

Conserved Residues Pro-109 and Asp-116 Are Required for Interaction of the Human Immunodeficiency Virus Type 1 Integrase Protein with Its Viral DNA Substrate

MEERA DRELICH, MARKUS HAENGGI, AND JAN MOUS*

Pharmaceutical Research New Technologies, F. Hoffmann-La Roche Ltd., CH-4002 Basel, Switzerland

Received 9 December 1992/Accepted 29 April 1993

The human immunodeficiency virus type 1 integrase protein can be specifically cross-linked to viral long terminal repeat substrate oligonucleotides in vitro by using UV light. Site-directed mutagenesis and deletion analyses were used to define the domains involved in the interaction of integrase with the viral DNA substrate. Our results showed that mutation of conserved residues Pro-109 and Asp-116, which are found to be critical for the endonuclease and integration activities of IN protein, abolished the ability of the protein to cross-link to its DNA substrate. Furthermore, deletion analysis experiments showed that removal of 39 amino acids from the amino terminus and deletion of 15 amino acids from the carboxyl terminus abolished DNA cross-linking.

The integration of retroviral DNA into host cell DNA is essential for the replication of retroviruses. This process is accomplished by the integrase (IN) protein, the only viral protein needed for a defined set of reactions that include endonucleolytic removal of two bases from the 3' terminus of linear viral DNA followed by joining of the processed viral DNA to nicked host (target) DNA (7, 8). The biochemical properties associated with human immunodeficiency virus (HIV) IN have been well defined by in vitro assays. These properties include DNA binding (9, 11, 16), specific DNA endonuclease activity (1, 4, 15, 17), DNA joining or strand transfer activity (1, 2, 4, 10, 12), and reversal of DNA joining, termed disintegration (3, 5).

Recent studies have pinpointed several key conserved residues (4, 5, 10, 13, 14) and regions (4) within the integrase that are essential for the processing and joining reactions mediated by the enzyme in vitro. It is now believed that residues located in the central domain of the protein form part of a single active site. However, not much is known about the parts of the protein that are involved in interaction with the viral DNA substrate. To address this question, we exposed DNA-IN protein complexes to UV light under conditions expected to permit covalent cross-linking between DNA and residues in close contact at the time of cross-linking.

HIV type 1 (HIV-1) IN protein was incubated with ³²P-labeled oligonucleotides representing the ends of viral long terminal repeat (LTR) sequences, exposed to UV light, and subjected to electrophoresis on a denaturing polyacrylamide gel. A cross-linked product with an electrophoretic mobility close to that of integrase protein was observed (Fig. 1). In addition, a higher-molecular-weight band which corresponded in approximate size to the dimeric form of IN protein was observed. That this protein band was related to IN protein was confirmed by Western immunoblot analysis (data not shown). The reaction conditions for the cross-linking reaction were similar to those used to assay the endonucleolytic and integration activities of IN protein (4). Typically, 10 pmol of IN was mixed with 2 pmol of 5'-end-labeled oligonucleotide substrate in a total reaction volume

of 10 μ l containing 20 mM Tris-HCl (pH 7.6), 1 mM dithiothreitol, and 2 mM MnCl₂. The reaction mixture was incubated on ice for 30 min and then exposed to UV light (254 nm) for 15 min at 1,200 J/cm²/min on ice with a UV Stratalinker 1800 (Stratagene).

Cross-linking was dependent on the presence of IN protein (Fig. 1A), divalent cations (Fig. 1B), and UV light (Fig. 1C). When IN protein was boiled prior to cross-linking, its ability to bind DNA was greatly diminished (Fig. 1B), reflecting the need for an optimal conformational state for binding. The reaction was complete within 10 min of exposure to UV light (Fig. 1C, lane 3). Cross-linking was sensitive to high salt (500 mM), as demonstrated in Fig. 1C, lane 8, and the addition of excess of cold oligonucleotide prevented the formation of a radioactive adduct (data not shown).

