In this investigation, we examine the interaction between the human immunodeficiency virus type I integrase and oligonucleotides that reflect the sequences of the extreme termini of the viral long terminal repeats (LTRs). The results of gel filtration and a detailed binding density analysis indicate that the integrase binds to the LTR as a high-order oligomer at a density equivalent to 10 ± 0.8 integrase monomers per 21-base pair LTR. The corresponding binding isotherm displays a Hill coefficient of 2, suggesting that the binding mechanism involves the cooperative interaction between two oligomers. This interaction is quite stable, exhibiting a prolonged half-life (t1/2 ~ 13 h) in the presence of Mn2⁺ cations. Complexes were less stable when formed with Mg2⁺ (t1/2 ~ 1 h). The role of Mn2⁺ appears to be in the induction of the protein-protein interactions that stabilize the bound complexes. In terms of the 3'-end processing of the LTR, similar catalytic rates (kcat ~ 0.06 min⁻¹) were obtained for the stable complex in the presence of either cation. Hence, the apparent preference observed for Mn2⁺ in standard in vitro integration assays can be attributed entirely to the augmentation in the DNA binding affinity of the integrase.

The integration of retroviral DNA involves two separable DNA strand cleaving events at positions that consequently form the sites of union between the viral and host DNA. The first occurs in the cytoplasm of the infected cell and results in the removal of the two outermost bases from each 3'-end of the pre-integrated linear DNA genome (1-3). This cleavage activity occurs invariantly adjacent to a conserved CA dinucleotide, embedded within a short inverted repeat sequence present at the termini of the viral LTRs. The second strand breakage involves the coordination of the recessed viral 3'-OH as a nucleophile in an enzyme-catalyzed attack on the phosphodiester backbone of the host target DNA. Concurrent with the exposure of a host 5'-phosphate, a single, direct transesterification with the attacking viral 3'-OH end results in the insertion of the viral DNA at this site (4).

Whereas the integration reaction is mediated in vivo by a large nucleoprotein complex known to contain several retroviral proteins (5), both the 3'-end processing and strand transfer activities can be carried out in vitro using short duplex oligonucleotides that mimic one of the viral LTR termini and the purified retroviral integrase protein alone (6-9). Hydrolysis of the viral substrates by the integrase is site-specific in the sense that it consistently cleaves 3' of the conserved CA dinucleotide, even when the 3'-LTR is artificially extended as single-stranded DNA (9). However, although several studies have confirmed that a level of sequence specificity is required for efficient cleavage and integration (9-12), this is not reflected in a corresponding DNA binding preference for the cognate LTR (6, 11, 13-17).

Several pertinent observations have been made with regard to the dominant nonspecific DNA binding properties of retroviral integrases. Deletion mapping studies have located the major determinant of the nonspecific DNA binding domain to a region within the C terminus of the integrase (18, 19). Although poorly conserved and of variable length in sequence comparisons between retroviral integrases, this region is rich in basic residues (Arg, Lys) and may interact nonspecifically with the DNA via general electrostatic interactions with the phosphate backbone. A central catalytic core domain of the integrase, which possesses a highly conserved D,D-35-E motif (15, 20), proposed to form the polynucleotidyl transfer catalytic site (21-23), also binds DNA but with an apparent requirement for branched structures that mimic the Y-type intermediate resulting from a single LTR integration event into a target DNA (17). The DNA binding properties of the core domain thus reflect the ability of this region to recognize and resolve such DNA structures by catalyzing the removal of the branch arm (LTR) and resealing the broken (target) strand in a process termed disintegration (24). It has been proposed that a second highly conserved region, a HHCC motif located near the N terminus of the integrase, may form a Zn²⁺-binding finger with similarities to those of the transcription factor IIIA-like DNA binding proteins (25). However, whereas this region has been shown to bind Zn²⁺ in vitro (22), no evidence has been presented to define a direct interaction with the substrate DNA in binding studies. Rather, this region has been implicated in forming protein-protein interactions required for the divalent cation-induced oligomerization of the integrase (26). Interference with this motif, either by mutagenesis (16, 21, 23, 27, 28) or chemical modification (28), is detrimental to the 3'-end processing and strand transfer activities of the integrase, but such modifications do not appear to block disintegration or the nonspecific interaction with DNA. The observation that the modification of this domain prevents the formation of the stable interaction between the integrase and the LTR suggests that protein-protein interactions may be essential in the assembly of the functional complexes (26, 29). The involvement of subunit interactions in the assembly of an active oligomeric form of the HIV-1 integrase has also been demonstrated by genetic

* This work was supported by Grant G9229152 from the Medical Research Council (UK) to I. K. P. and J. S. Oxford at London Hospital Medical College, UK. 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. Tel.: 331-40613002; Fax: 331-40613060.

* The abbreviations used are: HIV-1, human immunodeficiency virus type I; VD, variable domain; LTR, long terminal repeat; C-terminal, C terminus; DNA, deoxyribonucleic acid.
(30) and in vitro complementation (31–33) studies of integrase mutants.

