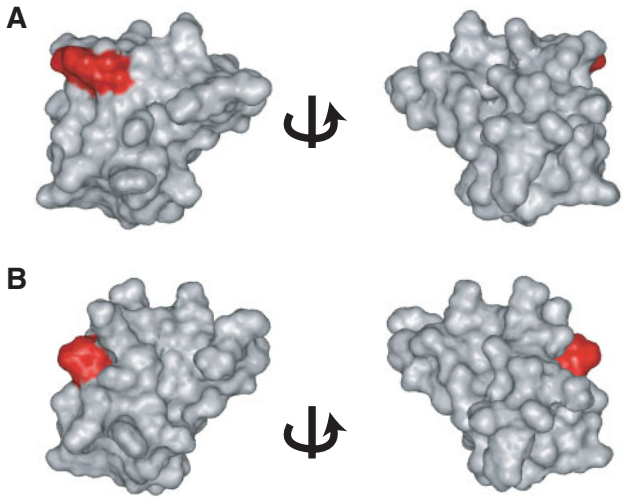

Fig. 6. **Molecular surface diagrams of the $IN^{1–55}$ monomer.** The side chain of Tyr15 shown in red is exposed to the solvent in each of the E (A) and D (B) forms. The molecular surfaces were created using the program InsightII (Accelrys Software Inc.).

Structure and Function Relationship of the HIV-1 Integrase Zinc Finger Domain—It has been shown that mutations of the N-terminal domain HHCC motif disrupt the 3′-end processing and the strand transfer activity of integrase *in vitro* (*24*, *25*), and inhibit viral replication in cultured cells by preventing the reverse transcription reaction (*26*). Thus, the zinc finger domain of integrase is required for full infectivity.

In general, the correct folding of a protein is required for the expression of its function. Although, Y15A is in one of the correctly folded conformations, the mutation abolishes the infectivity (Masuda *et al.*, unpublished data). This is different from the cases of CCHH type NCp7 (*21*) and H12C $IN^{1–55}$ (*20*) forming incorrect conformations. The structure of the N-terminal domain in the crystal structure of $IN^{1–212}$ is similar to that in the $IN^{1–55}$ E form of the isolated domain at the monomeric level (*3*). On the other hand, the HIV-2 integrase zinc finger domain forms only the E form in solution structure in spite of the presence of Tyr15 (*27*, *28*). It is notable that the Tyr15 residue is conserved at the same position and the structure is also conserved among the three zinc finger structures; the E form of the $IN^{1–55}$ solution structure, the N-terminal domain in the crystal structure of $IN^{1–212}$ and the solution structure of the HIV-2 integrase zinc finger domain (Fig. 7). Although the functional role of the D form remains to be determined, it is suggested that the HIV-1 integrase zinc finger domain is required to take the E form with Tyr15 for full infectivity.

As described above, the catalytic core and zinc finger domains interact to form the dimer of dimers in the crystal structure of HIV-1 $IN^{1–212}$ (*3*). Moreover, integrase forms a tetramer in solution, as judged on gel filtration, and its

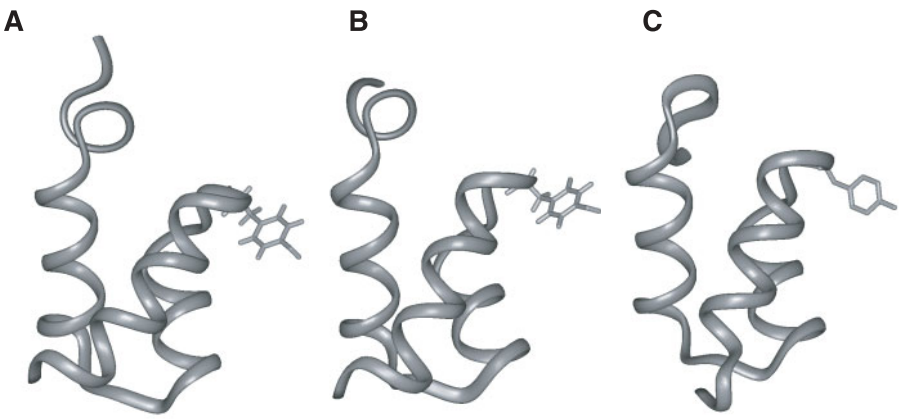

Fig. 7. **Location and structure of Tyr15 in several zinc finger structures.** (A–C) The orientation of the side chain of Tyr15 on the monomeric structure is similar in the three structures; E form of $IN^{1–55}$ (A; PDB code 1WJC), HIV-2 (B; PDB code 1E0E), and $IN^{1–212}$ (C; PDB code 1K6Y). Tyr15 is displayed as a stick. The diagrams were created using the program InsightII.

tetramer formation is required for integration activity (*24*). In the crystal structure of the $IN^{1–212}$ tetramer, Tyr15 is located at the dimer-dimer interface of the integrase tetramer and the side chain of Tyr15 stacks on the side chain of Lys186 of other subunit. It is notable that a mutation of Lys186 was also found to abrogate the infectivity (*29*). Thus, it was suggested that the interaction between Tyr15 and Lys186 is required for the optimal tetramerization of integrase, which is required for the integration activity. In the course of the integration reaction, formation of the D form may be required, for example, the dissociation of the tetramer into dimers.