The various oligonucleotide DNA substrates used to determine the sequence requirements for cross-linking are shown in Fig. 2A. Both strands were labeled with ³²P at the 5' terminus with T4 polynucleotide kinase to a specific activity of $\pm 10^5$ cpm/pmol and then annealed before use in the cross-linking reaction. The substrate sequence requirements were studied by cross-linking of HIV-1 IN to different HIV and murine leukemia virus LTR substrates. Equivalent amounts (2 pmol) of substrates, all labeled to approximately similar activities, were used in the cross-linking reaction. Our results (Fig. 3) showed that HIV-1 IN protein exhibits different affinities for the U₃ and U₅ LTRs of HIV-1, with a reduced affinity for the U₃ LTR (Fig. 3, lane 2). However, when a longer oligonucleotide, 35 nucleotides in length, representing the U₃ end, was used, cross-linking ability was increased, suggesting that integrase may recognize additional sequences at the U₃ end (Fig. 3, lane 7). Removal of two nucleotides (GT) from the 3' end of the U₅ LTR plus strand (Fig. 3, lane 5) partially affected DNA-protein cross-linking. However, mutation of the recognition dinucleotide CA to GG drastically reduced the cross-linking ability of IN protein (Fig. 3, lane 6). In addition, HIV-1 IN protein only weakly cross-linked to HIV-2 and murine leukemia virus U₅ LTRs, shown in Fig. 3, lanes 3 and 4, respectively. No evidence of protein-DNA cross-linking was observed when a labeled nonspecific oligonucleotide was mixed with IN pro-

* Corresponding author.