The refractory nature of this retroviral protein to analysis by conventional biochemical strategies, such as mobility shift and DNase I footprinting techniques, has previously precluded a detailed, quantitative account of the biophysical properties of integrase nucleoprotein complexes. Recently, highly constrained solution conditions allowing a specific mobility shift to be detected between the HIV-1 integrase and the LTR were defined (34). Notably, these required low concentrations of the DNA substrate and the omission of the divalent cation essential for the catalytic processing of the LTR. In the present investigation, we describe the application of laser-mediated single-pulse UV cross-linking and nuclease digestion techniques to obtain a quantitative description of the stable binding of the HIV-1 integrase to the LTR, under solution conditions more closely reflecting those optimized for 3'-end processing and DNA integration assays. The implications of this data on the quaternary structure of the nucleoprotein complex and the influence of divergent cations on the stability of this interaction are examined.

EXPERIMENTAL PROCEDURES

Materials

Oligonucleotides used in this study were synthesized and purified by Genset (Paris). Snake venom phosphodiesterase I was purchased from Boehringer Mannheim (France). All chemicals used were analytical grade. MnCl₂ and MgCl₂ were obtained from Probalo (Paris).

Methods

Expression and Purification of HIV-1 Integrase—The HIV-1 integrase encoded by the LAV-BRU isolate (35) was cloned and produced in Escherichia coli strain BL21(DE3) by induction with 0.5 mM isopropyl-1-thio-

β-D-galactoside (IPTG). The integrase was solubilized by resuspending the insoluble fraction of sonicated lysates in buffer A (40 mM Tris-HCl (pH 7.6), 1 mM NaCl, 10 mM imidazole, 10 mM CHAPS) and stirring for several hours at 4 °C. The solubilized protein was loaded onto a Ni²⁺-chelating column and further purified from the factor Xa enzyme by gel filtration on a Sephadex G-25 column (Pharmacia). Standard 3'-processing reaction conditions were 30 mM Tris-HCl (pH 7.6), 5 mM MnCl₂, 1 mM diethiothreitol, 0.05% Tween 20 with 50 mM NaCl, 1% glycerol, and 0.5 mM CHAPS contributed by the enzyme storage buffer in a 20-μl final volume. Incubations (pre-warmed) were conducted at 37 °C. The products were analyzed on 20% denaturing polyacrylamide gels.

Densityometry—Quantification was performed by phosphorimage densitometry (Molecular Dynamics).

Endonuclease Protection Assay—DNA binding reactions were conducted under conditions stated above for 3'-end processing reactions with a 5-min pre-binding step, nuclease digestions were performed by the addition of snake venom phosphodiesterase I to a final concentration of 100 μg/ml. This is equivalent to 0.8 units of phosphodiesterase I, where 1 unit is defined as the hydrolytic activity yielding 1 μmol of p-nitrophenyl thymidine 5'-phosphate min⁻¹ at 25 °C. Reactions were continued for specific time intervals and terminated by the addition of a 95% formamide/20 mM EDTA solution to a 1:4 dilution and heating to 80 °C. Products were analyzed on 20% denaturing polyacrylamide gels.

Cooperative Binding of HIV-1 Integrase to DNA

We suspect that in the case of the core domain of the integrase, the presence of a His tag (or other peptide sequence) may play a critical role in maintaining the native structure and activity of this domain rather than demonstrate a direct influence of the His tag on nonspecific DNA binding per se, as has been suggested elsewhere (37). The experiments detailed herein, therefore, were conducted directly with the His-tagged integrase, employing the processed integrase only as a control where appropriate.

Assay of Integrase 3'-Processing and Strand Transfer Activities—Oligonucleotide substrates reflecting the terminal U5 and U3 sequences of the HIV-1 long terminal repeats were 5'-end labeled on the 3'-processed strand (U5a or U3a, Table I) with T4 polynucleotide kinase (Amersham, France) and [γ-³²P]-ATP (3000 Ci/mmol, Amersham). Labeled oligonucleotides were purified from unincorporated nucleotides by gel filtration on a Sephadex G-25 column (Pharmacia). Standard 3'-processing reaction conditions were 30 mM Tris-HCl (pH 7.6), 5 mM MnCl₂, 1 mM diethiothreitol, 0.05% Tween 20 with 50 mM NaCl, 1% glycerol, and 0.5 mM CHAPS contributed by the enzyme storage buffer in a 20-μl final volume. Incubations (pre-warmed) were conducted at 37 °C. The products were analyzed on 20% denaturing polyacrylamide gels.

The binding mechanism is then characterized by a single equilibrium constant for the catalytic processing of the LTR. In the present study, we examined the quantitative description of the stable binding of the HIV-1 integrase to the LTR, under solution conditions more closely reflecting those optimized for 3'-end processing and DNA integration assays. The implications of this data on the quaternary structure of the nucleoprotein complex and the influence of divergent cations on the stability of this interaction are examined.

Theory and Data Fitting Procedures—We adopt here a simple Hill formulation adapted to cases where more than one ligand molecule can be reversibly to a polymeric receptor (we neglect, therefore, any cause of anti-cooperative behavior such as sequence preferences or binding configurations that may lead to lattice exclusion (40), as will be justified under "Results" and "Discussion"). The binding mechanism is then characterized by a single equilibrium constant for the catalytic processing of the LTR. In the present study, we examined the quantitative description of the stable binding of the HIV-1 integrase to the LTR, under solution conditions more closely reflecting those optimized for 3'-end processing and DNA integration assays. The implications of this data on the quaternary structure of the nucleoprotein complex and the influence of divergent cations on the stability of this interaction are examined.

