Interactions of the Human T-cell Leukemia Virus Type-II Integrase with the Conserved CA in the Retroviral Long Terminal Repeat End

Received for publication, August 24, 2000, and in revised form, January 26, 2001 Published, JBC Papers in Press, February 2, 2001, DOI 10.1074/jbc.M007754200

Tan Wang‡§, Andrew J. Piefer‡¶, and Colleen B. Jonsson§

From the Department of Chemistry and Biochemistry, New Mexico State University, Las Cruces, New Mexico 88003

The retroviral integrase (IN) catalyzes the covalent insertion of the retrovirus genome into the host chromosome following reverse transcription of the viral RNA genome to a DNA intermediate (1, 2). Successful integration of the viral DNA by IN requires three distinct enzymatic events: 3’-processing, strand transfer, and gap repair (3, 4). In the initial step of integration, IN catalyzes the hydrolysis of the phosphodiester bond between the phylogenetically conserved CA and the two or three terminal nucleotides from each 3’ end of the viral DNA. 3’-Processing of the newly synthesized viral DNA intermediate occurs in the cytoplasm (5). The strand transfer reaction follows in the nucleus where IN uses the 3’-hydroxyl group of each viral end to catalyze a direct nucleophilic attack at staggered 5’-phosphates in the phosphodiester backbone of the host DNA (6). IN can also catalyze the reverse of strand transfer, termed disintegration (7). Each of these reactions has been characterized in vitro with double-stranded oligonucleotide substrates that mimic the end of the viral DNA molecule (Fig. I A). The final step, gap repair, requires host cellular enzymes and perhaps IN to remove the unpaired nucleotides at the 5’ ends of the viral DNA and complete the filling-in and joining of the single-strand gaps between the viral and host DNA.

Numerous studies support the contributions of the conserved CA as necessary for recognition and catalysis by IN (5, 6–12). In addition, IN relies on at least one additional nucleotide upstream of the CA to distinguish its substrate from noncognate viral long terminal repeat ends. Although most INs have some level of activity with heterologous viral ends (13). IN shows an equal binding affinity to both nonspecific and specific (viral) DNA substrates (14–16), and therefore, precise recognition of the CA may occur during catalysis. Previously we showed the conserved CA/G/T, particularly at the guanine and adenosine residues, share a common scaffold of molecular features which are recognized by HIV-1, HTLV-2, and Moloney murine leukemia virus INs during 3’-processing or dinucleotide cleavage (17). Isoteric nucleotide analogue substitutions in these positions suggested both major groove and minor groove interactions in this region with IN (17). For example, insertion of an amino group into the minor groove at the Ade or position 3 in the plus strand or its removal at Gua or position 4 in the minus strand decreases activity in each of these INs. In contrast, these INs tolerate the insertion of a methyl group into the major groove of their U5 substrates at positions 3 and 4 in the plus strand and position 3 in the minus strand. However, the introduction of a methyl group using the nucleotide analogue O6-methylguanine (6-MeGua) at position 4 in the minus strand of the U5 substrate produces a dramatic decrease in the level of 3’-processing activity.

In this work, we extended the use of analogue nucleotide substitutions at Gua at position 4 in the minus strand to characterize the essential components of viral DNA substrate recognition and catalysis by HTLV-2 IN. Clear distinctions were observed in the major and minor groove contacts required for 3’-processing, strand transfer, and disintegration. However, in the process of these studies, a novel second site cleavage activity was uncovered in the HTLV-2 U5, but not the U3
Interaction of HTLV-2 IN with the Retroviral DNA Termini

14711

were assembled on ice with 1.0 pmol of substrate and 10 nM to 2.1 as described, and hybridized at a molar ratio of 1:1. Binding reactions

mM NaCl, 10 mM imidazole) supplemented with 1 mg/ml lysozyme

duplicate data sets.

above, and results quantified from a minimum of two experiments with

detergent. The blot was air-dried, and exposed to a phosphorscreen.

final volume of 20 l. The ratio of enzyme to substrate was varied from

1:5 to 40:1. Reactions were conducted at room temperature for 20 min

leakage, washed with 200 l of the reaction buffer without metal and

m of loading dye (95% formamide, 20 mM

results were required for 3’-processing at the upstream ACA.

were required for 3’-processing at the upstream ACA engineered into

the U3 substrate. As predicted, the modified substrate, U3–4/6-MG substrates by HTLV-2 IN and, of interest, this site follows a second internal CA in the HTLV-2 U5 end substrate. A low amount of additional heterogeneous cleavages were also noted adjacent to the major novel cleavage site.