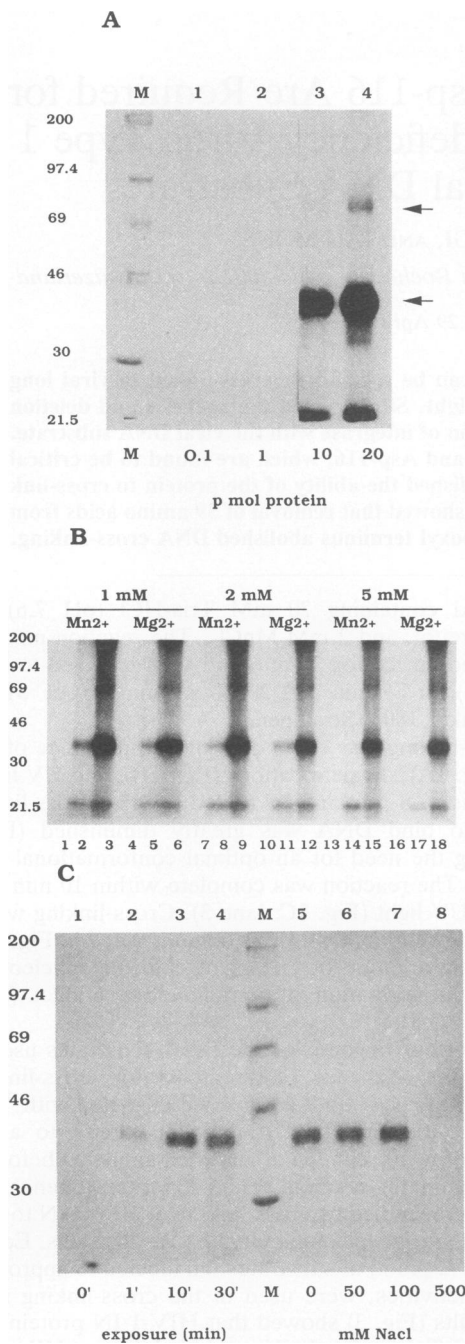


FIG. 1. (A) Characterization of the cross-linking reaction. Dependence on protein concentration. Cross-linking after addition of increasing concentrations of IN protein, with 2 pmol of labeled oligonucleotide substrate (HIV-1 U₅). Lane M, molecular size markers (in kilodaltons) (Rainbow markers) labeled with ¹⁴C, obtained from Amersham. Lane 1, 0.1 pmol; lane 2, 1 pmol; lane 3, 10 pmol; lane 4, 20 pmol of IN protein. The arrows indicate the positions of monomeric (lower) and dimeric (upper) IN proteins. (B) Requirement for divalent cations. The effects of 1 mM MnCl₂ (lanes 2 and 3) and 1 mM MgCl₂ (lanes 5 and 6), 2 mM MnCl₂ (lanes 8 and 9) and 2 mM MgCl₂ (lanes 11 and 12), and 5 mM MnCl₂ (lanes 14 and 15) and 5 mM MgCl₂ (lanes 17 and 18) on cross-linking of IN protein were determined. Divalent cation was omitted in lanes 1, 4, 7, 10, 13, and 16. IN protein was boiled for 10 min prior to cross-linking (lanes 2, 5, 8, 11, 14, and 17). The positions of molecular size markers are shown on the left (in kilodaltons). (C) Time course of

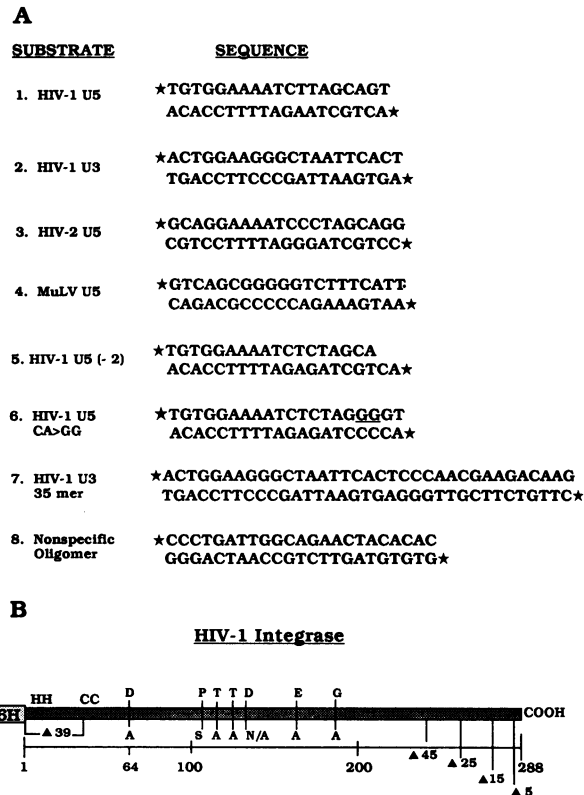


FIG. 2. (A) Oligonucleotide DNA substrates used to determine sequence requirements for cross-linking. The position of the ³²P label is indicated by a star. (B) Point mutations and the locations of in-frame deletion mutations within HIV-1 IN. The corresponding amino acid changes are indicated.

tein and exposed to UV light (Fig. 2, lane 8), demonstrating the specificity of the cross-linking reaction.

Although specific DNA binding could not be detected in DNA-binding competition experiments (16), our results suggest that some degree of terminal specificity exists in the formation of an initial DNA-protein complex and that it is trapped in this state by covalent cross-linking of DNA to IN protein by UV light. That terminal sequences influence IN function is supported by evidence from adduct interference studies performed with U₅ and U₃ DNA ends, in which it was found that specific DNA sites at the ends of HIV DNA are required for IN function (2). Similar findings have also been reported by other groups (11, 12, 17).

