Role of DNA End Distortion in Catalysis by Avian Sarcoma Virus Integrase*

Received for publication, May 21, 2001
Published, JBC Papers in Press, July 5, 2001, DOI 10.1074/jbc.M104632200

Richard A. Katz‡, Paul DiCandeloro§, George Kukolj¶, and Anna Marie Skalka
From the Fox Chase Cancer Center, Institute for Cancer Research, Philadelphia, Pennsylvania 19111

Retroviral integrase (IN) recognizes linear viral DNA ends and introduces nicks adjacent to a highly conserved CA dinucleotide usually located two base pairs from the 3′-ends of viral DNA (the “processing” reaction). In a second step, the same IN active site catalyzes the insertion of these ends into host DNA (the “joining” reaction). Both DNA sequence and DNA structure contribute to specific recognition of viral DNA ends by IN. Here we used potassium permanganate modification to show that the avian sarcoma virus IN catalytic domain is able to distort viral DNA ends in vitro. This distortion activity is consistent with both unpairing and unstacking of the three terminal base pairs, including the processing site adjacent to the conserved CA. Furthermore, the introduction of mismatch mutations that destabilize the viral DNA ends were found to stimulate the IN processing reaction as well as IN-mediated distortion. End-distortion activity was also observed with mutant or heterologous DNA substrates. However, further analyses showed that using Mn²⁺ as a cofactor, processing site specificity of these substrates was also maintained. Our results support a model whereby unpairing and unstacking of the terminal base pairs is a required step in the processing reaction. Furthermore, these results are consistent with our previous observations indicating that unpairing of target DNA promotes the joining reaction.

* This work was supported by National Institutes of Health Grants AI40385, CA71515, and CA06927 (to A. M. S.) and by an appropriation from the Commonwealth of Pennsylvania (to A. M. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
‡ To whom correspondence should be addressed: Fox Chase Cancer Center, Institute for Cancer Research, 7701 Burholme Ave., Philadelphia, PA 19111. Tel.: 215-728-3668; Fax: 215-728-2778; E-mail: R_Katz@fccc.edu.
§ Current address: DuPont Pharmaceuticals, Experimental Station, E336/1b, Wilmington, DE 19880.
¶ Current address: Boehringer Ingelheim (Canada) Ltd. Research and Development, Laval, Canada H7V 2C5.
† The abbreviations used are: IN, integrase; ASV, avian sarcoma virus; HIV-1, human immunodeficiency virus type 1; GST, glutathione S-transferase; RAG, recombination activation gene; WT, wild type.

Each DNA terminus (4, 5) (Fig. 1). Such “processing,” produces new recessed 3′-OH ends and releases a dinucleotide. In the second step, “joining,” IN catalyzes the insertion of the newly formed 3′-OH ends into target DNA (6, 7). The joining occurs through a coupled cleavage-ligation reaction whereby the 3′-OH oxygens of the viral DNA attack phosphorous atoms in both strands of the target DNA backbone, staggered by four to six base pairs (8). The product is a gapped intermediate in which only the 3′-ends of viral DNA are joined to target DNA. The gaps are repaired and the 5′-ends of viral DNA are joined to host DNA by undetermined mechanisms that depend on host repair functions (9).

In vitro, IN alone is sufficient to catalyze the processing and joining reactions. Only a metal cofactor (Mg²⁺ or Mn²⁺) and appropriate viral and target DNA substrates are required. The simplest in vitro systems employ model DNA substrates representing the terminal 15–20 base pairs of a single viral DNA end (3) (Fig. 1). The efficiency of processing is strongly dependent on the highly conserved CA dinucleotide sequence, whereas other less critical sequences reside within the first approximately seven base pairs (3).

The retroviral INs are three-domain proteins of ~300 amino acids (10, 11). The N-terminal region is a zinc-binding domain of ~50 amino acids that may function in multimerization and/or DNA recognition. The central catalytic domain is ~150 amino acids and contains a highly conserved constellation of acidic residues (the D,D(35)E motif) comprising the active site. The fold of this domain identifies IN as a member of a structural superfamily of polynucleotidyl transferases (12, 13). The carboxylate residues of the D,D(35)E motif are essential for IN activity, and structural analyses of the avian sarcoma virus (ASV) and human immunodeficiency virus type-1 (HIV-1) IN catalytic domains have confirmed their predicted role in binding the required metal ion(s), Mg²⁺ or Mn²⁺ (13). Mutagenesis experiments indicate that the D,D(35)E motif comprises a single active site for both the processing and joining reactions (14, 15). In addition, functional analyses indicate that the catalytic domain recognizes critical features of the viral DNA substrate and plays a role in target-site recognition as well (3). The C-terminal domain is ~100 amino acids and encodes a nonspecific DNA binding activity (16–18); however, the precise role of this activity in recognition of viral and target DNA is not well understood.