To further explore the mechanism involved in promoting cleavage at the internal site in the HTLV-2 U5 substrate, 6-MeGua was introduced in the minus strand at position 8 in addition to the original analogue substitution at position 4. It was expected that IN’s cleavage activity would be blocked at the conserved CA and the second internal CA. Surprisingly, we observed an additional distinct pattern of cleavage following catalysis with IN. Imprecise cleavage occurred at two sites in the substrate, following positions 5 and 9; both of which were 5’ to the conserved CA and the internal CA (Fig. 2, lane 8). A time course of this reaction revealed no relationship between the cleavage of these two sites; i.e. the two sites were cleaved independently. 2 The unique activities displayed by IN with the single and double modifications suggested that the effect of 6-MeGua may not cause a simple block in the hydrogen bonding between IN and Gua in the major groove as interpreted for other protein-DNA interactions (20). We hypothesized that 6-MeGua may generate a change in the geometry or conformation of van der Waals and hydrogen bonding interactions at the ACA through a structural alteration in the G-C base pair that affects the manner in which IN recognizes, binds, and cleaves the DNA substrate. In the following, we first address whether the upstream CA or ACA was required for cleavage.

ACA Is Sufficient for Catalysis by the HTLV-2 IN, But Not Cleavage at the Internal Site—To determine whether the ACA context was required for the observed internal cleavage activity, we examined HTLV-2 IN activity with the HTLV-2 U3 substrate and the U3 substrate with the 6-MeGua substituted opposite of the C in the conserved CA (U3–4/6-MG). The WT HTLV-2 U3 sequence was chosen because it does not have an internal CA at this position, although it does have a CA further upstream (Fig. 1C). The expected dinucleotide cleavage pattern was observed with the unmodified, WT U3 end substrate (Fig. 3A, lane 2), while the U3–4/6-MG substrate showed a 5-fold decrease in this product (Fig. 3A, lane 4) as compared with the WT U3 substrate. Neither the WT U3 nor U3–4/6-MG substrates showed internal cleavage when incubated with HTLV-2 IN. Previously, we showed that the ACA sequence was important for recognition and 3’-processing of the U5 substrate by HTLV-2 IN by comparison with its activity on heterologous retroviral substrates from HIV-1, HIV-1, and Moloney murine leukemia virus (13). Therefore, we hypothesized that mutation of the U3 substrate sequence from 5’-GTT to 5’-ACA in positions 7, 8, and 9 coupled with the 6-MeGua substitution at position 4 in the terminal conserved CA/G/T were required for cleavage of the ACA engineered into the U3 substrate. As predicted, the modified substrate, U3–4/6-MG-ACA, was processed by the HTLV-2 IN at the internal CA (Fig. 3A, lane 6). These experiments support the hypothesis that the modifications created at position 4 in the minus strand were required for 3’-processing at the upstream ACA.

The above results suggested that the ACA was necessary and sufficient for recognition and catalysis of the viral end sub-

virus, and HIV-1 (17). 6-MeGua and Hyp introduce changes in the major and minor groove, respectively (Fig. 1D). Further analysis of the HTLV-2 IN reactions catalyzed with these two substrates, 4-Hyp and 4/6-MG, revealed a novel pattern of cleavage (Fig. 2, lanes 4 and 6) as compared with the wild type (WT) U5 end substrate (Fig. 2, lane 2). In each modified substrate, a novel cleavage site occurred 3’ to position 7 (Fig. 1C), and, of interest, this site follows a second internal CA in the HTLV-2 U5 end substrate. A low amount of additional heterogeneous cleavages were also noted adjacent to the major novel cleavage site.

EXPERIMENTAL PROCEDURES

Oligonucleotide Substrates—Oligonucleotide sequences corresponding to the wild type (WT) U5 and/or U3 ends of the HTLV-2 genomes (Fig. 1C) were used as substrates in enzymatic assays as described previously (17). Oligonucleotide synthesis, introduction of 2’-deoxyxynucleoside phosphoramidite analogues, and high performance liquid chromatography purification of oligonucleotides were performed at Integrated DNA Technologies. Oligonucleotide size standards (8–32) were purchased from Glen Research. Oligonucleotides were further purified on 20% denaturing polyacrylamide gels, 5’-end-labeled with [γ-32P]ATP (PerkinElmer Life Sciences) and T4 DNA kinase (New England Biolabs), and hybridized to complementary strands as described previously (13).

Protein Purification—HTLV-2 IN was expressed in Escherichia coli BL21(DE3) cells and purified as hexahistidine-tagged fusion proteins from the insoluble fraction as described previously (17, 18). For native preparations of HTLV-2 IN, cultures were prepared as previously described (17) with the following modifications in the cell lysis. Pellets were resuspended in 10 ml of extraction buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 10 mM imidazole) supplemented with 1 mg/ml lysozyme (Sigma) and EDTA-free protease inhibitor tablets (Roche Molecular Biochemicals). The resuspended pellet was incubated on ice for 30 min. Triton X-100 was added to a 1% (v/v) concentration and the resulting solution was homogenized by douncing. The viscous cell lysate was then cleared by sonication, 5 × 30 s with 1 min rest in between pulses on ice. Clearing of lysates and native His-Tag IN purifications were performed as described (17). Protein concentrations were measured by the Bradford method (19) using the Bio-Rad Micro-Assay.