Theory and Data Fitting Procedures—We adopt here a simple Hill formulation adapted to cases where more than one ligand molecule can be reversibly to a polymeric receptor (we neglect, therefore, any cause of anti-cooperative behavior such as sequence preferences or binding configurations that may lead to lattice exclusion (40), as will be justified under "Results" and "Discussion"). The binding mechanism is then characterized by a single equilibrium constant for the catalytic processing of the LTR. In the present study, we examined the quantitative description of the stable binding of the HIV-1 integrase to the LTR, under solution conditions more closely reflecting those optimized for 3'-end processing and DNA integration assays. The implications of this data on the quaternary structure of the nucleoprotein complex and the influence of divergent cations on the stability of this interaction are examined.
where, in this simplified case, the Hill coefficient, $n$, corresponds to the number of integrase oligomers that interact upon binding to the DNA substrate. To account for these cooperative interactions, the titration curve, in terms of fractional saturation ($\theta$), were fitted directly by computer-assisted least squares procedures to the appropriate isotherm corresponding to the above mechanism:

$$
\theta = \frac{[D]}{[D]_0} = 1 + \left( \frac{[I]_0}{[I]_o} \right)^n
$$

where $[D]_0$ refers to the total DNA concentration and $[I]_o$ represents the concentration of free protein required to titrate the DNA to half saturation (for convenience, titrations are expressed in terms of integrase monomers, and consequently the equilibrium binding parameters obtained from the corresponding isotherms are also reported in monomeric units). For a first-order (independent site) binding mechanism ($n = 1$), the isotherm reduces to a hyperbola, and the dissociation constant ($K_d$) is given by $[I]_o$. Whereas $n = 2$, second-order binding is observed, i.e. the degree of polymerization of integrase oligomers is increased by two upon transfer from solution to DNA. In this case, the value $[I]_o$ is equivalent to $/K_c$. Transformed into the familiar Scatchard form, we obtain following the linear expression:

$$
\theta([I]_0) = \frac{1}{K_d} - \theta(K_d)
$$

To obtain the concentration of free integrase, $[I]$, binding isotherms (constructed in terms of total integrase monomer, $[I]_0$) were corrected for the stoichiometric number of binding sites (N) (again expressed in monomeric units). Since the concentration of bound integrase, $[I]_o$, is given by the fraction ($\theta$) of the total concentration of binding sites ($N[D]_0$) occupied, we then may write the following:

$$
[I]_0 = [I] + [I]_o = [I] + \theta N[D]_0
$$

For a series of titrations performed at different concentrations of DNA, at any given value of $\theta$, plotting $[I]_o$ versus $[I]$, gives a straight line from which $[I]_o$ and N may be determined directly from the intercept and slope of the line, respectively. The binding isotherm is generated by repeating this analysis throughout the range of $\theta$ values, essentially as described previously (41).

Dissociation kinetics were analyzed by the first-order exponential decay expression:

$$
[D]_0 = [D]_0 \exp(kt)
$$

where $[D]_0$ is the concentration of bound DNA at time 0, and $k$ is the corresponding dissociation rate constant.

RESULTS

Oligomeric Form of the HIV-1 Integrase—The HIV-1 integrase was expressed and purified as a His-tagged fusion protein (see “Experimental Procedures”). The oligomeric state of this protein was analyzed by gel filtration under conditions of sodium chloride (data not shown), at an intermediate salt concentration 1M NaCl (data not shown), at an intermediate salt concentration 0.4M both monomeric and dimeric species are observed (Fig.1). These data are consistent with the monomer-dimer equilibrium previously reported for both HIV-1 (37) and avian myeloblastosis virus (42) integrases on the basis of sedimentation studies conducted under similar salt concentrations. However, at a lower ionic strength more closely approximating that used in our DNA binding assays, we observed that the integrase exists exclusively in a higher order oligomeric state, minimally forming trimers, though more probably tetramers or larger aggregates.

DNA Binding Studies—We first studied the binding of the integrase to oligonucleotide mimics of the viral U5 and U3 LTRs. Nucleoprotein complexes formed between the HIV-1 integrase and the appropriate DNA substrate were irradiated with a single 5-ns pulse of high energy 266-nm UV light, and adduct formation was monitored by SDS-polyacrylamide gel electrophoresis. By this approach, a single, major cross-linked adduct is observed, exhibiting a molecular mass equivalent to that of the HIV-1 integrase (33 kDa) plus the oligonucleotide strand participating in adduct formation (6.8 KDa). A representative example is shown in Fig. 2, where the intensity of the cross-linked complex can be seen to increase as a function of the integrase concentration. The concentration dependence of the binding reaction does not, however, correspond a simple hyperbolic function expected for an independent, single-site binding mechanism. Instead, the sigmoidal nature of the saturation curve indicates the cooperative binding of the integrase to the LTR (Fig. 3A). This is readily seen when viewed in the form of a Scatchard plot (Fig. 3B). The data points display the humped characteristic of cooperatively associating systems (40). For the 21-bp U5 LTR duplex, the positive cooperativity curves obtained in this way correlate well with those obtained by photochemical reactivity of the occluded site (as detailed under “Experimental Procedures”). As shown below, the titration curves obtained in this way correlate well with those obtained by nuclease digestion techniques when measured under identical conditions (see inset to Fig. 6B), thereby confirming that $\lambda_c$ remains constant over the entire fractional saturation curve. The absolute value of $\lambda_c$ is however dependent on the photochemical reactivity of the DNA sequence occluded by the bound protein, which varies dependent on its base composition. This is seen readily from the comparison of the maximal cross-linked obtained with single-stranded homopolymers (Table I). The photochemical reactivity of the four nucleotide bases may be ordered (relative to poly(dT) = 1) as T (1) > C (0.09) > A (0.08).
Thus, information on the binding affinity of the integrase is obtained from the entire fractional saturation curve as described above and not from the absolute value of DNA cross-linking.