In addition to viral-specific sequence requirements, structural features of the viral DNA termini affect processing. If HIV-1 IN model viral DNA substrates are extended beyond the normal termini with duplex DNA, the efficiency of correct processing at the CA dinucleotide is reduced or eliminated (19, 20). Studies by Scottoline et al. (20) suggest a model in which the proximity of the processing site to the DNA terminus allows disruption of base pairing, and this is a critical step in the processing reaction. In addition, identification of cyclic dinucleo-

This paper is available on line at http://www.jbc.org

34213
otide products of the processing reaction indicates that the viral 3′-OH (as well as water) can act as a nucleophile for the processing reaction (8). Formation of this cyclic nucleotide suggests that DNA melting and distortion of the DNA substrate occurs, such that the 3′-OH is positioned to attack the phosphate at position -2 on the same strand. DNA distortion of the target DNA can also enhance the joining reaction (21), and it was proposed that the distorted viral DNA terminus and distorted target DNA could be recognized in a similar manner (20).

Here we have investigated the interaction between ASV IN and model viral DNA termini using a thymidine-selective chemical probe, potassium permanganate (KMnO₄), to detect protein-induced DNA distortions such as base unpairing and/or unstacking (22, 23). The results show that accessibility of three terminal base pairs to KMnO₄ is significantly increased in the presence of ASV IN, and this DNA-distorting activity maps to the catalytic domain. A positive correlation between the extent of IN-mediated distortion and processing activities was also observed.

**EXPERIMENTAL PROCEDURES**

**KMnO₄ Assay—**Assays were carried out under conditions similar to those described previously (24). Reactions mixtures contained 25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 100 μg/ml bovine serum albumin, and 1 μg/ml sonicated salmon sperm DNA. To 16 μl of buffer, -1 pmol (1 μl) of 32P-labeled DNA substrate (2–4 × 10⁶ dpm) and 1 μl of IN, KMnO₄ dilution, or buffer blank were added. After incubation at 37 °C for 10 min, 1 μl of 50 mM KMnO₄ was added, and incubation was continued at 37 °C for 4 min. Volume was adjusted with 78 μl binding buffer and 25 μl of stop solution (1.5 M NaOAc, pH 5.2, 1 μM β-mercaptoethanol, 100 μg/ml tRNA) was added. The nucleic acids were ethanol-precipitated, and the pellets were washed in ethanol, dried, and subjected to piperidine cleavage by resuspension in 1 M piperidine and heating at 92 °C for 30 min. The samples were lyophilized, resuspended in 10 μl of water, and lyophilized again. After a second lyophilization from 10 μl of water, the samples were resuspended in Maxam and Gilbert gel loading buffer and fractionated on 20% acrylamide-Urea gels. Processing activity was quantitated using a Fuji phosphorimager.

**IN Processing Assay—**Processing assays were carried out using conditions similar to those previously described (4). Reactions (10 μl) typically contained 1 × 10⁶ dpm 32P-labeled viral DNA substrate, 15 μM unlabeled viral DNA substrate, 2 μM IN, and 10 mM MnCl₂ (or 3 mM MgCl₂), 50 mM Hepes (pH 8.2), 2 mM β-mercaptoethanol, 50 mM NaCl, 0.1% thiodiglycol, 10 μM EDTA and 4% glycerol. The latter four components are contributed from the IN storage buffer. Samples were removed at the indicated times and subjected to fractionation on 20% polyacrylamide-Urea gels. Processing activity was quantitated using a Fuji phosphorimager.

**Proteins—**Wild type (WT) ASV IN and individual domains were purified from *Escherichia coli* as described previously (25). Proteins were dialyzed against a buffer containing 500 mM NaCl, 50 mM Hepes, pH 8.1, 1 mM EDTA, and 5% glycerol.

**Vector Construction and Purification of Catalytic Domain Mutants—**Mutations encoding the Arg-161 and Lys-164 substitutions were introduced into the expression plasmid, pET28b (Novagen). The bacterially expressed proteins—

**RESULTS**

**Detection of ASV IN-mediated DNA Distortion at Viral DNA Terminals**—Several studies have shown that IN recognizes specific viral sequences efficiently only when they are present at DNA termini (19, 20). To investigate potential DNA structure-specific interactions between IN and viral DNA substrates we used model ASV DNA substrates comprising a 19-mer duplex that corresponds to the U3 end of ASV DNA. The experimental design is summarized in Fig. 1. The strand that is normally cleaved in the processing reaction to release the two Ts was 5′-end-labeled with 32P. KMnO₄, a reagent frequently used to detect regions of DNA that are unpaired or distorted as a result of protein binding (22, 23), was added to the reactions to detect potential IN-mediated distortion of viral DNA. KMnO₄ is an oxidizing agent that attacks the 5,6 double bonds of thymine. In B-DNA, this bond is shielded by base stacking interaction and, thus, T residues in such DNA duplexes are relatively resistant to oxidation.