Integration and Disintegration Assays—Reaction buffer for HTLV-2 IN contained 25 mM MOPS (pH 7.2), 10 mM β-mercaptoethanol, 10% (v/v) glycerol, 0.75 mM CHAPS, and 7.5 mM MnCl2. Reactions were assembled in reaction buffer with 1 pmol of substrate and 0.013 µg/µl HTLV-2 IN in a final volume of 15 µl at 37 °C for 1 h. Reactions were terminated by addition of 10 µl of loading dye (95% formamide, 20 mM EDTA, 0.05% bromphenol blue, 0.5% xylene cyanol). Reaction products were separated on 20% polyacrylamide denaturing gels, subjected to autoradiography or PhosphorImager screens (Molecular Dynamics). Products were quantified with ImageQuant software (Molecular Dynamics).

The amount of product was calculated from a minimum of three separate trials for each experiment performed in duplicate or triplicate.

Filter Binding Assays—U5 end substrates were purified, 32P-labeled as described, and hybridized at a molar ratio of 1:1. Binding reactions were assembled on ice with 1.0 pmol of substrate and 10 nm to 2.1 µM HTLV-2 IN in reaction buffer (25 mM MOPS, pH 7.2, 10 mM β-mercaptoethanol, 10% (v/v) glycerol, 7.5 mM MnCl2, and 0.75 mM CHAPS) in a final volume of 20 µl. The ratio of enzyme to substrate was varied from 1.5 to 40:1. Reactions were conducted at room temperature for 20 min for complex formation. Each reaction was slot blotted onto a nitrocellulose, washed with 200 µl of the reaction buffer without metal and detergent. The blot was air-dried, and exposed to a phosphorscreen. Reactions were quantified from PhosphorImager data as described above, and results quantified from a minimum of two experiments with duplicate data sets.

RESULTS

Substitution of Hyp or 6-MeGua at for Gua at Position 4 Blocks Precise Recognition and Catalysis of the U5 End Substrates by HTLV-2 IN—Previously, we noted that the replacement of guanine (Gua) at position 4 in the minus strand of the U5 substrate with 6-MeGua (4/6-MG) or Hyp (4-Hyp) decrease 3’-processing activity as compared with wild type U5 end substrate for several INs, HTLV-2, Moloney murine leukemia

2 T. Wang and C. B. Jonsson, unpublished observations.
strates by HTLV-2 IN, and that such a sequence may be recognized internally in a non-U3 or U5 context. To determine whether the ACA was alone sufficient for 3'-processing, non-U3 or U5 end substrates (NUS) with ACA (3-ACA NUS) and without (NUS) were designed and tested (Fig. 1C). 3'-Processing of the 3-ACA NUS was observed (Fig. 3B, lane 7). A random substrate, included as a control (NUS), was recognized, but was processed at much lower level compared with WT (Fig. 3B, compare lanes 5 and 3). Three additional random substrates, 5-ACA NUS, 6-ACA NUS, and 7-ACA NUS (Fig. 1B), were designed to test whether HTLV-2 IN could recognize and process the ACA when placed internally in a random substrate context. These substrates were examined in standard reactions along with the WT U5 end substrate (Fig. 3B, lane 2). HTLV-2 IN mediated catalysis was noted in reactions in which the substrate, ACA-NUS, had ACA placed at positions 3 through 5 (Fig. 3B, lane 7), positions 5 through 7 (Fig. 3B, lane 9), or positions 6 through 8 (Fig. 3B, lane 11). These reactions produced products that were chiefly composed of internal cleavages 3' to the internal ACA. Thus, the 3-ACA reaction produced an 18-mer, the 5-ACA reaction resulted in a 16-mer, and the 6-ACA reaction produced a 15-mer. NUS substrates that contained ACA at positions 7 through 9 were not 3'-processed by the HTLV-2 IN (Fig. 3B, lane 13). This suggests that IN can only gain access to internal ACA sequences if they are placed within the first 6 bases of the 3' end. Strand transfer was evident with both the 5-ACA NUS and 6-ACA NUS substrates (Fig. 3B, lanes 9 and 11). This serves as further confirmation that the HTLV-2 IN relies solely on the ACA for complete processing of these substrates.

Major Groove Modifications or a Base Pair Mismatch at Position 4 Promote Internal Cleavage—To explore whether a simple lack in the hydrogen bonding contact in the major groove in Gua at position 4 was promoting catalysis at the second site, three additional modifications were made. 6-Thio-Gua, 2-aminopurine (2-AP), or Cyt (Fig. 1D) were substituted for Gua in the U5 substrate (Fig. 1C). The 6-thio-Gua is a conservative substitution in that the sulfur atom of the thio-keto group can act as a weak hydrogen bond acceptor, and unlike the 6-MeGua substitution, 6-thio-Gua maintains Watson-Crick base pairing with Cyt (21, 22). However, the hydrogen bond length for N-H–S distance is about 0.4 Å.
greater than the N-H–O distance (21, 22). Furthermore, the van der Waals radius of sulfur is 0.45 Å greater than oxygen and this may introduce propellar distortion. 2-AP simply removes the carbonyl functional group from Gua, and the Cyt substitution created a mismatch at this position.

