Formation of a stable complex between the human immunodeficiency virus integrase protein and viral DNA

Cornelis Vink, Ramon A.Puras Lutzke and Ronald H.A.Plasterk*
Division of Molecular Biology, the Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

Received July 11, 1994; Revised and Accepted August 30, 1994

ABSTRACT

The integrase (IN) protein of the human immunodeficiency virus (HIV) mediates two distinct reactions: (i) specific removal of two nucleotides from the 3' ends of the viral DNA and (ii) integration of the viral DNA into target DNA. Although IN discriminates between specific (viral) DNA and nonspecific DNA in physical in vitro assays, a sequence-specific DNA-binding domain could not be identified in the protein. A nonspecific DNA-binding domain, however, was found at the C terminus of the protein. We examined the DNA-binding characteristics of HIV-1 IN, and found that a stable complex of IN and viral DNA is formed in the presence of Mn2+. The IN–viral DNA complex is resistant to challenge by an excess of competitor DNA. Stable binding of IN to the viral DNA requires that the protein contains an intact N-terminal domain and active site (in the central region of the protein), in addition to the C-terminal DNA-binding domain.

INTRODUCTION

Integration of human immunodeficiency virus (HIV) DNA into DNA of the infected human cell is essential for replication of HIV (1,2). The retroviral DNA integration reaction can be divided in two steps (for reviews see refs 3—9). In the first step, two nucleotides are removed from the 3' termini of the blunt-ended viral DNA. In the second step, the newly generated 3' OH ends of the viral DNA are coupled to sites in the target DNA. Both steps are carried out by a single protein, the viral integrase (IN) protein. Purified, recombinant IN can mediate both site-specific cleavage (or donor cut) and integration (DNA strand transfer) of oligonucleotide substrates that mimic the viral DNA termini (10—15). The chemistry of IN-mediated reactions is similar. In the donor cut reaction, IN makes a specific phosphodiester bond at the viral DNA ends accessible for nucleophilic attack. Although several hydroxyl-group containing compounds can serve as nucleophile in vitro (16,17), it is likely that in vivo H2O is the primary nucleophile. Specific hydrolysis of the viral DNA by IN results in release of a dinucleotide with 5' phosphate-group and a 3' OH group (16,17). DNA strand transfer is a one-step transesterification reaction (16): the viral 3' OH ends are coupled to phosphates in opposite strands of the target DNA. In the case of HIV DNA integration, these phosphates are five nucleotides separated from each other (18,19). The two bases at each 5' end of the viral DNA in the integration intermediate remain free; they are probably removed in the final stage of the integration process, which is mediated by cellular enzymes. The cellular repair machinery is probably also responsible for filling-in of the five nucleotides single-stranded gaps that flank the integrated retroviral DNA. Consequently, the integrated retrovirus has lost 2 bp from both ends and is flanked by a 5 bp direct duplication of the target DNA.

Three distinct regions have been identified in the IN protein. (i) The N terminus of IN is required for donor cut and DNA strand transfer. This region contains four conserved residues, two histidines and two cysteines, and is able to coordinate a Zn2+ ion (20,21). The function of the N terminus is not yet known. (ii) The central part of the protein is highly conserved among all IN proteins and contains a motif which is also found in transposases of several bacterial transposons (22,23). Analysis of both point mutants and deletion mutants of IN has indicated that the central region comprises the active site of the protein. A polypeptide of amino acids 50—194 of the 288 residue HIV-1 IN protein is sufficient to promote disintegration (21,24), an activity which can be regarded as the reversal of DNA strand transfer (25). Three amino acids within the central region appear to be essential for all catalytic activities of IN: Asp64, Asp116 and Glu152 (2,15,26—29). (iii) The C terminus of IN contains a nonspecific DNA-binding domain, which is essential for donor cut and DNA strand transfer (23,24,30,31).