The KMnO₄ modification reaction can be affected by a variety of buffer conditions; for example, glycerol and reducing agents can quench the reaction (22). To control for such variation, buffer blanks (0 μM IN) were included in all experiments. In addition, all proteins were dialyzed against the same batch of buffer, which was also used as a blank. The sensitivity of the reaction can be adjusted by varying KMnO₄ concentration, reaction times, and conditions that affect DNA breathing (i.e. NaCl concentration). IN requires a divalent metal for activity (Mg²⁺ or Mn²⁺) and, thus, divalent metal was excluded from these reactions to prevent IN-mediated cleavage of substrate DNA.

After the mixtures of DNA substrates and ASV IN were exposed to KMnO₄, they were treated with piperidine and fractionated on polyacrylamide-Urea gels. As shown in Fig. 2A (lane 3), in the absence of IN, the two terminal Ts (positions 1 and 2) are partially accessible to KMnO₄ modification, as might be expected due to breathing at the DNA ends (29). However, in the presence of increasing concentrations of IN, these two thy-
midines become significantly more accessible (lanes 4–6). The IN concentration at which distortion is detected is within the range used in the standard IN processing assay. As expected, cleavage requires treatment with both KMnO₄ and piperidine, demonstrating that the cleavage is chemically mediated rather than IN endonuclease-mediated. At the concentration of KMnO₄ used in these experiments, the basal level for IN-independent modification was low (Fig. 2A, lane 3). Under these conditions, some substrate remained unreacted in the presence of IN (Fig. 2A, lane 6), allowing quantitation of the IN-dependent affects by phosphorimaging (data not shown). Sensitivity of the thymidine at position 3 from the tip on the complementary strand was also examined. As shown in Fig. 2B, ASV IN-mediated distortion could also be detected at this position. In this case, very little background breathing was observed.

A basic protein, equine cytochrome c (pI = 9.58) was also included as control for ASV IN (pI = 9.94); no significant activity was detected with the control protein, even at a 3-fold higher concentration (Fig. 2C).

**Substrate Requirements for IN-mediated DNA Distortion Activity**—To examine the DNA sequence requirements for this distortion activity, several mutant substrates were assayed (Fig. 3A; mutant substrates are shown in Fig. 1). The highly conserved CA at positions 3 and 4 from the terminus (and corresponding complementary base pairs) was changed to GA (lanes 8–10) or GT (lanes 11–15). As expected from previous studies (3), a significant reduction (~10-fold) in processing activity (~2 cleavage) was observed with these substrates using Mg²⁺ as a cofactor (Fig. 3A). However the IN-mediated DNA distortion activity was similar to wild type with both mutant substrates (Fig. 3B). With the CA to GT change, the thymidine introduced at position 3 from the end became accessible to KMnO₄ in the presence of IN (lane 12). This is consistent with the reactivity of the thymidine at the third position in the complementary strand of the wild type substrate (Fig. 2B).

**Fig. 2. Effect of ASV IN on KMnO₄ sensitivity of viral DNA substrates.** ³²P-end-labeled substrates were incubated with ASV IN or control protein. After treatment with KMnO₄ and piperidine cleavage, the products were fractionated on 20% acrylamide-UREA gels. IN concentration is shown below each lane. A, viral DNA substrate is shown at top. The substrate used in this experiment was identical to the WT shown in Fig. 1 except that it lacked a C/G base pair at the labeled end. Asterisk indicates 32P-5' end label. The highly conserved CA is indicated in bold. Filled circles on substrate mark T residues that are sensitive to KMnO₄. Adjacent to the autoradiogram, S indicates migration of labeled substrate strand and filled circles denote cleavage at terminal T residues. Presence or absence of piperidine (PIP) and potassium permanganate (PER) are indicated by + and − below. A, asterisk indicates 32P-3' end labeled substrate (WT + 3; also see Fig. 1). Closed circle on the substrate marks the T residue that is sensitive to KMnO₄, C, KMnO₄ sensitivity of T residues using cytochrome c (CytC) as a control. WT + viral DNA substrate is diagrammed in abbreviated form.