Effect of deletion mutations on the cross-linking ability of IN protein. The generation of recombinant HIV-1 IN protein, deletion mutations, and site-directed mutations has been

cross-linking of IN protein to viral substrate DNA (lanes 1 to 4) and effect of NaCl (lanes 5 to 8). Ten picomoles of IN was mixed with 2 pmol of oligonucleotide substrate labeled at the 5' end of both strands, incubated on ice, and exposed to UV light (1,200 J/min) for 0 min (lane 1), 1 min (lane 2), 10 min (lane 3), or 30 min (lane 4). The reaction was stopped by the addition of sodium dodecyl sulfate (SDS) sample buffer and a 15-fold excess of cold substrate DNA. The samples were then heated at 95°C for 5 min and loaded onto 10.5% polyacrylamide gels containing SDS. The gels were visualized by autoradiography. Molecular size markers (in kilodaltons) are shown in lane M.

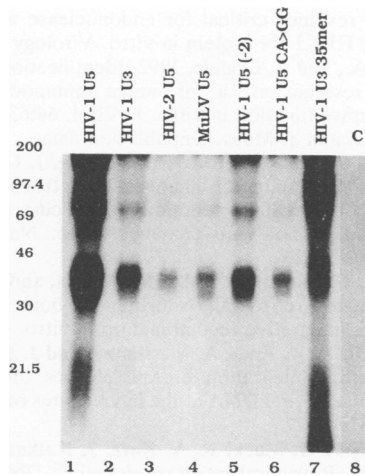


FIG. 3. Substrate sequence requirements for cross-linking to IN. Lanes: HIV-1 U₅ (lane 1), HIV-1 U₃ (lane 2), HIV-2 U₅ (lane 3), murine leukemia virus (MuLV) U₅ (lane 4), HIV-1 U₅(-2) (lane 5), HIV-1 U₅ with CA mutated to GG (lane 6), and HIV-1 U₃ 35-mer (lane 7); lane 8, cross-linking reaction with a nonspecific oligomer as a control. The sequences of all substrates used are shown in Fig. 2A. Positions of molecular size markers are shown on the left (in kilodaltons).

described previously (4). The mutant IN proteins (Fig. 2B) were purified in the same way as the wild-type protein and exhibited identical chromatographic profiles, yielding equivalent quantities of protein (4). To determine whether sequences located at the amino and carboxyl ends of HIV-1 IN protein were required for interaction with DNA, several C-terminally truncated forms and one N-terminal deletion mutant were affinity purified and analyzed for their ability to cross-link to labeled HIV-1 U₅ oligonucleotide substrate. We found that removal of 39 amino acids from the amino terminus completely abolished DNA cross-linking and that deletion of 5 amino acids from the carboxyl terminus partially affected but deletion of 15 amino acids drastically

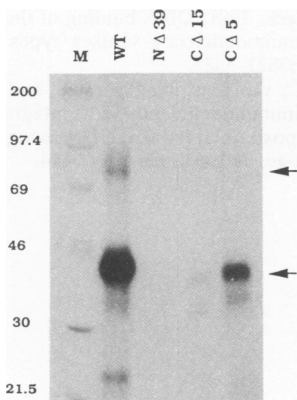


FIG. 4. Effects of in-frame deletions on cross-linking ability of IN protein. All reactions were performed in a 10- μ l total volume with equal amounts (10 pmol) of purified wild-type (WT) or mutated IN proteins and 2 pmol of labeled HIV-1 U₅ LTR DNA. Arrows indicate the positions of cross-linked full-length monomeric and dimeric IN proteins. Lane M, molecular size markers (in kilodaltons).

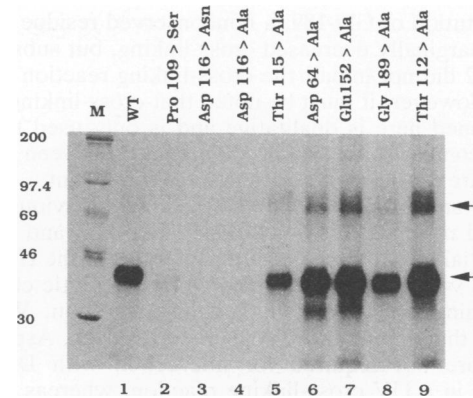


FIG. 5. Cross-linking ability of IN proteins with site-directed mutations. All reactions were performed in a 10- μ l total volume with equal amounts (10 pmol) of purified wild-type (WT) or mutant IN proteins and 2 pmol of labeled HIV-1 U₅ LTR DNA. Arrows indicate the positions of cross-linked monomeric and dimeric IN proteins. Lane M, molecular size markers (in kilodaltons).

affected DNA cross-linking ability (Fig. 4). These results suggest that residues in the N-terminal and C-terminal portions of the IN protein contribute to DNA recognition either directly or indirectly by influencing protein folding.