IN requires specific sequences at the ends of the viral DNA for cleavage and integration of these ends (32—37). Specific binding of IN to these sequences, however, could not be detected (30,31,34,38,39). In order to study the interaction of purified IN protein with viral and target DNA, we performed a series of experiments in which viral DNA, target DNA and divalent cations were added at different time-points to reaction mixtures containing IN. IN–DNA interactions were studied both directly, by protein–DNA cross-linking, and indirectly, in assays for catalytic functions of IN. These experiments indicate that a stable complex

*To whom correspondence should be addressed
of IN and viral DNA can be formed when Mn$^{2+}$ is present. Stable, Mn$^{2+}$-dependent binding by IN requires that the N-terminus, C-terminus and central active site of the protein are intact.

**MATERIALS AND METHODS**

**Oligonucleotides**

The strand transfer substrate (substrate #1 in Figure 2A) is a double-stranded oligonucleotide which is identical in sequence to the HIV-1 U5 DNA end from which the two 3′-terminal nucleotides have been removed. The sequence of this substrate is 5′-TTAGTCAGTGGAAAATCTCTAGCA-3′ (the top strand in Figure 2A) and its complement. The donor cleavage substrate (Figure 3A) is identical in sequence to the strand transfer substrate, but contains two additional nucleotides (5′-GT-3′) at the 3′ end of the top strand in Figure 3A. The competitor or target DNA (CP or substrate #2 in Figure 2A) is an oligonucleotide that contains a biotin moiety at the 3′ end of one of its strands. The sequence of this oligonucleotide is 5′-GGGTTCCCTAGATTTCCCACTGT-bio-3′ and its complement. Substrate U5 WT is identical in sequence to the donor cleavage substrate in Figure 3A, but contains a 32P label at the 5′ end of the top strand, instead of at the 3′ end. Oligonucleotide U5 CA > TG is identical to substrate U5 WT, except for mutation of the conserved 5′-CA-3′ into TG (and the complementary mutation in the complementary strand). Substrate U5 MoMLV is identical in sequence to the U5 DNA end of MoMLV: 5′-GACTACCCGTCAGGGGGTCTTTCATT-3′ except for mutation of the conserved 5′-CA-3′ into TG (and the complementary mutation in the complementary strand). Substrate C(10)>G is identical to U5 WT except for a C to G mutation at position 10 from the 3′ end of the top strand in Figure 2A (and the complementary mutation in the complementary strand). The nonspecific substrate (nonspec.) in Figure 4B is identical in sequence to substrate #2 in Figure 2A. Oligonucleotides were labelled with 32P, either at the 5′ end or 3′ end as described previously (17, 35).

**Proteins**

Wild-type HIV-1 IN, the deletion mutants NΔ50 and CΔ66, and active-site mutant D116I were expressed and purified as fusions to maltose-binding protein (MBP) as previously described (24,40).

**Protein–DNA UV cross-linking**

UV cross-linking reactions were done in 10 μl volumes, and contained 20 nM 32P-labelled viral oligonucleotide, 0.3 μM MBP–HIV-1 IN, 20 mM MOPS pH 7.2, 3 mM DTT, 3 mM MnCl$_2$ (or, where indicated, 5 mM MgCl$_2$), 0.1 mg/ml BSA, 0.1% Tween20, 4% glycerol, 50 mM NaCl, 1 mM Tris, 50 μM EDTA and 0.1 mM β-mercaptoethanol. Where indicated, biotinylated competitor oligonucleotide was added to reaction mixtures to a final concentration of 0.5 μM. Preincubation experiments were performed as follows. The IN protein was incubated in reaction buffer in the presence of either divalent cation, the viral oligonucleotide or competitor oligonucleotide, or any combination of these three components. The components that were not added at time-point 0 to the IN protein, were added after 5 or 10 min incubation at room temperature. After preincubation, the eppendorf tubes containing the reaction mixtures were placed on an ice-cooled ‘Chromato-vue’ transilluminator (254 nm, 50 Hz, 0.6 A, Ultraviolet products, Inc.), and irradiated for 8 min. After cross-linking, 10 μl of protein sample buffer was added to the reaction mixtures, which were incubated at 100°C for 5–10 min. Five μl of the samples were loaded onto either 10% or 7.5% SDS–polyacrylamide gels. Cross-linking was analyzed after drying of the gels and autoradiography. Quantitation was done by densitometry, using an Ultrasean XL Enhanced Laser Densitometer (LKB).