**Fig. 3. Assay of IN-induced KMnO₄ sensitivity and IN activity on WT and mutant substrates.** Symbols and annotations as in Fig. 1. A, IN processing activity on 5'-labeled WT and mutant substrates. Substrate ends are shown above each group of samples in abbreviated form as in Fig. 2C. Arrow indicates cleavage at correct processing site, yielding a “2” product. Reactions in lanes 1–15 included Mg²⁺ as a co-factor. Substrates were: lanes 1–5, WT + lanes 6–10, CA-GA; lanes 11–15, CA-GT (see Fig. 1). Reactions in the panel on right included Mn²⁺ as a cofactor. In right panel, substrates are from left to right: WT, CA-GA, CA-GC. A reproducible mobility difference was noted between substrates, which correlates with sequence differences at the termini. The secondary −3 cleavage site is observed using Mn²⁺ as a cofactor (see “Results”). B, KMnO₄ assay using mutant substrates from panel A. Mutations are boxed; other symbols are as in Fig. 2. C, KMnO₄ activity on HIV-1 substrates (HIV-U5–21, HIV-U5–21, see Fig. 1).
These results suggested that the distortion activity can be uncoupled from the processing activity. However, in the presence of Mn$^{2+}$ as a cofactor, and with the incubation time extended to 60 min, -2 processing activity could be observed with the mutant substrates (Fig. 3A, right panel). In general, the endonuclease activity of ASV IN increases with Mn$^{2+}$ while the specificity decreases. However, we note that the preferred IN-cleavage sites in the mutant substrates remain at the -2 position, suggesting that a feature other than the conserved CA dinucleotide can contribute to selection of the correct cleavage site.

To investigate DNA sequence requirements for distortion activity further, we tested the activity of the ASV IN protein with heterologous, HIV-1 DNA substrates that differ in sequence from the ASV DNA substrate at most positions (substrates shown in Fig. 1). The U3- and U5-derived HIV-1 DNA substrates were mutated to introduce additional thymidine residues that could be monitored for KMnO$_4$ modification (the G/C base pair at position 2 was changed to T/A). Thus, on all substrates tested, we observed a relationship between the ability of ASV IN to distort ends and select the correct processing site.

**Analysis of DNA Substrates with Destabilized Termini**—The results described above suggested that although the sequence requirement for DNA end distortion by ASV IN is quite relaxed; such distortion may be important for processing activity. As KMnO$_4$ is believed to detect base unpairing or unstacking, it seemed possible that such distortion may be required for processing as proposed previously for HIV-1 IN (20). If so, base mismatching at the termini should stimulate processing by ASV IN. To test this prediction, a set of substrates was designed in which the terminal two base pairs were mismatched. In these test substrates, the two terminal Ts (on the strand with the processing site) were maintained, and non-complementary bases (CC, GG, TT) were introduced on the opposite strand. In this way the KMnO$_4$ sensitivity of the two terminal Ts could be monitored as with the WT substrate.

We first monitored KMnO$_4$ sensitivity of the mismatched substrates (Fig. 4A). The percentage of cleaved products corresponding to the terminal two Ts was quantified by phosphorimaging (data not shown). To establish the range of KMnO$_4$ sensitivity for the terminal Ts, we compared the WT + substrate to one that contained unpaired Ts at positions 1 and 2 (ΔΔA, Fig. 1). Surprisingly, in the absence of ASV IN, the KMnO$_4$ reactivity of the terminal Ts in this substrate differed by only ~2-fold from the substrate containing paired terminal nucleotides (Fig. 4A, compare lanes 1 and 5). We interpret these results to mean either that there is significant breathing in the WT substrate under these conditions or that the unpaired bases in the ΔΔA substrate maintain stacking interaction and thus are relatively resistant to KMnO$_4$ modification. As observed previously, addition of increasing amounts of IN resulted in increased KMnO$_4$ sensitivity with the WT substrate (lanes 1–4). Furthermore, addition of IN also increased the sensitivity of the terminal Ts in the substrate with unpaired Ts (lanes 5–8). This suggests that IN is able to distort the DNA terminus in a manner that is distinct from, but may include, base unpairing. As expected, an increase in KMnO$_4$ sensitivity of the terminal Ts (~3-fold) was observed in all the mismatched substrates (lanes 9–20) compared with the WT substrate (lanes 1–4).

We next tested the ability of ASV IN to process the mismatched substrates. As shown in Fig. 4B, the rate and extent of processing of all three mismatched substrates was significantly higher than the WT substrate. One interpretation of this positive correlation between end distortion and enzymatic activity is that mismatching at the termini lowers an energetic barrier for processing (20).

**Analysis of Substrates with Stabilized Termini**—If unpairing of ASV termini is required for efficient processing, then stabilization of base pairing at the substrate tips might be expected to inhibit this activity. As a corollary to the experiments with mismatched substrates, we examined substrates in which the terminus was stabilized by replacing the two T/A base pairs with more stable G/C base pairs (see Fig. 1 for substrates). Replacement of the two terminal T residues with G residues removes the KMnO$_4$-sensitive T bases, and therefore position 3 was monitored for KMnO$_4$ modification on the opposite strand (as in Fig. 2B). As shown in Fig. 5A (lanes 1 and 2) the T at position 3 in the WT substrate shows increased susceptibility to KMnO$_4$ modification in the presence of IN as in Fig. 2B. Modification of position 3 in the substrate with the GG/CC ends (lane 3) is significantly lower than the WT, presumably due to reduced breathing of these ends. KMnO$_4$ accessibility in the presence of IN was also significantly reduced as compared with the WT substrate (compare lanes 1 and 4). As shown in Fig. 5B, the rate and extent of processing of the GG/CC stabilized substrate (lanes 6–10), was also reduced as compared with the WT substrate (lanes 1–5, and graph). To distinguish between sequence and structural effects of the GG/CC mutation, we introduced a change at position 3 in which the highly conserved A was substituted with a T creating a TT mismatch. This mismatch is predicted to destabilize the terminus structure and promote CA-independent processing as described previously for HIV-1 DNA substrates (20). As shown in Fig. 5B (lanes 11–15, and graph), this mismatch indeed restored near WT processing activity. KMnO$_4$ sensitivity was also significantly enhanced at the mismatch site in the presence of IN.
(Fig. 5A, compare lanes 4 and 6). These results also show that the extent of IN distortion at the terminus and the efficiency of processing are correlated. In addition, mismatching at position 3 can overcome the deleterious effect of mutating the conserved CA dinucleotide.