Absence of Binding Specificity—In terms of the total concentration of integrase monomers required to reach half-maximal saturation, no preference was observed for the binding of the duplex viral LTR sequences (Table I). The HIV-1 integrase displays a similar apparent affinity for both single- and double-stranded forms of the LTR and for nonspecific single-stranded DNA. Curiously, although no sequence specificity appears to operate in the binding mechanism, some level of preference is evident from the comparison of the titration of homopolymeric single-stranded sequences. Single-stranded polypurine sequences (poly(dA)) are bound with equal efficiency as the LTR, but not polypurine sequences (poly(dA)). The lower affinity seen for poly(dA) when compared to the polypuridine sequences has been confirmed in direct competition experiments, where poly(dA) fails to compete out the binding of poly(dC) (data not presented). Similarly, poly(dG), which is essentially refractory to UV cross-linking at 266 nm, competes poorly for the binding of poly(dT) or the U5 LTR (data not shown).

The binding of both LTR and nonspecific sequences is notably weaker in the absence of a divalent cation or when Mn$^{2+}$ is replaced by Mg$^{2+}$. In the absence of a detectable specificity for the cognate LTR sequences, we measured the affinity of binding to poly(dT) as a function of the ionic strength (Fig. 4). Several features of this nonspecific interaction are noteworthy. The binding affinity is markedly sensitive to the concentration of monovalent cations. Consistent with the polyelectrolyte theory (43), log $K_{obs}$ (the observed association constant, in this case reported in terms of total monomer) decreases linearly with log [NaCl] (in the range 0.05–0.2 M NaCl). For simple equilibria, this approach may be used to extract information on the number of DNA-phosphateNa$^+$ ion pairs extruded by the binding of a protein ligand to DNA. However, it should be kept in mind that other factors, such as the influence of NaCl on the oligo-
Cooperative Binding of HIV-1 Integrase to DNA

Summary of UV cross-linking titration data for single and double strand oligonucleotides

| Oligonucleotide (5'-3')<sup>a</sup> | Mn<sup>2+</sup> | Mg<sup>2+</sup> | $K_{app}$<sup>b</sup> | $S_{max}$<sup>c</sup> |
|-----------------------------------|--------------|--------------|--------------|--------------|
|                                    | mm          | mm          | $\times 10^{-7}$ | %            |
| U5a:                              |             |             |              |              |
| GATGCAAAGATCTCTAGCAGT              | 10          |             | 1.48 ± 0.36   | 3.38 ± 0.33  |
| U5b:                              |             |             | 0.62 ± 0.18   | 2.54 ± 0.21  |
| U5ds:                             |             |             | 2.34 ± 0.45   | 2.74 ± 0.27  |
| CACAGTTTTTAGAGATGCTCA              | 5           |             | 6.77 ± 1.89   | 2.94 ± 0.25  |
| U3a:                              |             |             | 11.44 ± 2.86  | 2.20 ± 0.46  |
| AGCTAGTGCTTGCT                     | 10          |             | 1.65 ± 0.59   | 2.87 ± 0.36  |
| U3b:                              |             |             | 1.33 ± 0.42   | 2.49 ± 0.56  |
| U3ds:                             |             |             | 1.54 ± 0.26   | 2.67 ± 0.17  |
| TACCTAAACTGCGGAGTGCA               | 10          |             | 1.23 ± 0.41   | 1.89 ± 0.19  |
| DisU5:                            |             |             |              |              |
| -TTGGGCAAGCCTGCTGCCCTCTCTCTCCCTCT | 5           |             | 1.06 ± 0.24   | 0.84 ± 0.28  |
| Poly(dT) (21)                      |             |             |              |              |
| TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT | 10          |             | 3.77 ± 0.56   | 8.63 ± 0.43  |
| Poly(dA) (21)                      |             |             | 16.27 ± 2.29  | 8.32 ± 0.72  |
| Poly(dC) (21)                      |             |             | 4.35 ± 1.84   | 0.34 ± 0.05  |
| Poly(dG) (21)                      |             |             | 1.55 ± 0.79   | 0.80 ± 0.12  |
|                                 |             |             | 0.001         | 0.01         |
|                                 |             |             | Log [cation] (M) | [cation] (M) |
|                                 |             |             | Log [cation] (M) | [cation] (M) |
|                                 |             |             | 0.59          | 0.27         |
|                                 |             |             | 0.28          | 0.36         |
|                                 |             |             | 0.34          | 0.26         |
|                                 |             |             | 0.36          | 0.27         |
|                                 |             |             | 0.27          | 0.34         |

<sup>a</sup> Oligonucleotides are shown in single-stranded form or hybridized to their complements except for disU5, which is a single oligonucleotide capable of forming a branched (dumbbell) structure (24).