Dinucleotide cleavage of the 4/6-thio-Gua substrate by HTLV-2 IN showed levels similar to the WT substrate (Fig. 4A, compare lanes 2 and 4). This suggested that carbonyl oxygen in the major groove might not be an important hydrogen bonding acceptor during recognition or catalysis of the U5 end substrate by IN. A smaller amount of cleavage, 1.8-fold decreased as compared with WT dinucleotide level, was also noted at the second site (Fig. 4A, lane 4). From this result, one might predict that complete removal of the carbonyl group would not effect 3'-processing. However, catalysis of the substrate with the 2-AP substitution showed a 4-fold decrease in 3'-processing (Fig. 4B, lane 4). This argues for interactions at the major groove in this position. Furthermore, this suggests that we are observing two distinct activities with these substitutions. One activity results from the absence of a major groove acceptor, which causes a decrease in dinucleotide cleavage. The other results from the presence of the isosteric modification, which promotes cleavage at the upstream ACA.

Previously, Scottoline et al. (23) has shown that the insertion of 2-base pair mismatches immediately 3’ to an internalized CA end will promote internal cleavage of viral DNA substrates. In contrast to this result, we observed internal cleavage at a distance of 3 base pairs when we introduced 6-MeGua in the position 4 of the minus strand, i.e., the third base pair downstream from the internal CA. To further probe the influence of distortion at the position 4, a mismatch was made by substitution of Cyt for Gua (4C/C, Fig. 1C). A mismatch can create much more distortion in a B-DNA than isosteric nucleotide modifications. 3'-Processing of this substrate by HTLV-2 IN produced cleavage at the internal CA (Fig. 4A, lane 6). Similar results were also observed for a G-G base pair substitution at this site.3 A summary of activities of several of the modified substrates used in the cleavage assays is presented in Fig. 5. The result obtained with the 4 C/C substrate, combined with the isosteric substitutions presented above, suggests that the greater the local structural distortions at position 4, the greater the decrease in the dinucleotide product and increase in catalysis at the upstream site. In summary, these results reflect two modes by which IN may interact with the viral end: one which relies on interactions at the major groove contact at this position for 3'-processing, and second, a mechanism that directs recognition and cleavage to the upstream site in the absence of a “correct end.”

Effect of Minus Strand Substitutions on Strand Transfer and Disintegration Activity of HTLV-2 IN—We were also interested in what effect both major and minor groove alterations had on recognition and catalysis of the strand transfer and disintegration substrates. Previous work with the disintegration substrates has strongly argued in favor of an indirect as opposed to direct sequence readout of the substrate (13, 24–26). 6-MeGua and Hyp were substituted at position 4 in the minus strand of precleaved and disintegration substrates of HTLV-2 (Fig. 1, A and B). In contrast to our results with the blunt end substrates in the 3’-processing reaction which showed a 5-fold decrease with Hyp (Fig. 2), the precleaved substrate containing the Hyp substitution, 4/Hyp, showed levels of activity similar to the WT substrate in HTLV-2 IN mediated reactions (Fig. 6A, compare lanes 2 and 4). Moreover, strand transfer activity was not detected when the 6-MeGua substitution was introduced into position 4 (Fig. 6A, lane 6) or positions 4 and 8 (Fig. 6A, lane 8). Comparison of the strand transfer reaction products generated by HTLV-2 IN with WT, 4/6-MG, and the 4,8/6-MG blunt substrates showed cleavage at the internal CA in the 4/6-MG, but not the 4,8/6-MG or WT substrate (see Fig. 2). This suggests the internally processed 4/6-MG substrate was not viable for subsequent strand transfer events. The major groove substitution, 2-AP, was also introduced into precleaved substrate. This substrate showed no strand transfer products when incubated with HTLV2 IN as compared with the WT substrate (Fig. 4C, lane 4).

Reactions with HTLV-2 IN and modified disintegration substrates were also examined. Similar to IN catalyzed strand transfer reactions with the precleaved substrate containing the Hyp substitution, the Hyp replacement at position 4 in the disintegration substrate had little effect on IN catalysis as compared with activity with the WT disintegration substrates. This substrate yielded a 1.3-fold decrease in disintegration products as compared with the WT substrate (Fig. 6B, compare lanes 2 and 4). PhosphorImager analysis of the products generated from the 4/6-MG and the 4,8/6-MG disintegration substrates showed a greatly reduced level of activity, an 8–10-fold decrease, as compared with the WT disintegration substrate (Fig. 6B, lanes 6 and 8). The introduction of 2-AP into position 4 reduced disintegration activity 10-fold (Fig. 6D, lane 4) as compared with WT (Fig. 6D, lane 2).