Effects of Extending the DNA Terminus on ASV IN-mediated DNA Processing and Distortion—Previous studies have shown that extending the DNA terminus severely inhibited processing by HIV-1 IN (19, 20). One possible interpretation is that the terminal extension prevents unpairing or distortion at the processing site. To test this hypothesis, the ASV U3 terminus was extended by two additional base pairs (TT/AA). KMnO₄ analysis revealed significant DNA breathing of the terminal two base pairs in the absence of IN (Fig. 6A, lane 1). In the presence of IN, sensitivity of the two terminal Ts was enhanced and extended to the T adjacent to the processing site (lane 2). A control substrate in which the four T stretch was entirely unpaired showed equal sensitivity of all unpaired Ts in the absence of IN (lane 3), but greatly enhanced KMnO₄ sensitivity was observed at the T adjacent to the CA in the presence of IN (lane 4). Correct IN processing (−4 position) of the extended substrate was reduced as compared with the control substrate containing the unpaired T extension (Fig. 6B, compare lanes 1–5 with lanes 6–10). However, initial cleavage (lanes 2 and 3) occurred at an aberrant site in the extended substrate, which was two base pairs from the terminus (−2), followed by cleavage at the correct site (−4) at later times (lanes 4 and 5). As noted above, KMnO₄ modification of the extended duplex substrate is most efficient at the two terminal positions in the presence of IN (Fig. 6A, lane 2). Under these conditions, it is possible that initial cleavage is dictated by IN-mediated distortion at positions 1 and 2. The initial cleavage may expose a new 3′-end, possibly allowing further distortion at the correct −4 site. This experiment provides further evidence that an important feature for processing site selection is the distance from the terminus and that processing depends on the ability of the DNA ends to become distorted or unpaired.

Mapping of DNA Distortion Activity—We next asked which domain or domains of ASV IN were responsible for the DNA distortion activity. Four truncated proteins were assayed: IN 1–207, 39–286, 52–207, and 208–286 (Fig. 7A). As shown in Fig. 7B, distortion activity could be detected with the two proteins that retained the catalytic domain (1–207 and 39–286). The isolated catalytic domain (52–207) did not display distortion activity when assayed at the same concentrations, but activity could be detected at higher concentrations (Fig. 7C). The C-terminal domain, 208–286, assayed as a GST fusion protein (30), did not display detectable activity even at a higher concentration (7B, lane 18). We conclude that the catalytic domain contains the distorting activity, but both the N- and C-terminal domains can contribute to the activity either by stabilizing the catalytic domain or by participating in DNA binding.

The isolated catalytic domain of ASV IN (52–207) lacks de-
residues in ASV IN (Arg-161 and Lys-164) are candidates for contacting the viral DNA substrate within the region showing IN-mediated distortion (positions 1, 2, 3). To test this hypothesis we introduced alanine and glutamic acid substitutions into the analogous positions, i.e. (R161A, R161E, K164A, K164E) in the isolated ASV IN catalytic domain (positions 52–207), and determined the effects on activities of these proteins. The R161E/K164E double substitution is predicted to have the most severe effect and was studied further, but other substitutions were also assayed in a similar manner (data not shown).

The isolated ASV catalytic domain displays nonspecific DNA binding, dissociation (25, 31) and the aforementioned −3 specific endonuclease activity, both of which require an intact active site (31). The −3 endonuclease activity is a minor activity of the full-length IN and is most prominent when Mn2+ is used as a metal cofactor. The core domain that included the R161E/K164E double substitution showed a severe reduction (−10-fold) in −3 endonuclease activity (Fig. 8A, compare lanes 1–5 and 6–10), which is consistent with the previously described effect on HIV-1 IN activity (33). However, although compromised for endonuclease activity, the R161E/K164E mutant retained robust DNA distortion activity (compare Fig. 8B with Fig. 7C). These results indicate that Arg-161 and Lys-164 contribute to catalysis but not to DNA distortion activity.