<sup>b</sup> The concentration of integrase in total monomer required to achieve half-maximal saturation.

<sup>c</sup> The maximal percent of DNA cross-linked, equivalent to $k_{AC}$, a quantum yield factor dictated by the sequence composition of the oligonucleotide (see under "Experimental Procedures").

Stable Complexes Resistant Attack by Nucleases—Binding of the LTR by HIV-1 integrase results in the formation of a stable complex that resists challenge by DNA or heparin (26, 29, 44). To investigate the assembly of such complexes, a novel approach was developed whereby their formation was probed by way of exposure to the 3'-5'-specific exonuclease, snake venom phosphodiesterase I. The LTR is rapidly degraded when incubated directly with the nuclease; greater than 95% of the DNA is attacked (shortened by one nucleotide or more) within approximately 1–2 s. Complete reduction to single nucleotides occurs within 30 s (data not shown). The degradation is, however, significantly reduced upon the pre-incubation of the LTR with the integrase (Fig. 5, left panel, lane 1). Competitor DNA completely eliminates the protection of the LTR if added before (lane 2), but not after (lane 3), the pre-binding step. The protection of the LTR by the integrase is observed in an all-or-none fashion, irrespective of the strand bearing the radiolabel. The absence of intermediate length products demonstrates that both ends of the DNA molecule are bound coincidentally in the stable complex. Furthermore, the retention of the -2 base deavage product confirms that the substrate 3'-end remains tightly bound and protected from the nuclease following processing. Indeed, the respective dissociation kinetics for the U5 LTR when present in blunt-ended or pre-processed form (lacking the two terminal bases) are indistinguishable ($k_{diss} = 1.5 \times 10^{-5} s^{-1}$). Stable binding does not appear to be sequence specific. Complexes with oligonucleotide substrates bearing no obvious sequence similarity to the LTR exhibit equivalent stability (data not shown). Notably, however, neither strand of the LTR may be protected when present in single-stranded form (Fig. 5, right panel). This contributes clear evidence that the nonspecific binding of HIV-1 integrase to single- and double-stranded DNA involves different binding conformations that may be distinguished on the basis of their stability in dissociation experiments.

Stoichiometry of the Stable Nucleoprotein Complex—In accordance with the UV cross-linking data (above), the binding reaction leading to nuclease-resistant complexes also displays a second-order dependence on the total integrase concentration (Fig. 6). To obtain a quantitative analysis of the binding density of the integrase, titrations were performed with several different concentrations of the DNA substrate, thereby allowing the bound and free forms of the integrase to be discriminated on the basis of the conservation of mass (41). An example of this...
Type of analysis is shown in Fig. 6C. The concentration of integrase required to achieve half-maximal saturation (in terms of the fraction of DNA protected) indicates that the bound complexes are constructed from approximately 10 monomers per 21-bp DNA duplex. The free enzyme concentration is obtained by extrapolation of the data to the intercept with the vertical axis, where the concentration of bound integrase is effectively zero. The broken line in Fig. 6B traces the theoretical binding isotherm, in terms of the free integrase concentration, obtained from the systematic repetition of this type of analysis throughout the fractional saturation curve. The solid lines in Fig. 6B also represent theoretical second-order binding curves. These were calculated from the binding isotherm after transformation to account for a binding density, at saturation, of 10.5 monomers per LTR. In this way, good accordance between data and theory is seen, demonstrating that the binding density obtained is a fair approximation of that existing over the entire binding curve.

Divalent Cation Requirement—The formation of the stable complex does not exhibit an absolute requirement for the presence of a divalent cation. Nuclease-resistant complexes appear, albeit at low efficiency, when the integrase is incubated with the US LTR in the absence of the cation (Fig. 7). The protection of the LTR is enhanced considerably in the presence of Mn$^{2+}$ but only poorly with Mg$^{2+}$ cations. These results are entirely consistent with the influence of the respective divalent cations on the apparent DNA binding affinity reported in Table I.

A sigmoidal response, in terms of activity versus cofactor concentration, is predicted to arise from a cofactor-induced oligomerization of a protein where the oligomeric form interacts preferentially with the substrate. In the case of the integrase, Mn$^{2+}$ clearly enhances binding in a non-hyperbolic fashion. The Hill coefficient ($n^H$) measured for the Mn$^{2+}$-induced transition observed in Fig. 7 is $4.2 \pm 1.0$. A similar value ($n^H = 5.0 \pm 1.2$) was obtained after the proteolytic removal of the His-tag sequence from the recombinant protein, confirming that this property is inherent to the integrase. Consequently, the kinetics of 3'-end processing are also markedly non-linear with respect to the Mn$^{2+}$ concentration (data not shown). These data are thus consistent with the cooperative binding mechanism and indicate that the metal ion promotes the protein-protein interactions required for the stable binding of the DNA substrate.