Major Groove, not Minor Groove or Mismatch Substitutions in the U5 End Substrate, Enhance the Binding Affinity of HTLV-2 IN—The activities observed in the previous experiments could be due to binding and/or catalytic interactions with the various substrates. To define the contribution of the initial DNA binding interactions, we have employed a nitrocellulose filter binding assay to measure the DNA binding affinity of IN for several of the U5 substrates used herein. It has been demonstrated that IN displays increased specificity for viral DNA ends in the presence of divalent metal ions (27). There-

---

3 T. Wang, A. J. Piefer, and C. B. Jonsson, unpublished results.
fore, binding reactions were performed in the presence of metal under standard reaction conditions using natively purified HTLV-2 IN and HTLV-2 U5 and modified substrates.

Four types of modified U5 substrates were analyzed; those with major groove and minor groove modifications and those with mismatches and substitutions in the conserved C/G base pair at position 4. An additional random substrate, Ran, was also included. The minor groove substitution Hyp, located at position 4 in the minus strand, bound significantly less than the WT HTLV-2 U5 substrate at all concentrations (Fig. 6). Its level of binding was similar to the random substrate, Ran. Substitutions in the major groove of the conserved G included 4/6-MeGua and 2-AP (Fig. 7). Both of these substrates behaved equivalently in the filter binding assay. They bound more than 10 to 15%, respectively, higher than the WT substrate at 3.6 × 10⁻² and 7.2 × 10⁻² µg/µl, and modestly higher at lower concentrations of IN. The concentration, 3.6 × 10⁻² µg/µl, was that used in all previous integration and disintegration experiments shown herein. The 4/6-MeGua and 2-AP substrates also bound at significantly higher levels than the Ran substrate. The mismatch substrate, 4C/C, bound nearly the same as the Hyp and Ran substrates, but significantly less than the WT HTLV-2 U5 substrate (Fig. 7). Substrates containing the context of a viral U5 end with a substitution at the conserved 4 position in the minus strand, 4G/C and 4T/A, showed the lowest binding of all the substrates examined; more than 2-fold less than the WT substrate at the two higher concentrations studied. These substrates have no 3'-processing activity.⁴ These substrates were also bound significantly less than the Ran substrate, 2–3-fold. Complexes were not observed when IN was not added to the reaction.⁵ In summary, the major groove modifications showed an enhanced binding of the U5 substrate to IN, while the minor groove modification had a decreased binding affinity. The mismatch substrate did not show the enhanced binding, and this suggests that the interaction of IN with this substrate is through a mechanism distinct from that of the 4/6-MeGua and 2-AP substrates. The low binding affinity of IN to the 4G/C, and 4T/A, also indicates an interaction with IN that is unique to these substrates. These data suggest that IN was able to recognize some, but not all, of the key nucleotide functional groups in position 4. Since some determinants were present, it is possible that these substrates were not recognized as targets, but as faulty viral ends.

Fig. 3. 3'-Processing activity of HTLV-2 IN with HTLV-2 WT U3, modified U3, and NUS substrates. A and B, HTLV-2 IN reactions with the WT U5 substrates are shown in lane 2 of each panel. A, HTLV-2 IN reactions are shown for the WT U3 substrate (lane 2), and modified U5 substrates with a 6-MeGua substitution at position 4 in the minus strand (4/6-MG, lane 4), and the 6-MeGua at position 4 in the minus strand and ACA substitutions in positions 7–9 of the U3 plus strand (4/6-MG ACA, lane 6). B, reactions with either the NUS substrate or NUS substrates with substitutions of ACA at varying positions in the plus strand are shown for: lane 5, NUS; lane 7, positions 5 through 7 (5-ACA NUS); lane 9, positions 5 through 7 (5-ACA NUS); lane 11, positions 6 through 8 (NUS 6-ACA); lane 13, positions 7 through 9 (NUS 7-ACA). Even lanes 2–12 show reactions done in the presence of substrate and the absence of protein. Symbols: Prd, products; sub, substrate. Lanes 1 and 14 contain a labeled oligonucleotide size marker.

Fig. 4. 3'-Processing activity of HTLV-2 IN with HTLV-2 WT and modified U5 substrates. In panels A and B, HTLV-2 IN reactions are shown with the WT U5 substrate (lane 2), and modified U5 substrates at position 4 in the minus strand. Modifications included substitution of Gua for: A, lane 3, 6-thio-Gua (4/6-thio) and lane 6, Cyt (4C/C); B, lane 3, 2-AP. Odd lane numbers in each panel show reactions performed with substrate in the absence of protein.

⁴ T. Wang and C. B. Jonsson, unpublished information.
⁵ A. J. Piefer and C. B. Jonsson, unpublished observations.
**DISCUSSION**