The ASV IN catalytic domain displays robust nonspecific DNA binding activity in a standard filter-binding assay. As shown in Fig. 8C, the R161E/K164E mutant retained WT DNA binding activity. Thus, it is possible that the activities measured by the filter binding and KMnO4 sensitivity are related. As also shown in Fig. 8C, the isolated HIV-1 catalytic domain does not bind to DNA in this assay, which is consistent with previous studies that mapped the major HIV-1 IN DNA binding determinants to the C terminus (17, 18).

**DISCUSSION**

Retroviral IN specifically recognizes and processes viral DNA termini. This recognition is dependent on both viral DNA sequences and the presence of a DNA terminus. Here we demonstrate that ASV IN can distort DNA termini as indicated by enhanced accessibility to KMnO4. The distortion activity was detected within the terminal three base pairs and included the processing site adjacent to the conserved CA. ASV IN is also able to distort heterologous and mutant DNA substrates, and, in the presence of Mn2+, the ends of such substrates are nicked by IN preferentially at the normal processing site (Fig. 3A). Thus, distortion and cleavage appear to be coupled with respect to site selection. Furthermore, substrates containing mismatches in the first three positions showed enhanced distortion as well as enhanced processing activity (Fig. 4). This result suggests that mispairing lowers an energetic barrier for both distortion and enzymatic activity as previously observed for HIV-1 IN (20). Conversely, ASV IN catalytic activity is not required for distortion as indicated by results obtained with enzymatically deficient or inactive proteins (R161E/K164E and D64N) (Fig. 8; results not shown for the D64N catalytic domain). These results are consistent with the fact that the metal co-factor (Mg2+ or Mn2+) was not required for distortion activity.

KMnO4 detects base unpairing or unstacking of T residues. Under the conditions used, we detected basal breathing of the terminal T residue and minimal breathing of Ts at positions 2 and 3 (Fig. 2). Previous NMR-based studies indicate that significant breathing at DNA ends is limited to two or three terminal base pairs (29). As expected, substrate mispairing within the first three positions enhanced KMnO4 sensitivity (Figs. 4 and 5). However, we observed that ASV IN could distort paired as well as mispaired and unpaired terminal T...
residues as determined by KMnO4 modification (Figs. 4 and 6). Sensitivity to KMnO4 is probably dependent on unstacking of bases on both sides of the target T residue (34). Thus, the KMnO4 sensitivity detected in the presence of ASV IN, may represent an activity which exposes both faces of the T base. NMR studies of DNA duplexes have detected base stacking interactions between a 5′-TT extension and the adjacent core duplex (35). It is possible that the T extensions in the ASV DNA duplex substrates (ΔAA, 2T/ΔAA-EXT) are similarly structured under the reaction conditions used here. Taken together, our results suggest that DNA end breathing enhances the ability of ASV IN to bind and distort the termini such that the terminal bases are highly exposed to KMnO4 modification and are in optimal position for catalysis.

The sequence-independent DNA binding activity of retroviral integrases is believed to reflect their ability to recognize target DNA. Several studies have revealed specific interactions between IN and viral DNA ends (3), but the mechanisms by which retroviral IN proteins discriminate these ends and host DNA are poorly understood. IN functions as a multimer, and DNA binding surfaces may span subunits within or between protomers. Recently, a model for ASV IN docking to a viral DNA end was proposed based on the crystal structure of a two-domain protein (the catalytic domain plus the C-terminal domain) (36). In this model, the viral DNA end contacts basic residues in both the catalytic and C-terminal domains with the CA end placed at the active site, and this is consistent with cross-linking data obtained with HIV-1 IN (37).

Here we report that the ASV IN distortion activity was mapped to the catalytic domain (Fig. 7), which is consistent with the coupling observed between distortion and cleavage site selection. However, substitution of two basic residues thought to be involved in orienting the CA in the active site (R161E/K164E) reduced enzymatic activity, but had no detectable effect on DNA binding or distortion activities (Fig. 8). Thus, other residues in the core domain must serve to promote unpairing or stabilize the unpaired ends. The observation that the R161E/K164E catalytic domain maintains DNA binding suggests that this domain retains extensive contacts with DNA (Fig. 8). Additional substitutions may allow identification of these contact residues.

Experimental evidence strongly suggests that processing by HIV-1 IN is facilitated by disruption of terminal base pairs (20), but we could not detect the predicted distortion activity with this protein under a variety of conditions with KMnO4 as a probe (data not shown). We note that ASV IN is significantly more enzymatically active than HIV-1 IN prepared in similar heterologous expression systems. It is unclear if this difference is relevant to virus biology, but the more robust activity of ASV IN may increase the ability to detect distortion activity. We also note that the predicted pI of the ASV IN catalytic domain (52–207, pI 10.4) is considerably higher than that of HIV-1 IN (50–212, pI 7.4), perhaps reflecting a less extensive basic surface for DNA binding. This may account for the inability to detect DNA binding with the isolated HIV-1 IN catalytic domain (Fig. 8). The major DNA binding determinants of HIV-1 IN map to the C-terminal domain and DNA contacts with the catalytic domain are only detected by protein-DNA cross-linking (3). In contrast, the ASV IN catalytic domain shows strong DNA binding (Fig. 8). Thus, the isolated HIV-1 IN catalytic domain displays neither DNA binding nor distortion activity, whereas the ASV IN catalytic domain has both activities.