A further remarkable feature of the transition is the concentration range over which it is effected. The apparent binding constant ($K_{app}$) for Mn$^{2+}$ estimated from several such titrations, for both His-tagged and factor Xa cleaved proteins, is $64 \pm 10 \mu M$. Whereas the optimum in vitro 3'-end processing and disintegration activities of the integrase are obtained at high (non-physiological) concentrations of Mn$^{2+}$, comparatively poor activity is exhibited with Mg$^{2+}$ (27, 37, 45). To further investigate the role of the divalent cation in the formation of the stable complex, the respective dissociation and 3'-end processing kinetics of nuclease-challenged complexes were compared (Fig. 8). Complexes formed with Mn$^{2+}$ are considerably more stable than the corresponding complexes formed with Mg$^{2+}$, exhibiting half-lives of approximately 13 and 1 h, respectively. However, the first-order rate constant obtained for the 3'-end processing reaction is similar in the presence of either cation. By eliminating the uncleaved DNA, we follow only the enzymatic activity of the stable complex. The rate constant obtained then gives a direct measurement of $k_{cat}$, which, under these conditions, was $0.046 \pm 0.011\min^{-1}$ and $0.069 \pm 0.008\min^{-1}$ for Mn$^{2+}$ and Mg$^{2+}$, respectively. Such low rates are consistent with the delayed appearance of processed 3'-termini in virus-infected cells, which accumulates over a period of several hours (1), and is probably an intrinsic function of the catalytic properties of the integrase. Nonetheless, we may conclude that the overall higher efficiency of 3'-end processing in Mn$^{2+}$ observed in vitro may be explained simply by the augmentation in the stability of the interaction with the substrate and not to an inherent preference for this cation for catalytic function.

**DISCUSSION**

In this report, we have described the development of novel procedures to probe the interaction between HIV-1 integrase and DNA in vitro as the preliminary step in a detailed study of the structural architecture of these nucleoprotein complexes. Whereas the first approach, laser UV cross-linking, detects binding via the intimate protein-DNA contacts formed within an occluded photoreactive site, for nonspecific binding it is indiscriminately as to the position of that site on the DNA molecule. On the other hand, protection of the DNA from hydrolysis by exonucleases detects principally the stable contacts formed with the ends of the DNA substrate. The fact that identical binding isotherms are generated by these two different techniques, combined with the absence of intermediate
length nuclease digestion products, strongly implies that the entire DNA molecule, including both ends, is occluded upon the formation of the stable nucleoprotein complex. Since non-occluded regions and unbound ends appear to be absent (both of which, if present, could potentially have introduced anti-cooperative effects (40)), the binding isotherm is expected to be influenced only by the positive cooperativity involved in this mechanism. In this case, the corresponding Hill coefficient should be representative of the number of oligomers participating in the assembly of the stable complex (see "Experimental Procedures"). For both techniques, the second-order binding observed with the 21-bp LTR is consistent with a mechanism involving the cooperative dimerization of a high-order oligomeric form of the protein. From a density of $10.5 \times 10^6$ monomers per LTR and a Hill coefficient of 2, the oligomeric form of the integrase in solution is estimated to be pentameric, though, within the limits of reasonable experimental error, the average degree of oligomerization (M) in solution may be extended to incorporate tetrameric or hexameric forms ($4 \leq M \leq 6$). Clearly, the resultant high-order nucleoprotein complex should exhibit a substantial molecular mass ($>300$ kDa). This may, in part, provide an explanation for the sedimentation properties of such complexes (13) and may indicate the reason why such complexes are not particularly suited to gel retardation studies.

We have evaluated the cooperativity for the binding of short oligonucleotide duplexes only. In this limiting case, the observed cooperativity may be restricted, by reason of the number of bases occluded by a single oligomer, to a single cooperative (apparent dimerization) event. Recent experiments have enabled us to observe the complexes formed with larger DNA fragments directly by electron microscopy. In the presence of $\text{Mn}^{2+}$, large clusters of bound integrase molecules, extending inwards from the DNA ends, have been detected, indicating that the associated cooperativity parameter may be quite large. These findings appear to comply with the extended, nucleosome-like DNase I footprints observed for the interaction between the avian myeloblastosis virus integrase on plasmid-borne LTRs (46). Accordingly, the cooperativity parameters will require further investigation with longer DNA substrates. We may, however, infer from these data, that in the absence of other cellular factors, the integrase oligomer binds contiguously to naked DNA rather than to distinctly separated sites. For the reasons explained above, the binding of the short LTR.

---

2 I. K. Pemberton and E. LeCam, unpublished observations.
substrates thus presents itself as a useful model system with which to examine binding under conditions of limited cooperativity.

In agreement with several previous reports (6, 11, 13–17), no sequence specificity was observed in the binding of the viral LTR nor was a preference apparent for the double-stranded substrate in UV cross-linking titrations. Stable complexes are, however, not formed with single-stranded DNA, suggesting the underlying mechanism of binding may be different for these two forms of DNA. Our recent data suggest that the binding density of integrase for single-stranded DNA (poly(dT)) is considerably lower than that observed for the US LTR duplex. We expect that the differential dissociation rates obtained with such complexes may reflect this observation. We are currently investigating this in more detail.