**Donor cleavage and strand transfer reactions**

Reaction mixtures were similar as in the cross-linking experiments, and contained Mn$^{2+}$ at 3 mM. After preincubation at room temperature, the reaction mixtures were placed at 37°C for 30 min. Reactions were terminated by addition of 10 μl of formamide loading dye (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) and analyzed on either 12% (strand transfer) or 24% (donor cleavage) denaturing polyacrylamide gels. Reaction products were visualized by autoradiography. Alternatively, in order to quantitate strand transfer activity, reactions were stopped by addition of 70 μl of 5.7 M Tris pH 7.5, 0.6 M EDTA, 1.1 M NaCl. The mixtures were subsequently added to wells of a Streptavidin-coated microtiter plate (Combiplate 8, Labsystems). The biotinylated DNA was bound to the wells by gentle agitation of the microtiter plates for 1 h at room temperature. The wells were washed six times for 5 min with 100 μl of 5 mM Tris, pH 7.5, 0.5 mM EDTA, 1 M NaCl. Finally, biotinylated DNA was eluted from the wells by addition of 40 μl of formamide loading dye, and agitation of the wells for 45 min at room temperature. The amount of radioactivity of the eluted DNA was determined by scintillation counting. A part of the eluted DNA was also analyzed on gel. Donor cleavage activity was quantitated by densitometry.

**RESULTS**

**Effect of Mn$^{2+}$ on DNA binding activity of IN**

The activity of IN on viral DNA ends in vitro is strongly inhibited if an excess of other DNA or RNA species is present in the reaction mixture (34,38,41). Inhibition of IN activity by competitor DNA is less strong, however, if the competitor DNA is added to the reaction mixture after addition of the viral DNA substrate (38). To examine whether this ‘order-of-addition’ effect acts at the level of DNA binding by IN, we sequentially added radiolabelled viral DNA (U5) and competitor DNA to solutions containing HIV-1 IN protein, and tested for DNA binding by UV cross-linking analysis. As competitor DNA (CP), a double-stranded oligonucleotide of random sequence was used which was added to reaction mixtures in a 12.5 times molar excess over viral DNA. To address the role of divalent cations on DNA binding by IN, we also varied the time-points of addition of Mn$^{2+}$ in these preincubation experiments. Purified, recombinant HIV-1 IN requires Mn$^{2+}$ for activity (13,14). After preincubation, the samples were irradiated with UV, and were analyzed by SDS–polyacrylamide gelelectrophoresis and autoradiography. Figure 1A shows the pattern of DNA cross-linking in the various reactions. Efficient cross-linking was only observed when both Mn$^{2+}$ and the viral DNA (U5) were added before the competitor DNA (CP) (lanes 2, 3 and 5 in Figure 1A). The highest efficiency of cross-linking was observed when IN was preincubated with Mn$^{2+}$, and subsequently the viral DNA was added (lane 5). Cross-linking was inefficient when both viral DNA and an excess of competitor DNA were added simultaneously, after preincubation of IN with Mn$^{2+}$ (data not shown). The apparent molecular weight of the major cross-linked
The UV cross-linking experiments suggest that IN stably binds to the viral DNA termini in the presence of Mn$^{2+}$. To determine whether the conditions that result in efficient formation of IN–viral DNA complexes also result in high catalytic activities, we performed a preincubation/strand transfer experiment. An outline of the DNA strand transfer reaction is shown in Figure 2A. Integration of the $^{32}$P-labelled viral DNA substrate (substrate #1 in Figure 2A) into the biotinylated target DNA (substrate #2) results in coupling of the radiolabel to the biotin label. Thus, the amount of radioactivity that is associated with the biotinylated DNA (which can be isolated using Streptavidin-coated microtiter plates) is a direct measure for DNA strand transfer activity of IN. Preincubation of IN with DNA substrates and Mn$^{2+}$ was done for 10 min at room temperature. DNA strand transfer was subsequently allowed to proceed for 30 min at 37°C. No activity was detected if the incubation step at 37°C was omitted (data not shown). Figure 2B shows the result of the preincubation/DNA strand transfer experiment. The strand transfer activity of IN in the various reactions approximately reflects the DNA-binding activity of the protein in the UV cross-linking assay (Figure 1A). The highest activity was found when IN was preincubated with Mn$^{2+}$, before addition of the viral substrate and the target DNA, respectively (lane 5). These data...
Figure 2. Effect of various preincubations on DNA strand transfer activity of IN. (A) Outline of the strand transfer reaction. Substrate #1 is a double-stranded oligonucleotide which is identical in sequence to the HIV-1 U5 DNA end after donor cleavage has occurred. The $^{32}$P label at the 5’ end of the strand that contains the conserved 5’-CA-3’ sequence is indicated with a closed circle. In strand transfer, IN promotes the coupling of the 3’ OH of the top strand of substrate #1 to phosphates in the target DNA, in this case substrate #2 (or CP). Note, that integration can take place in both strands of the target DNA. (B) DNA strand transfer activity after various preincubation reactions. Nomenclature is as in Figure 1. After preincubation at room temperature for 10 min, the reactions were at 37°C for 30 min. Samples were separated on a 12% denaturing polyacrylamide gel. Lane 1 is without IN. ST, strand transfer products. Biotinylated strand transfer products were isolated using a Streptavidin-coated microtiter plate, and were subsequently quantitated. The amount of radioactivity of the eluted DNA was 0, 3.4, 1.2, 0.7, 4.4, 1.4, 1.2, 1.4, 1.6, 1.2 and 1.0% of the total amount of radioactivity that was present in reactions 1–11, respectively.