Our previous studies (38) with ASV IN and HIV-1 IN suggested that unpairing of bases in target DNA also promotes the joining step of the integration reaction; an extruded cruciform DNA structure in a covalently closed supercoiled plasmid target was highly preferred integration site. As the integration sites within the cruciform mapped to the junctions between the stems and loops, we proposed that the extruded stem-loop allows DNA unpairing that may overcome an energetic barrier for the joining step of the integration reaction. The results presented here indicate that the DNA termini bound by the catalytic domain are distorted, which is consistent with the proposed role of the catalytic domain in target site selection (3).

Thus, in addition to using the same catalytic residues for viral DNA end processing and target DNA joining, the recognition of predistorted target DNA, or distortable viral DNA ends, may involve the same catalytic domain contacts. Previous studies with HIV-1 IN also suggested a common recognition mechanism for ends and distorted target DNA (20).

Retroviral IN belongs to a structural and mechanistically related superfamily of recombinases that catalyze phosphoryl

FIG. 8. Activity assay, KMnO4 assay, and DNA binding assays of IN catalytic domain containing substitutions in putative DNA binding residues. A, endonuclease assay for WT (lanes 1–5) and the R161E/K164E substituted (lanes 6–10) ASV IN catalytic domain. Symbols are as in Fig. 3. The catalytic domain cleaves between the C and A (−3) using Mn2+ as a co-factor (see Fig. 3A). B, KMnO4 assay of R161E/K164E ASV IN catalytic domain. Symbols are as in Fig. 3. C, nonspecific DNA-protein filter binding assay using 32P-end-labeled DNA and WT and mutant catalytic domains. Retention of DNA on the filter reflects protein binding (see under "Experimental Procedures").
DNA End Distortion by ASV Integrase

transferase reactions (12, 13). Two other members are the bacteriophage Mu transposase and the RAG-1/2 recombinase that catalyzes the initiation of V(D)J recombination. In both these systems, DNA distortion or unpairing may also be a requisite step. Mu DNA ends are cleaved from adjacent host sequences during transposition and distortion or unpairing of flanking DNA occurs during metal-dependent assembly of the Mu transpososome (39). For V(D)J recombination, cleavage occurs between recombination signal sequences and coding sequences (40). This first cleavage reaction, mediated by RAG-1/2, produces a new 3'-hydroxyl at the end of the coding sequence. RAG-1/2 then mediates an attack by this hydroxyl on the coding-recombination signal sequences junction on the opposite strand, forming a hairpin end. Several lines of evidence indicate that RAG-mediated DNA distortion and unpairing occurs at the coding-recombination signal sequences junction (41–43). Thus, DNA distortion activity may be an additional shared feature of these superfamily members.

In summary, we have shown that ASV IN has a pre-processing activity that distorts linear DNA ends. Though this activity is distinctive, it appears to be required for the subsequent endonucleolytic processing reaction. DNA ends with mispair mutations may serve as mimics of this pre-processed stage and may be regarded as “transition-state” analogues susceptible to increased IN processing activity. These results identify an important function of IN that can now be studied independently of the catalytic activity. It is possible that this distortion function can be targeted for inhibition.

Acknowledgments—We thank Kim Boland for excellent technical assistance and Marie Estes for preparing the manuscript. We are also grateful to George Merkel and Mark Andrade for protein reagents. We also thank Mark Andrade for helpful discussions and Ken Zaret, Yoshi Matsumoto, and Tony Yeung for critical review of the manuscript.