Similar to other INs, the 3'-processing activity of HTLV-2 IN tolerates the insertion of a methyl group into the major groove of the HTLV-2 U5 substrate at positions 3, 4, 5, and 7 in the plus strand and position 3, 5, and 7 in the minus strand, but not at position 4 in the minus strand (17). Introduction of a methyl group at the carbonyl oxygen at position 6 in the carbonyl ring decreased 3'-processing activity and promoted cleavage activity at an upstream CA. The sequence ACA was shown to be required for this novel activity. Two plausible hypotheses for the effect of the 6-MeGua on the 3'-processing reaction were explored. Briefly, our first hypothesis suggested the O6-methyl group might interfere with 3'-processing of the U5 substrate because it blocks the potential hydrogen bonding acceptor site in the major groove (20, 28). However, it has also been reported that 6-MeGua can change the local shape of a base pair (29). Therefore, we alternatively proposed that this substitution could distort the O-P torsion angle, and thereby, interfere with 3'-processing of the U5 substrate through a change in the local structure of the DNA. Several lines of evidence suggest the loss of the 3'-processing activity was due to a loss in the hydrogen bonding in the major groove by the introduction of the methyl group with the 6-MeGua substitution, and not due to a decreased affinity of IN for the substrate. First, U5 substrates substituted with 2-AP for Gua showed a decrease in 3'-processing, and showed WT levels of binding. Second, while the Cyt mismatch at this position resulted in a loss of activity, which indicated that the local structure was not recognized, filter binding experiments showed that the ability of HTLV-2 IN to bind the substrate was poor. In contrast, filter binding analysis of IN with U5 substrates containing 6-MeGua or 2-AP showed a very high binding affinity as compared with the WT U5 substrate. Therefore, the data suggest that the loss of 3'-processing activity by the 2-AP or 6-MeGua substitutions were due to a loss in hydrogen bonding and not disruption of the local structure. A summary of these results is shown in Table I.

The available data suggest that the major groove at position 4 does contribute to hydrogen bonding interactions, however, the data also suggest that the distortion in the structure of the DNA molecule promotes cleavage of the upstream ACA. We propose the following model for the retroviral integrase reaction pathway based on these observations (Fig. 8). IN first binds to the retroviral DNA end (Fig. 8, step 1). Our data suggests that minor groove contact at the amine at position 2 of the Gua was a major determinant for a stable IN/substrate interaction of HTLV-2 IN with the Retroviral DNA Termini

![Image](https://example.com/image)

**FIG. 5. Summary of HTLV-2 IN 3'-processing activity with U5 WT, modified U5, and modified U3 substrates.** HTLV-2 IN reactions are shown with the WT U5 substrate (lane 3), an oligonucleotide size marker (lanes 1 and 16), and modified U5 and U3 substrates at positions 4 and/or 8 in the minus strand. Modifications included substitution of Gua for: lane 5, Cyt (4C/C); lane 7, 2-AP; lane 9, 6-thio-Gua (4/6-thio); lane 11, 6-MeGua (4/6-MG); lane 13, 6-MeGua substituted at both positions 4 and 8 (4,8/6-MG); and lane 15, 6-MeGua substituted at position 4 of the U3 minus strand and containing ACA substituted at positions 7–9 of the U3 plus strand (U3 4/6-MG ACA). Even lane numbers 2–14 show reactions performed with substrate in the absence of protein. Lanes 1 and 16 contain a labeled oligonucleotide size marker. The asterisk denotes the cleavage product of the U3 4/6-MG ACA substrate. Due to the different composition of this substrate, both the substrate and the product do not run true to the size marker.

**FIG. 6. Strand transfer and disintegration activity of HTLV-2 IN with modified U5 substrates.** HTLV-2 IN reactions are shown with the WT U5 substrate, and modified strand transfer (panels A and C) and disintegration U5 substrates (panels B and D). Modifications in substrates were made at Gua in position 4 in the minus strand. In panels A and C, HTLV-2 IN strand transfer reactions are shown for: A, lane 2, the WT U5 precleaved substrate (Wt); and modified strand transfer substrates: lane 4, Hyp (4/Hyp); lane 6, 6-MeGua (4/6-MG); lane 8, 6-MeGua substituted at positions 4 and 8 in the minus strand (4,8/6-MG). C, lane 2, the WT U5 precleaved substrate (Wt); and modified strand transfer substrates: lane 4, 2-AP. In panels B and D, HTLV-2 IN disintegration reactions are shown for: B, lane 2, the WT U5 disintegration substrate (Wt); and modified disintegration substrates: lane 4, Hyp (4/Hyp); lane 6, 6-MeGua (4/6-MG); and lane 8, 6-MeGua substituted at positions 4 and 8 of the Y substrate (4,8/6-MG). D, lane 2, the WT U5 disintegration substrate (Wt) and modified disintegration substrates: lane 4, 2-AP. Odd lane numbers in each panel show reactions performed with substrate in the absence of protein.
Interaction of HTLV-2 IN with the Retroviral DNA Termini

**Fig. 7.** Binding affinity of HTLV-II IN for HTLV-2 U5 and modified substrates. HTLV-2 IN was incubated with the HTLV-2 U5 and modified substrates and examined for their binding affinity with a filter binding assay as described under “Experimental Procedures.” Shown is the mean and S.E. of a minimum of two assays done in duplicate of the fraction of substrate retained on the membrane. Legend: HTLV-2 U5 substrate (Wt, □), random 20-mer duplex substrate (Ran, □), major groove modification Hyp replacing the conserved G in the HTLV-2 U5 substrate (4/Hyp, □), major groove modifications 6-MeGua and 2-aminopurine located in the 4 position of the HTLV-2 U5 substrate (4/6-MG, □, and 2-AP, □), Cyt mismatch in the 4 position of the HTLV-2 U5 substrate (4C/C, □). The double mutants in the 4 position of the HTLV-2 U5 substrates (4G/C, □, and 4T/A, □).

