Folding of the multidomain human immunodeficiency virus type-I integrase

DUANE P. GRANDGENETT AND GOODARZ GOODARZI

Institute for Molecular Virology, St. Louis University, St. Louis, Missouri 63110

(RECEIVED December 7, 1993; ACCEPTED March 28, 1994)

Abstract

Protein folding conditions were established for human immunodeficiency virus integrase (IN) obtained from purified bacterial inclusion bodies. IN was denatured by 6 M guanidine·HCl–5 mM dithiothreitol, purified by gel filtration, and precipitated by ammonium sulfate. The reversible solvation of precipitated IN by 6 M guanidine·HCl allowed for wide variation of protein concentration in the folding reaction. A 6-fold dilution of denatured IN by 1 M NaCl buffer followed by dialysis produced enzymatically active IN capable of 3' OH end processing, strand transfer, and disintegration using various human immunodeficiency virus-1 (HIV-1) long terminal repeat DNA substrates. The specific activities of folded IN preparations for these enzymatic reactions were comparable to those of soluble IN purified directly from bacteria. The subunit composition and enzymatic activities of IN were affected by the folding conditions. Standard folding conditions were defined in which monomers and protein aggregates sedimenting as dimers and tetramers were produced. These protein aggregates were enzymatically active, whereas monomers had reduced strand transfer activity. Temperature modifications of the folding conditions permitted formation of mainly monomers. Upon assaying, these monomers were efficient for strand transfer and disintegration, but the oligomeric state of IN under the conditions of the assay is indeterminate. Our results suggest that monomers of the multidomain HIV-1 IN are folded correctly for various catalytic activities, but the conditions for specific oligomerization in the absence of catalytic activity are undefined.

Keywords: protein folding; retrovirus integrase

Integration of the retrovirus DNA genome into the host genome is necessary for virus replication (Lewis et al., 1992). Considerable progress has been made toward our understanding of the molecular mechanisms involved in retrovirus integration through the use of bacterially expressed retrovirus integrase (IN) (Bushman et al., 1990; Katz et al., 1990; Sherman & Fyfe, 1990; Vink et al., 1991a, 1991b). Purified IN accomplishes long terminal repeat (LTR) DNA 3' OH end processing and DNA joining reactions observed *in vivo*. Expressed IN is also capable of reversing the joining step using DNA substrates that mimic the insertion of a single LTR terminus into a target molecule (Chow et al., 1992). Expressed IN can efficiently perform the insertion of a single LTR DNA end into a single strand of target DNA (LTR half-site) (Mizuuchi, 1992). However, the enzyme cannot efficiently perform the concerted insertion of both viral DNA ends into a target DNA molecule (LTR full-site) (Bushman & Craigie, 1991; Engelman et al., 1993). Integrase purified from avian myeloblastosis virus can efficiently perform the concerted integration of viral-like DNA substrates (Fitzgerald et al., 1992).

Both conserved amino acid sequence motifs and direct functional assays have identified several domains in the integrase of retroviruses (Fig. 1). From the amino-terminus, a "zinc finger" motif (HX₃₋₇-HX₂₃₋₃₂CX₂C) can physically bind zinc (Burke et al., 1992), and this region is required for efficient 3' end processing and strand transfer (Engelman & Craigie, 1992; Vincent et al., 1993; Vink et al., 1993). The adjacent motif, DX₃₉₋₅₈DX₃₅E, is essential for catalysis (3' end processing, strand transfer, disintegration) (Drelich et al., 1992; Kulkosky et al., 1992; Bushman et al., 1993), and this catalytic core, which is independently capable of disintegration, maps between residues 50 and 186 of human immunodeficiency virus-1 (HIV-1) IN. This catalytic core also possesses a region responsible for dimerization of IN, which is an essential property for catalysis (Bushman et al., 1993; Engelman et al., 1993; van Gent et al., 1993a). Lastly, a non-specific DNA binding region maps approximately between residue 177 and the near COOH-terminus (Kahn et al., 1991; Mumm & Grandgenett, 1991; Woerner et al., 1992) and is required for efficient 3' end processing and strand transfer but not for the disintegration reaction (Engelman et al., 1993; Vink et al., 1993).

We wanted to explore protein folding procedures that would permit the control of folding of HIV-1 IN *in vitro* instead of *in*

Reprint requests to: Duane P. Grandgenett, Institute for Molecular Virology, St. Louis University, 3681 Park Avenue, St. Louis, Missouri 63110.

Regions Required For:

Fig. 1. Mapping of the multidomain HIV-1 integrase. **Top:** Regions required for specific functions associated with the integrase are defined (see introduction for specific references). **Bottom:** The dots on top of the box indicate specific amino acids that are highly conserved among various retrovirus integrases (Lin & Grandgenett, 1991; Kulkosky et al., 1992).

the bacterial environment. The high-level expression of eukaryotic proteins in *Escherichia coli* sometimes results in the formation of insoluble inclusion bodies, including HIV-1 IN (Sherman & Fyfe, 1990). Some purification techniques to produce soluble and enzymatically active IN involve a high-salt extraction of the insoluble pellet containing IN derived from lysed bacterial cells. In this report, solubilization of HIV-1 IN from purified insoluble particles was accomplished by treatment with 6 M guanidine·HCl (Gdn·HCl)–5 mM dithiothreitol (DTT). Integrase was further purified by gel filtration in the presence of 6 M Gdn·HCl–5 mM DTT and stored as an ammonium sulfate precipitate. For protein folding reactions, various amounts of precipitate were dissolved in 6 M Gdn·HCl–5 mM DTT. Folding parameters were investigated for production of enzymatically active IN. The simplest and most efficient folding procedure identified heretofore involved a dilution of denatured IN to 1 M Gdn·HCl with 1 M NaCl buffer followed by incubation on ice and dialysis against the same buffer. The enzymatic activities previously shown to be associated with bacterially expressed HIV-1 IN—3' end processing, half-site strand transfer reactions, and disintegration activity—were associated with folded IN. The subunit composition of the various enzymatically active IN preparations was analyzed by glycerol gradient centrifugation. To produce enzymatically active monomers in solution, it was necessary to maintain ice temperatures at the dilution step. These monomers are apparently capable of dimerization under conditions of assay, since dimerization is necessary for catalysis (Bushman et al., 1993; Engelman et al., 1993; van Gent et al., 1993b).