REFERENCES

1. Brown, P. O. (1998) in Retroviruses (Coffin, J. M., Hughes, S. H., and Varmus, E., eds), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
2. Katz, R. A., and Skalka, A. M. (1994) Annu. Rev. Biochem. 63, 133–173
3. Katzman, M., and Katz, R. A. (1999) Adv. Virus Res. 52, 371–395
4. Katzman, M., Katz, R. A., and Skalka, A. M., and Leis, J. (1989) J. Virol. 63, 5319–5327
5. Roth, M. J., Schwartzberg, P. L., and Goff, S. P. (1988) Cell 58, 47–54
6. Craigie, R., Fujiiwa, T., and Bushman, F. (1990) Cell 62, 829–837
7. Katz, R. A., Merkel, G., Kulkosky, J., Leis, J., and Skalka, A. M. (1990) Cell 63, 87–95
8. Engelman, A., Mizuuchi, K., and Craigie, R. (1991) Cell 67, 1211–1221
9. Daniel, R., Katz, R. A., and Skalka, A. M. (1999) Science 284, 644–647
10. Andrade, M. D., and Skalka, A. M. (1996) J. Biol. Chem. 271, 19633–19636
11. Esposito, D., and Craigie, R. (1999) Adv. Virus Res. 52, 319–333
12. Yang, W., and Steitz, T. A. (1995) Structure 3, 131–134
13. Wlodawer, A. (1999) Adv. Virus Res. 52, 335–350
14. Kulkosky, J., Jones, K. S., Katz, R. A., Mack, J. P., and Skalka, A. M. (1992) Mol. Cell. Biol. 12, 2331–2338
15. Engelman, A., and Craigie, R. (1993) J. Virol. 66, 6361–6369
16. Khan, E., Mack, J. P., Katz, R. A., Kulkosky, J., and Skalka, A. M. (1991) Nucleic Acids Res. 19, 851–860
17. Vink, C., Oude Groenening, A. M., and Plasterk, R. H. (1993) Nucleic Acids Res. 21, 1419–1425
18. Woerner, A. M., and Marcus-Sekura, C. J. (1993) Nucleic Acids Res. 21, 3507–3511
19. Vink, C., van Gent, D. C., Elgersma, Y., and Plasterk, R. H. A. (1991) J. Virol. 65, 4636–4644
20. Scottoline, B. P., Chow, S., Ellison, V., and Brown, P. O. (1997) Genes Dev. 11, 371–382
21. Müller, H. P., and Varmus, H. E. (1994) EMBO J. 13, 4704–4714
22. Holstege, F. C., and Timmers, H. T. (1997) Methods 12, 203–211
23. Niesen, P. E. (1990) J. Mol. Recognit. 3, 1–25
24. Kukolj, G., and DuBow, M. S. (1992) J. Biol. Chem. 267, 17827–17835
25. Kulkosky, J., Katz, R. A., Merkel, G., and Skalka, A. M. (1995) Virology 206, 448–456
26. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds) (1999) Short Protocols in Molecular Biology, 4th Ed., John Wiley & Sons, Inc., New York
27. Asante-Appiah, E., and Skalka, A. M. (1997) J. Biol. Chem. 272, 16196–16205
28. Müller, B., Bizub-Bender, D., Andrade, M. D., Jones, K. S., and Skalka, A. M. (1995) J. Virol. 69, 5631–5639
29. Leroy, J. L., Kochoyan, M., Huynh-Dinh, T., and Guerou, M. (1988) J. Mol. Biol. 200, 223–238
30. Andrade, M. D., and Skalka, A. M. (1995) J. Biol. Chem. 270, 28299–28306
31. Bujaac, G., Jakobiiski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R. A., and Skalka, A. M. (1995) J. Biol. Chem. 253, 333–346
32. Lukowski, J., Yang, F., Alexandratos, J., Merkel, G., Katz, R. A., Gravuer, K., Skalka, A. M., and Wlodawer, A. (1995) J. Biol. Chem. 270, 32685–32699
33. Jenkins, T. M., Esposito, D., Engelman, A., and Craigie, R. (1997) EMBO J. 16, 6849–6859
34. Declais, A.-C., and Lilley, D. M. J. (2000) J. Mol. Biol. 299, 421–433
35. Senior, M., Jones, R. A., and Breslauer, K. J. (1988) Biochemistry 27, 3879–3885
36. Yang, Z. N., Mueser, T. C., Bushman, F. D., and Hyde, C. C. (2000) J. Mol. Biol. 296, 535–548
37. Esposito, D., and Craigie, R. (1998) EMBO J. 17, 5823–5834
38. Katz, R. A., Gravuer, K., and Skalka, A. M. (1998) J. Biol. Chem. 273, 24190–24196
39. Saviliash, H., Rice, P. A., and Mizuuchi, K. (1995) EMBO J. 14, 4983–4988
40. Fugmann, S. D., Lee, A. I., Shoket, P. E., Vinley, I. J., and Schatz, D. G. (2000) Annu. Rev. Immunol. 18, 495–527
41. Cuomo, C. A., Mundy, C. L., and Ottinger, M. A. (1996) Mol. Cell. Biol. 16, 6583–6590
42. Akamatsu, Y., and Oettinger, M. A. (1998) Mol. Cell. Biol. 18, 4670–4678
43. Kale, S. B., Landree, M. A., and Roth, D. B. (2001) Mol. Cell. Biol. 21, 459–466
Role of DNA End Distortion in Catalysis by Avian Sarcoma Virus Integrase
Richard A. Katz, Paul DiCandeloro, George Kukolj and Anna Marie Skalka

J. Biol. Chem. 2001, 276:34213-34220.
doi: 10.1074/jbc.M104632200 originally published online July 5, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M104632200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 41 references, 18 of which can be accessed free at
http://www.jbc.org/content/276/36/34213.full.html#ref-list-1