Figure 3. Effect of various preincubations on donor cleavage activity of IN. (A) Outline of the donor cleavage reaction. The donor cleavage substrate contains a $^{32}$P label (closed circle) in between the two 3’-terminal nucleotides of the top strand. Donor cleavage by IN in the presence of glycerol and Mn$^{2+}$ gives rise to four major products: the recessed viral DNA (in this case unlabelled), a simple dinucleotide (D), a cyclic dinucleotide (C), and a product in which glycerol is coupled to the 3’-terminal dinucleotide via a 5’-phosphodiester bond (16,17). (B) Donor cleavage after various preincubations. Samples were separated on a 24% denaturing polyacrylamide gel. The three radiolabelled cleavage products from (A) are indicated. ‘U5’ depicts the viral DNA substrate from (A); other nomenclature is as in Figure 1. Lane 1 is without IN. The relative amount of activity is 50, 11, 37, 100, 6, 5, 8, 23, 12 and 5% in lanes 2–11, respectively.
suggest that the stable interaction of IN with the viral DNA that is seen by UV cross-linking is a functional interaction, which results in high catalytic activities of the protein. This suggestion is supported by the similar divalent cation preference of IN both for UV cross-linking and strand transfer; HIV-1 IN is not active in strand transfer in the presence of either Mg$^{2+}$ or Ca$^{2+}$ (data not shown).

In the experiments described above, the competitor DNA served both as competitor for viral DNA binding and as target for integration. The results of these experiments are, therefore, the combined effect (of stable) binding of IN to the viral DNA termini and the availability of competitor DNA to serve as target. To exclusively examine the interaction between IN and the viral DNA ends, we also performed the various preincubation experiments with a donor cleavage substrate, i.e. a viral DNA substrate of which the two 3'-terminal nucleotides of one of its strands can specifically be removed by IN. A schematic representation of the donor cut reaction is shown in Figure 3A. Specific products D, C and G are end products of this reaction (16,17) and are, therefore, the result of interaction of IN with the viral DNA only. The activities of IN in the donor cut reaction (Figure 3B) reflect those in the strand transfer reaction and UV cross-linking assay: donor cleavage is highly efficient when IN is preincubated with Mn$^{2+}$, and viral DNA is added before competitor DNA (lane 5 in Figure 3B).