Since the N-terminal deletion also removed the residues of a putative Zn²⁺-binding motif (6), another highly conserved motif among retroviral endonucleases, site-directed substitution of these residues would reveal their role in DNA recognition binding, if any. It has been reported previously that substitutions of these conserved His and Cys residues do not completely abolish the catalytic activity of the integrase (5, 10). Our preliminary results with a Cys-43 \rightarrow Ser mutant also confirm that substitution of this conserved residue does not influence the catalytic activity of IN or the cross-linking of IN to its substrate DNA (5a).

Cross-linking ability of mutant IN proteins. We used several different site-directed HIV-1 IN mutations to determine the ability of the mutant proteins to cross-link to the HIV-1 U₅ LTR. Point mutations were introduced in a central region of the protein containing the sequence Thr-115-Asp-Asn-Gly-118, which is well conserved among retroviral endonucleases (6, 10). We selected point mutations Pro-109 \rightarrow Ser, Thr-115 \rightarrow Ala, Asp-116 \rightarrow Ala, and Asp-116 \rightarrow Asn, which substituted highly conserved residues, whereas Thr-112, a nonconserved residue, was changed to Ala. Additional site-directed changes were Asp-64 \rightarrow Ala and Glu-152 \rightarrow Ala, two conserved residues, and Gly-189 \rightarrow Ala, a nonconserved residue. The production and purification of wild-type and mutant IN proteins have been described previously (4).

Our results with cross-linking experiments showed that mutation of conserved residues Pro-109 and Asp-116, which were found to be critical for the endonuclease and integration activities of IN protein in vitro (4), also abolished the DNA cross-linking ability of the protein (Fig. 5, lanes 2 to 4). However, other highly conserved residues, Asp-64 and Glu-152, also found to be essential for IN activity and located in the amino- and carboxyl-terminal halves of the protein, respectively, did not appear to be involved in interacting with DNA, since the cross-linking ability of the proteins with these residues was not impaired (Fig. 5, lanes 6 and 7). Mutation of Thr-115 to Ala, another conserved residue, slightly decreased cross-linking ability (Fig. 5, lane

5). Substitution of Gly-189, a nonconserved residue (Fig. 5, lane 8), marginally decreased cross-linking, but substitution of Thr-112 did not impair the cross-linking reaction (Fig. 5, lane 9). However, it must be noted that cross-linking ability as mentioned here is qualitative and is only used to make relative comparisons. Slight differences between mutant proteins are not necessarily functionally relevant.

We (4) and others (5, 10) have shown previously that conserved residues Asp-64, Pro-109, Asp-116, and Glu-152 are essential for all three catalytic activities of the HIV-1 IN protein in vitro, namely, specific endonucleolytic cleavage, DNA joining or integration, and disintegration. We find, however, that among these conserved residues, Asp-64 and Glu-152 are not required for interaction with DNA, as measured in a UV cross-linking reaction, whereas Pro-109 and Asp-116 appear to be essential. Obviously, these results do not reveal whether those critical residues contact the substrate directly or influence the local conformation of a putative binding domain. Nevertheless, we suggest, from these and previous observations, that Asp-116, as part of the active site of the integrase, may be responsible for binding Mg^{2+} and DNA to form a metal coordination complex. Thus, mutants containing a substitution of Asp-116 are unable to bind DNA, becoming catalytically inactive. This is in contrast to the roles of Asp-64 and Glu-152, which also likely form part of the active site but apparently do not contribute directly to the formation of metal-DNA coordinate complexes.