**Table I**

| U5 substrate | 3′-Processing | 3′-Processing upstream cleavage | Strand transfer | Disintegration | Filter binding |
|--------------|---------------|-------------------------------|----------------|---------------|---------------|
| 4/Hyp        | +             | ±                             | ++             | ND            | ++            |
| 4/6-MG       | +             | +                             | ++             | ND            | ++            |
| 4/6-Thio     | ++            | +                             | ++             | ND            | +             |
| 2-AP         | +             | +                             | ND             | ND            | NA            |
| 4C/C         | +             | ±                             | ND             | ND            | NA            |
| 4G/C         | ND            | ND                            | NA             | ND            | NA            |
| 47/NA        | ND            | ND                            | NA             | ND            | +             |
| Ran          | ND            | ND                            | NA             | NA            | NA            |

a. ±, below 10% WT; +, 10–39% WT; ++, 40–79% WT; ++++, 80–99% WT; +++++, greater than 100% WT activity.

b. Upstream cleavage product level as compared to WT 3′-processing.

c. ND, not determined.

d. NA, not applicable.

Interaction. Several groups have shown that certain proteins, for example, SRY and LEF-1, can alter DNA structure through the minor groove. These proteins target certain sequences that may have an intrinsic tendency to unwind and roll one or more base steps. SRY and LEF use an intercalative wedge to pry open a single base step and distort the DNA (30). There is ample evidence that this type of DNA distortion occurs in retroviral integration (23, 31). Interestingly, the target of SRY (AACAAA) and LEF (TTCAAA) have a high similarity to the viral termini recognized by retroviral IN. Nevertheless, herein we show the importance of the minor groove in binding of the viral end by the HTLV-2 IN. However, this was not observed for the strand transfer and disintegration substrates. We propose that only WT substrates that contain a conserved C/A/G/T sequence will orient the IN complex in a way as to ensure the ends are processed correctly. The reduced binding affinity IN displayed with the mutations in the 4 position of the HTLV-2 U5 long repeat, the 4G/C and 4T/A substrates supports this hypothesis; i.e. IN does not recognize the key nucleotides at position 4, and therefore releases the substrate. Yi et al. (27) has reported that flipping of the C/A/G/T causes a decrease in binding in the presence of metal ion. We have reported herein that this reduction in binding can be traced, at least in part, to minor groove interactions at position 4. This mechanism apparently does not hold true for the random substrate, Ran, which showed only a slight reduction in binding. This in turn implies a different set of contacts for the target DNA (sequences without a CA) as proposed by recent modeling studies by Yang et al. (32).

A correct orientation in the binding step proceeds to unwinding of substrate (Fig. 7, step 2), and correct cleavage at the 3-2 phosphodiester bond, if there is a major groove O6 contact available at position 4 (Fig. 8, step 4). A transition state that could be present at step 3 would most probably reflect the melting of the three terminal base pairs as proposed by Chen et al. (31). Those distortions caused by our modified substrates may be viewed by IN as such a transition state described by step 2 in this model for the integration reaction mechanism. In effect, these modified substrates may orient the complex further upstream on the viral DNA, and thus, promote aberrant cleavage of the substrate at an upstream ACA. Interestingly, the greater the distortion at the Cyt/Gua, the greater the increase in the percentage of cleavage at the internal ACA motif. In the case of HIV-1 IN, when a 2-base pair mismatch was placed immediately downstream of an internalized CA, precise cleavage of the substrate occurred (23). Herein, we have shown that a single mismatch located at the third base pair, downstream of the CA, could promote internal cleavage. Finally, the major groove O6 contact at Gua may also allow the IN-substrate complex to change conformation, and achieve the transition state of the complex, which leads to strand transfer as shown in step 4 of our model. Conformational changes initiated by divalent metal binding have been mapped to all three domains of IN (33, 34). Although a metal binding step is not depicted in our
Interaction of HTLV-2 IN with the Retroviral DNA Termini

model, these conformational changes may allow IN to productively interact with substrate. We hypothesize that other substrate-induced conformational changes may occur immediately before or concurrent to 3' -processing as shown in step 4 of our model, although as of yet this has not been shown. These substrate induced conformational changes could also include reorganization of the IN-substrate complex as well as intramolecular changes in the tetrameric or higher ordered IN structure in preparation for the strand transfer event.