Specific interactions between IN and the viral DNA termini

The IN–DNA cross-linking experiments demonstrated stable binding of IN to the viral DNA ends. The assays for catalytic IN functions suggested that this binding is productive, leading to relatively high levels of strand transfer or cleavage products. To examine the specificity of DNA binding by IN, we tested several different oligonucleotide substrates in the UV cross-linking assay (Figure 4A and B). One of the substrates we tested is a viral DNA substrate in which a 5'-CA-3' sequence is changed into 5'-TG-3' (substrate U5 CA>TG in Figure 4A). This 5'-CA-3' dinucleotide is essential for cleavage and integration by IN (34—37). Substrate U5 CA>TG is, however, cross-linked to IN, albeit somewhat less efficiently than the wild-type substrate (U5 WT) (compare lanes 5 and 11 in Figure 4A). A possible explanation for the binding of IN to mutant U5 CA>TG is that sequences outside the 5'-CA-3' sequence contribute to specific sequence recognition by IN. A substrate that is cleaved at very low levels by HIV IN is a substrate that resembles one of the ends of the Moloney murine leukemia virus (MoMLV) DNA (U5 MoMLV in Figure 4A) (35). This substrate is cross-linked less efficiently than the wild-type substrate after competition with an excess of nonspecific DNA (compare lanes 5 and 17). Substrate C(10)>G in Figure 4B resembles a wild-type viral DNA end except for the presence of a single point mutation. This mutation does not have an effect on donor cut and strand transfer activity of IN (35). In accordance with this, substrate C(10)>G is cross-linked in a similar way as the wild-type substrate (compare lanes 2—6 with lanes 8—12 in Figure 4B). We also tested an oligonucleotide of random sequence (nonspec. in Figure 4B). This oligonucleotide was not efficiently bound by IN, i.e. binding of this substrate by IN could easily be competed by addition of competitor DNA (compare lanes 5 and 17). These results show that although IN possesses sequence specificity for stable DNA binding, this specificity is limited. Limited substrate sequence specificity was also found for both the donor cut and strand transfer activity of HIV-1 IN (34—37).

DISCUSSION

Mn$^{2+}$ is required for stable DNA binding by IN

Purified IN proteins bind DNA in a nonspecific fashion (23,24,30,31,34,38,42). In contrast, the viral DNA ends are cleaved and integrated by IN in a specific manner (34—37). We show here that in the presence of Mn$^{2+}$ stable IN–viral DNA complexes are formed which are resistant to challenge by an excess of competitor DNA. These complexes are less efficiently generated with nonspecific DNA substrates.

What is the role of divalent cations?

Previously, a DNA-binding domain of HIV-1 IN was identified between amino acids 200 and 270 of the 288 residue protein (24). Proteins that contain this domain are able to bind DNA nonspecifically. As a consequence, deletion mutant NA50 and active-site mutant D116I bind DNA in UV cross-linking assays, whereas mutant CA66 does not. However, the stable DNA binding that is seen for wild-type HIV-1 IN, is not observed for mutants NA50 and D116I. This indicates that the C-terminal DNA-binding domain of IN is sufficient for nonspecific (Mn$^{2+}$-independent) DNA binding, but that for efficient, stable complex formation between IN and the viral DNA ends both the N-terminus and central active site of the protein are necessary. Residue Asp116 was shown to form part of a triad of amino acids (22,23). The role of these acidic amino acids in the catalytic activities of IN was suggested to be the coordination of divalent metal ions, in analogy to the situation for other enzymes that catalyze phosphoryl transfer reactions (27,28). The preincubation/UV cross-linking experiments we report here indicate that at least one of the invariant acidic amino acids (Asp116) plays an important role in stable binding of IN to its viral DNA substrate. We speculate
that the three essential amino acids function in the coordination of one or more divalent metal ions, which is important for the correct positioning of the viral DNA ends in the active site of IN.