Further analysis of the role of Tyr15 might be useful also for elucidation of the function of integrase as well as the development of the integrase inhibitor.

Database Deposition—The chemical shift assignments have been deposited in the Biological Magnetic Resonance Data Bank (accession code 10015).

We are indebted to Drs. T. Someya, N. Nameki and T. Sakamoto for the valuable advice regarding protein purification and NMR analyses. We also thank Mr. H. Sato for his assistance in the NMR measurements. This work was supported, in part, by a Grant-in-Aid for High Technology Research from the Ministry of Education, Science, Sports and Culture, Japan.

REFERENCES

1. Hindmarsh, P. and Leis, J. (1999) Retroviral DNA Integration. *Microbiol. Mol. Biol. Rev.* **63**, 836–843
2. 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
3. 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
4. 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
5. Chen, J.C.H., Krucinski, J., Miercke, L.J.W., 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
6. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **20**, 71–75
7. Bodenhausen, G. and Ruben, D.J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69**, 185–189
8. Jeener, J., Meier, B.H., Bachmann, P., and Ernst, R.R. (1979) Investigation of exchange processes by two- dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546–4553
9. Grzesiek, S. and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins. *J. Biomol. NMR* **3**, 185–204
10. Kay, L.E., Ikura, M., Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* **89**, 496–514
11. Grzesiek, S. and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.* **96**, 432–440
12. Grzesiek, S. and Bax, A. (1992) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson.* **99**, 201–207
13. Grzesiek, S. and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291–6293
14. Clubb, R.T., Thanabal, V., and Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue $^{1}H^{N}$, ^{15}N, and $^{13}C'$ chemical shifts in ^{15}N-^{13}C-labelled proteins. *J. Magn. Reson.* **97**, 213–217
15. Kay, L.E., Xu, G.Y., Singer, A.U., Muhandiram, D.R., and FormanKay, J.D. (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples of proteins. *J. Magn. Reson.* **B101**, 333–337
16. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289–302
17. Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham, III, T.E., DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to

simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.* **91**, 1–41
18. Fraczkiewicz, R. and Braun, W. (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comp. Chem.* **19**, 319–333
19. Pellecchia, M., Sebbel, P., Hermanns, U., Wüthrich, K., and Glockshuber, R. (1999) Pilus chaperone FimC–adhesin FimH interactions mapped by TROSY-NMR. *Nat. Struct. Biol.* **6**, 336–339
20. Cai, M., Huang, Y., Caffrey, M., Zheng, R., Craigie, R., Clore, G.M., and Gronenborn, A.M. (1998) Solution structure of the His12 → Cys mutant of the N-terminal zinc binding domain of HIV-1 integrase complexed to cadmium. *Protein Science* **7**, 2669–2674
21. Ramboarina, S., Moreller, N., Fournié-Zaluski, M.C., and Roques, B.P. (1999) Structural investigation on the requirement of CCHH zinc finger type in nucleocapsid protein of human immunodeficiency virus 1. *Biochemistry* **38**, 9600–9607
22. Engelman, D.M., Steitz, T.A., and Goldman, A. (1986) Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* **15**, 321–353
23. Mitaku, S., Hirokawa, T., and Tsuji, T. (2002) Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. *BioInfomatics* **18**, 608–616
24. Zheng, R., Jenkins, T.M., and Craigie, R. (1996) Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc. Natl. Acad. Sci. USA* **93**, 13659–13664
25. 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
26. Masuda, T., Planelles, V., Krogstad, P., and Chen, I.S. (1995) Genetic analysis of human immunodeficiency virus type 1 integrase and the U3 *att* site: unusual phenotype of mutants in the zinc finger-like domain. *J. Virol*. **69**,6687–6696
27. Eijkelenboom, A.P., van den Ent, F.M., Vos, A., Doreleijers, J.F., Hård, K., Tullius, T.D., Plasterk, R.H.A., Kaptein, R., and Boelens, R. (1997) The solution structure of the amino-terminal HHCC domain of HIV-2 integrase: a three-helix bundle stabilized by zinc. *Curr Biol.* **7**, 739–746
28. Eijkelenboom, A.P., van den Ent, F.M., Wechselberger, R., Plasterk, R.H.A., Kaptein, R., and Boelens, R. (2000) Refined solution structure of the dimeric N-terminal HHCC domain of HIV-2 integrase. *J. Biomol. NMR* **18**, 119–128
29. Tsurutani, N., Kubo, M., Maeda, Y., Ohashi, T., Yamamoto, N., Kannagi, M., and Masuda,T. (2000) Identification of critical amino acid residues in human immunodeficiency virus type 1 IN required for efficient proviral DNA formation at steps prior to integration in dividing and nondividing cells. *J. Virol.* **74**, 4795–4806