Interestingly, when Gua in the conserved and internal AC (AT/GT sequences was substituted with 6-MeGua or 4,8/6-MG, IN no longer made a precise cleavage at either of the ACA sequences. Cleavages were noted 5' to the conserved and upstream CA sequences. Imprecise cleavage of a substrate by IN has not been previously documented. However, nonspecific alcoholysis of DNA has been reported, and differs from the cleavage noted with the 4,8/6-MG substrate in that it occurs at most positions in nonviral DNA substrates (35). Based on the activity of the 4,8/6-MG substrate we hypothesize that HTLV-2 IN does not recognize the sequence of this substrate, but instead has read the substrate by its structure or rather the distortion introduced by the base modification. The conformation of the substrate may reposition IN on the substrate in such a way as to promote cleavage at the alternative sites. It is clear that conformational changes play an essential role in retroviral integrase catalysis. Additional experiments will further map the chronological and spatial locations of these structural changes as they relate to substrate binding and catalysis by IN.

Acknowledgment—We thank Dr. Monica Roth, Department of Biochemistry, University of Medicine and Dentistry of New Jersey, for critical reading of the manuscript.

REFERENCES

1. Katz, R. A., and Skalka, A. M. (1994) Annu. Rev. Biochem. 63, 133–173
2. Whitcomb, J. M., and Hughes, S. H. (1992) Annu. Rev. Cell Biol. 8, 275–306
3. Craigie, R., Fujitwara, T., and Bushman, F. (1990) Cell 62, 829–837
4. Katz, R. A., Merkel, G., Kulskowski, J., Leis, J., and Skalka, A. M. (1990) Cell 63, 87–95
5. Roth, M. J., Schwartzberg, P. L., and Goff, S. P. (1989) Cell 58, 47–54
6. Engelman, A., Mizuuchi, K., and Craigie, R. (1991) Cell 67, 1211–1221
7. Chow, S. A., Vincent, K. A., Ellison, V., and Brown, P. O. (1992) Science 255, 723–726
8. Katzman, M., Katz, R. A., Skalka, A. M., and Leis, J. (1989) J. Virol. 63, 5319–5327
9. Bushman, F. D., and Craigie, R. (1990) J. Virol. 64, 5645–5648
10. Leavitt, A. D., Rose, R. B., and Varumus, H. E. (1992) J. Virol. 66, 2539–2558
11. Marshall, T., Planellas, V., Krogstad, P., and Chen, I. S. Y. (1995) J. Virol. 69, 658–668
12. Brown, H. E. V., Chen, H., and Engelman, A. (1999) J. Virol. 73, 9011–9020
13. Balakrishnan, M., and Jonsson, C. B. (1997) J. Virol. 71, 1025–1035
14. Pemberton, I. K., Buckle, M., and Buc, H. (1996) J. Biol. Chem. 271, 1498–1506
15. van Gent, D. C., Eigerma, Y., Bolk, M. W. J., Vink, C., and Plasterk, R. H. A. (1991) Nucleic Acids Res. 19, 3821–3827
16. Khan, E., Mack, J. P. G., Katz, R. A., Kulskowski, J., and Skalka, A. M. (1991) Nucleic Acids Res. 19, 851–860
17. Wang, T., Balakrishnan, M., and Jonsson, C. (1999) Biochemistry 38, 3624–3632
18. Balakrishnan, M., Zastroz, D., and Jonsson, C. B. (1996) Virology 219, 77–86
19. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
20. Sett, T. E., and Levy, D. E. (1979) Nucleic Acids Res. 25, 3354–3361
21. Bugg, C. E., and Thewalt, U. (1970) Biochemistry 9, 7441–7445
22. Thewalt, U., and Bugg, C. (1972) J. Amer. Chem. Soc. 94, 8892–8898
23. Scoccoline, B. P., Chow, S., Ellison, V., and Brown, P. O. (1997) Genes Dev. 11, 371–382
24. Donzella, G. A., Jonsson, C. B., and Roth, M. J. (1993) J. Virol. 67, 7077–7087
25. Jonsson, C. B., and Roth, M. J. (1993) J. Virol. 67, 5583–5591
26. Chow, S., and Brown, P. O. (1994) J. Virol. 68, 3969–3977
27. Yi, J., Asante-Appiah, E., and Skalka, A. M. (1999) Biochem. 38, 8458–8468
28. Voigt, J. M., and Topal, M. D. (1990) Biochemistry 29, 1632–1637
29. Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L., and Jones, R. A. (1986) Biochemistry 25, 1027–1036
30. Werner, M. H., Bronenborn, A., and Clare, G. M. (1996) Science 271, 778–784
31. Chen, Z., Yan, Y., Minishi, S., Li, Y., Zigag-Murphy, J., Xu, B., Wintner, M., Felsak, P., Wolfe, A., Sardana, V., Emin, E. A., Hazuda, D., and Kuo, L. C. (2000) J. Mol. Biol. 296, 521–533
32. Yang, Z., Musser, T. C., Bushman, F. D., and Hyde, C. C. (1999) J. Mol. Biol. 296, 535–540
33. Asante-Appiah, E., and Skalka, A. M. (1997) J. Biol. Chem. 272, 16196–16205
34. Asante-Appiah, E., Seelholzer, S. H., and Skalka, A. M. (1998) J. Biol. Chem. 273, 35578–35587
35. Katzman, M., and Sudol, M. (1996) J. Virol. 70, 2595–2604