Stable donor DNA binding by transposases other than IN

One of the most thoroughly studied transposition reactions is that of phage Mu. The transposase of Mu (MuA), has been shown to bring the two ends of the phage DNA together, to cleave these ends, and mediate the strand transfer reaction (43—45). There are clear similarities between phage Mu transposition and retroviral DNA integration. For instance, the chemical steps that occur in both donor cleavage and DNA strand transfer are similar in these systems (16,46). In both cases a single-strand cut is made at the ends of the elements to expose a 3' OH group. Subsequently, the 3' OH group is coupled in one step to phosphates in the target DNA. To date, there is little information about protein—DNA intermediates in retroviral DNA integration. In contrast, various stages in the phage Mu transposition reaction have been analyzed in detail. Three stable complexes between MuA transposase and Mu DNA were identified (45,47,48). Stable binding of MuA to the Mu DNA ends is dependent on the presence of divalent metal ion (48). A stable MuA-DNA complex (the stable synaptic complex or SSC) is generated in the presence of either Mg2+, Mn2+ or Ca2+. Ca2+, however, does not allow donor cleavage to occur, resulting in accumulation of the SSC (48). This complex is resistant to challenge by an...
excess of competitor DNA, and is converted into the so-called cleaved donor complex (CDC) and strand transfer complex (STC), respectively, after addition of Mg$_2^+$. The stable complex between HIV-1 IN and the viral DNA termini which we describe here, has similar properties as the stable complexes that are formed during Mu transposition. Similarly as for MuA–Mu DNA complexes, the generation of IN–viral DNA complexes is dependent upon the presence of divalent metal ions (Mn$^{2+}$). In addition, IN–viral DNA complexes are resistant to challenges with an excess of competitor DNA; this is irrespective of whether the complex is formed with a donor cut substrate or a strand transfer substrate. The role of Ca$_2^+$ in assembly of the SSC of Mu is not seen in the formation of IN–viral DNA complexes; only a marginal increase in strand transfer activity was seen when IN was preincubated with Ca$_2^+$, and the reaction was done in the presence of Mn$^{2+}$ (data not shown).

What is the high order IN–viral DNA complex that is involved in cleavage and integration?

Transposition is generally carried out by multimeric protein complexes, e.g. in the case of Mu (see above), Tn10 (49) and Tn7 (50). The active form of the retroviral IN proteins is probably also a multimer; this was supported by both kinetic data (51) and in vitro complementation experiments (40,52). Models have been proposed in which dimers (53) or tetramers of IN (9,16) carry out the concerted integration of both viral DNA ends into target DNA. However, it is difficult to test these models since purified, recombinant HIV IN almost exclusively mediates single-end strand transfer events (for a recent discussion see ref. 9). The stable complex that we describe here is probably not a full equivalent of the complexes that are generated in vivo. For example, a difference between the in vitro generated complexes that we describe here, and the preintegration complexes that are found in vivo is the divalent cation requirement. In vitro Mn$^{2+}$ is required for activity, in vivo probably Mg$^{2+}$. In addition, the concerted integration of both viral DNA ends is still not seen in vitro (data not shown). It is possible that factors other than viral DNA and IN are required in formation of a functional complex which is able to undergo double-end strand transfer events, or alternatively that the correct architecture of the preintegration complexes is not achieved by merely mixing of the constituents in vitro.

ACKNOWLEDGMENTS

We thank Drs P.Borst, K.Simmen, G.Rudenko, J.C.Vos and F.M.I.van den Ent for critical reading of the manuscript. This work was supported by the Netherlands Organization for Scientific Research (N.W.O.), and by a grant from Glaxo Group Research Limited to C.V.

REFERENCES