This study shows, for the first time, possible differences in the roles played by putative active-site residues in DNA binding versus cleavage and integration activities. Our approach has been to identify residues that are in close contact with DNA as being functionally significant. Studies are under way to confirm the importance of these amino acid residues in DNA binding, by performing partial proteolysis of covalently cross-linked DNA-protein complexes and then pinpointing the individual residues cross-linked to DNA by peptide sequencing.

REFERENCES

- Bushman, F. D., and R. Craigie. 1991. Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. *Proc. Natl. Acad. Sci. USA* **88**:1339-1343.
- Bushman, F. D., and R. Craigie. 1992. Integration of human immunodeficiency virus DNA: adduct interference analysis of required DNA sites. *Proc. Natl. Acad. Sci. USA* **89**:3458-3462.
- Chow, S. A., K. A. Vincent, V. Ellison, and P. O. Brown. 1992. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science* **255**:723-726.
- Drelich, M., R. Wilhelm, and J. Mous. 1992. Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro. *Virology* **188**:459-468.
- Engelman, A., and R. Craigie. 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J. Virol.* **66**:6361-6369.
- Haenggi, M., and J. Mous. Unpublished data.
- Johnson, M. S., M. A. McClure, D. F. Feng, J. Gray, and R. F. Doolittle. 1986. Computer analysis of retroviral *pol* genes: assignment of enzymatic functions to specific sequences and homologies with non-viral enzymes. *Proc. Natl. Acad. Sci. USA* **83**:7648-7652.
- Katz, R. A., G. Merkel, J. Kulkosky, J. Leis, and A. M. Skalka. 1990. The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. *Cell* **63**:87-95.
- Katzman, M., R. A. Katz, A. M. Skalka, and J. Leis. 1989. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. *J. Virol.* **63**:5319-5327.
- Khan, E., J. P. G. Mack, R. A. Katz, J. Kulkosky, and A. M. Skalka. 1990. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res.* **19**:851-860.
- Kulkosky, J., K. S. Jones, R. A. Katz, J. P. G. Mack, and A. M. Skalka. 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol. Cell. Biol.* **12**:2331-2338.
- Lafemina, R. L., P. L. Callahan, and M. G. Corgingley. 1991. Substrate specificity of recombinant human immunodeficiency virus integrase protein. *J. Virol.* **65**:5624-5630.
- Leavitt, A. D., R. B. Rose, and H. E. Varmus. 1992. Both substrate and target oligonucleotide sequences affect in vitro integration mediated by human immunodeficiency virus type 1 integrase protein produced in *Saccharomyces cerevisiae*. *J. Virol.* **66**:2359-2368.
- Roth, M. J., P. L. Schwartzberg, and S. P. Goff. 1989. Structure of the termini of DNA intermediates in the integration of retroviral DNA. Dependence on IN function and terminal DNA sequence. *Cell* **58**:47-54.
- Roth, M. J., P. Schwartzberg, N. Tanese, and S. P. Goff. 1990. Analysis of mutations in the integration function of Moloney murine leukemia virus: effect on DNA binding and cutting. *J. Virol.* **64**:4709-4717.
- Sherman, P. A., and J. A. Fyfe. 1990. Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleaving activity. *Proc. Natl. Acad. Sci. USA* **87**:5119-5123.
- van Gent, D. C., Y. Elgersma, M. W. J. Bolk, C. Vink, and R. H. A. Plasterk. 1991. DNA binding of the integrase proteins of human immunodeficiency viruses types 1 and 2. *Nucleic Acids Res.* **19**:3821-3827.
- Vink, C., D. C. van Gent, Y. Elgersma, and R. H. A. Plasterk. 1991. Human immunodeficiency virus integrase protein requires a subterminal position of its viral DNA recognition sequence for efficient cleavage. *J. Virol.* **65**:4636-4644.