We have observed that Mn$^{2+}$ can stimulate binding at concentrations between one and two orders of magnitude lower than that previously reported as the optimum concentration for $3'$-processing for both the HIV-1 (27) or Moloney murine leukemia virus (45) integrases. An explanation for this disparity could be in the existence of two distinct divalent cation binding sites on the integrase oligomer: one involved in mediating protein-protein interactions, and hence stimulating cooperative binding, and another incorporated at the catalytic site. Indeed, a region distinct from the catalytic site, the HHCC finger domain, has recently been implicated in the oligomerization of the HIV-1 integrase in the presence of divalent cations, a property further shown to be essential for stable complex formation (26). Such interactions probably underlie the cooperativity of DNA binding, much as the coordination of Zn$^{2+}$ by the HXCXCXHX$_2$C motif (47) of the T4 bacteriophage gene 32 protein (gp32) has been inferred to stimulate cooperative binding to single-stranded DNA via enhanced protein-protein interactions (48). The presence of Zn$^{2+}$ is not necessary to obtain cooperative binding of gp32 but stimulates it approximately 30-fold (48). By analogy, the coordination of a metal ion by the integrase HHCC motif may stabilize a particular conformation of this subdomain and thereby enhance a property already inherent in the integrase to self-associate. Although the precise Mn$^{2+}$ interaction site remains to be resolved, the general assumption would appear to be correct. Stable complexes form in the absence of the divalent cation but only efficiently at much higher protein concentrations. Thus, the cation appears to favor the polymerization process by a mechanism that lowers the activation energy barrier to this association. A similar observation has been made for the self-association equilibrium of the RecA protein, which is stimulated in the presence of Mg$^{2+}$ (49). The capacity of such proteins to associate in the absence of DNA appears to be a recurring theme for proteins that are known to form direct and functional interactions in the context of a highly structured DNA-protein complex.

We have demonstrated that the ternary nucleoprotein complex formed with Mn$^{2+}$ is considerably more stable than the corresponding complex formed with Mg$^{2+}$. The absolute rate of $3'$-end processing exhibited by the stable complex is, however, similar with either cation. This allows us to deduce that the marked preference exhibited in vitro for Mn$^{2+}$ is simply a consequence of the enhanced stability of binding to the LTR. Furthermore, it was reported (after submission of this manuscript) that an increase in the efficiency of the $3'$-end process-

---

3 I. K. Pemberton, M. Buckle, and H. Buc, unpublished observations.

---

**Fig. 7.** Cooperative stimulation of the formation of stable complexes by Mn$^{2+}$. Integrase (350 nM) was pre-incubated with the US LTR (12.5 nM) for 5 min prior to incubation with phosphodiesterase for 10 min at 37°C. The plot shows the consequence of altering the concentration of MnCl$_2$ (squares) and MnCl$_2$ (triangles) on the extent of DNA protected.

**Fig. 8.** Comparison of the kinetics of dissociation and $3'$-end processing exhibited by stable complexes formed in the presence of Mn$^{2+}$ or Mg$^{2+}$. Integrase (390 nM) was pre-incubated with the US LTR for 3 min at 37°C prior to addition of phosphodiesterase. Incubations were stopped at appropriate time intervals and analyzed by gel electrophoresis to discriminate the percentage of the original substrate remaining and the extent of $3'$-end processing of the stable complexes (as detailed under "Experimental Procedures"). Left panel, kinetics of dissociation (circles) and $3'$-end processing (triangles) in the presence of 5 mM MnCl$_2$. Right panel, kinetics of dissociation (circles) and $3'$-end processing (triangles) in the presence of 5 mM MgCl$_2$. The time reported for dissociation is with respect to addition of the nuclease. For $3'$-end processing, this refers to total incubation time. Data points are fitted by non-linear least squares analysis to the appropriate first-order expression.
Cooperative Binding of HIV-1 Integrase to DNA

Cooperative binding of the LTR substrate in Mg\(^{2+}\) can be achieved either by altering the solution conditions (notably to include PEG) (50) or by increasing the length of the DNA substrate (51). In both cases, these effects can be explained simply by a concomitant stimulation in the binding interaction with the DNA substrate and add further evidence to support our conclusion that the bias for Mn\(^{2+}\) in standard in vitro reactions operates at the level of DNA binding and not catalysis. The requirement for Mn\(^{2+}\) is not seen with the pre-assembled viral nucleoprotein complex purified from infected cells, which undergoes efficient integration in vitro in the presence of Mg\(^{2+}\) (52). Clearly, since the preference exhibited for Mn\(^{2+}\) in vitro is to stimulate efficient binding, the recombination protein to the DNA substrate, the necessity for this cation has already been circumvented when using the preformed integration machinery. Other cellular components, potentially host and/or viral accessory proteins, may be required to regulate the assembly of the authentic nucleoprotein complex. In this respect, the presence of several viral proteins other than integrase has been noted in the pre-integration complexes (5). In the absence of such factors, the physiological role of the metal ion-induced interaction between integrase oligomers remains unclear. It is possible that the integrase binds to the viral DNA ends as a structured and highly stable aggregate. It also appears likely that some interaction should take place between integrase molecules bound at the extreme ends of the viral genome to allow the coordinated integration of the viral termini at a discrete target site in the host genome. This is accomplished in the analogous transposition mechanism of bacteriophage Mu by coordination of distant MuA binding sites into a tetrameric MuA synaptic complex (53). By comparison, a direct interaction between the distally bound integrase complexes may bring the LTR termini together to form the legitimate strand-transfer complex. Indeed, our recent observations and those of Mazumder et al. (33) suggest that the assembly of complexes is not restricted to lateral, cooperative interactions, but may also involve intramolecular interactions between bound complexes. Undoubtedly, a greater understanding of the process of retroviral integration will necessitate a more detailed exploration of the structural architecture and cofactor requirements of the functional integrase nucleoprotein complexes. We expect that the approaches outlined in this report will greatly assist these studies.

Acknowledgments—We thank Jane MacDougall, Richard Buckingham, and J. M., Eric LeCam, and John Oxford for advice and stimulating discussions during the course of this work.

REFERENCES

1. Roth, M. J., Schwartzberg, P. L., and Goff, S. P. (1989) Cell 58, 47-54
2. Brown, P. O., Bowerman, B., Varum, H. E., and Bishop, J. M. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 2529-2532
3. Vora, A. C., Fitzgerald, M. L., and Grandgenett, D. P. (1990) J. Virol. 64, 6566-6569
4. Engelman, A., Mizuochi, K., and Craige, R. (1990) Cell 67, 1211-1221
5. Burinsky, M. I., Sharova, N., McDonald, T. L., Pushkarskaya, T., Tarpley, W. G., and Stevenson, M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6125-6129
6. Sherman, P., and Fyfe, J. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 5119-5123
7. Katzen, M., Katz, R. A., Skalka, A. M., and Leis, J. (1989) J. Virol. 63, 5319-5327
8. Bushman, F., Fujivaiwa, T., and Craige, R. (1990) Science 249, 1555-1558
9. Vink, C., Van Gent, D. C., Elgersma, Y., and Plasterk, R. H. A. (1991) J. Virol. 65, 4636-4644
10. Leavitt, A. D., Rose, R. B., and Varmus, H. E. (1992) J. Virol. 66, 2359-2368
11. LaFemina, R. L., Callahan, P. L., and Cordingly, M. G. (1991) J. Virol. 65, 5624-5630
12. Bushman, F., and Craige, R. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 1339-1343
13. van Gent, D. C., Elgersma, Y., Bolk, M. W. J., Vink, C., and Plasterk, R. H. A. (1991) Nucleic Acids Res. 19, 851-860
14. Schauer, M., and Billich, A. (1992) Biochim. Biophys. Res. Commun. 185, 874-880
15. Khan, E., Mack, J. P. G., Katz, R. A., Kulkosky, J., and Skalka, A. M. (1991) Nucleic Acids Res. 19, 851-860
16. Sherman, P., and Fyfe, J. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 3051-3059
17. Katzman, M., Katz, R. A., Kulkoisky, J., and Skalka, A. M. (1991) Nucleic Acids Res. 19, 851-860
18. Vora, A. C., Fitzgerald, M. L., and Grandgenett, D. P. (1990) Proc. Natl. Acad. Sci. U. S. A. 90, 3429-3432
19. Engelman, A., and Craige, R. (1992) J. Virol. 66, 6361-6369
20. Chow, S. A., Vincent, K. A., Ellison, V., and Brown, P. O. (1992) Science 255, 723-726
21. Burke, C. J., Sanayei, G., Bruner, M. W., Ryan, J. A., LaFemina, R. L., Robbins, H. L., Zett, A. S., Middaugh, C. R., and Cordingly, M. G. (1992) J. Biol. Chem. 267, 9639-9644
22. Ellison, V., Gerton, J. V., Vincent, K. A., and Brown, P. O. (1995) J. Biol. Chem. 270, 3320-3326
23. Vincent, K. A., Ellison, V., Chot, S. A., and Brown, P. O. (1993) J. Virol. 67, 425-437
24. Jonsson, C. B., and Roth, M. J. (1993) J. Virol. 67, 5562-5571
25. Vink, C., Puras Lutzke, R. A., and Plasterk, R. H. A. (1994) Nucleic Acids Res. 22, 4103-4110
26. Kajjana, G., and Goff, S. P. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 10593-10597
27. Engelman, A., Bushman, F. D., Craige, R. (1993) EMBO J. 12, 3269-3275
28. van Gent, D. C., Vink, C., Oude Groeneveld, A. A. M., and Plasterk, R. H. A. (1993) EMBO J. 12, 3261-3265
29. Mazumder, A., Engelman, A., Craige, R., Fesen, M., and Pommier, Y. (1994) Nucleic Acids Res. 22, 1037-1043
30. Farnet, C. M., and Haseltine, W. A. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 1419-1425
31. Misra, T. K., Grandgenett, D. P., and Parsons, J. T. (1982) Biochemistry 21, 2525-2529
32. Misra, T. K., Grandgenett, D. P., and Parsons, J. T. (1982) Biochemistry 21, 2525-2529
33. Mishra, T. K., Grandgenett, D. P., and Parsons, J. T. (1982) J. Virol. 44, 330-343
34. Qi, H., and Giedroc, D. P. (1994) Biochemistry 33, 8139-8148
35. Nadler, S. G., Roberts, W. J., Shao, Y., and Williams, K. R. (1990) J. Biol. Chem. 265, 10389-10394
36. Huiskok, R. W., and DiCapua, E. (1991) Biochemistry (Paris) 73, 191-198
37. Engelman, A., and Craige, R. (1995) J. Virol. 69, 5908-5915
38. Paul Lee, S., Kim, H. G., Censullo, M. L., and Han, M. K. (1995) Biochemistry 34, 10205-10214
39. Farnet, C. M., and Haseltine, W. A. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 3051-3059
40. Lavoie, B. D., Chan, B. S., Allison, R. G., and Chaconas, G. (1991) EMBO J. 10, 3031-3039