Results

Purification of HIV-1 IN

Purification of expressed HIV-1 IN (pNY clone) from *E. coli* was facilitated by isolation of insoluble inclusion bodies. The lysis of cells expressing HIV-1 IN was accomplished by hypotonic treatment. The inclusion bodies were separable from the majority of the contaminating proteins by multiple washing and centrifugation steps. The physical passage of the suspended aggregates in buffer through a 20-gauge needle enhanced the

purity of IN in the final inclusion body preparations. IN constituted approximately 50% of the total protein mass in the purified inclusion bodies as analyzed by SDS-PAGE (data not shown).

Further purification of HIV-1 IN was accomplished by chromatography of 6 M Gdn·HCl–5 mM DTT treated inclusion bodies on a gel filtration column (Fig. 2). DTT was included to disrupt disulfide bonds and protein aggregates. IN was localized by absorbance at 280 nm and SDS-PAGE. IN constituted approximately 80% of the protein at peak absorbance. Selected fractions were pooled (Fig. 2). Routinely, between 3.0 and 5.0 mg of IN at approximately 80% purity were obtained from 2 L of cells (6 different preparations). This protein concentration is based on the calculated extinction coefficient of 1.59 OD/mg of IN in 6 M Gdn·HCl at 280 nm (Gill & von Hippel, 1989). Pooled fractions of Sephacryl-S200 purified IN were dialyzed overnight against 2.65 M (NH₄)₂SO₄, and the precipitate was stored on ice. This storage procedure provided IN with a stable environment for months.

Standard folding reaction

Storage of IN as an (NH₄)₂SO₄ precipitate provided a versatile tool for controlling the protein concentration in the folding reaction. Various amounts of the precipitate were centrifuged at 12,000 × *g* and the supernatant removed. The precipitate was solubilized with 60 μL of 6 M Gdn·HCl–5 mM DTT and incubated for 90 min at 22 °C. Under standard folding conditions, the concentration of Gdn·HCl solubilized IN was varied from 15 to 90 μM. The folding reaction was initiated by the rapid addition of 300 μL of ice-cold 1 M NaCl buffer, followed by incubation on ice for 3 h. Decreasing this ice incubation time to 1.5 h had no noticeable effect on the folding reaction. The sample was dialyzed in an inverted 1.5-mL microfuge tube as described in the Materials and methods. This standard folding procedure resulted in enzymatically active IN that was compared to "control" IN purified directly from *E. coli* (Engelman & Craigie, 1992) (Figs. 3, 4).

Several parameters were modified to investigate the folding reaction. Light scattering is a convenient tool for measuring the appearance of protein aggregates in solution. An increase in light

Fig. 2. Purification of pNY HIV-1 IN using Sephacryl S-200 HR gel filtration under denaturing conditions. Purified inclusion bodies from 2 L of cells expressing HIV-1 IN were treated with 6 M Gdn·HCl–5 mM DTT. The sample was chromatographed as described in the Materials and methods. The absorbance of each fraction was recorded at 280 nm. The fraction volume was 0.81 mL, and 5 μ L was used for SDS-PAGE (see *inset*). The protein was stained with Coomassie blue. The void volume was determined by using blue dextran. Chymotrypsinogen (Worthington) was used for a 25,000-Da molecular weight marker.

Fig. 3. Catalytic activities of pNY HIV-1 IN derived from folded enzyme or purified directly from bacteria. Shown are 3' end processing, strand transfer, and disintegration activities. The DNA substrates are illustrated above each panel. **A:** 3' end processing and strand transfer. The substrate (5.5 pmol) is a blunt-ended 20-bp oligonucleotide containing U5 end sequences. The arrow indicates the cleavage of the dinucleotide. The DNA was 5' end labeled with 32 P on the processed strand. IN is absent in lane 1. Lanes 2 and 3 contained 80 nM and 160 nM folded IN, while lanes 4 and 5 contained 65 nM and 130 nM purified IN from bacteria (Engelman & Craigie, 1992). The processed LTR DNA strand, 18 nucleotides in length, and the DNA integration products (IP) formed by IN are indicated. Disintegration assays were carried out with a dumbbell (5.2 pmol) (**B**) or with a Y (6.2 pmol) substrate (**C**). The structure of each DNA substrate used is shown above its panel, with the bold lines representing viral sequences. The input dumbbell substrate (38-mer) was processed to a 14-bp labeled product, whereas the 5' end labeled 16-mer of the Y substrate was joined to unlabeled target sequences by a "disintegration" reaction to form a labeled 30-mer product (Vincent et al., 1993). The labeled 16-mer oligonucleotide was not purified by a sequencing gel prior to hybridization. The reactions in lanes 6–10 and 11–15 contained the same amounts of IN as did lanes 1–5, respectively. After electrophoresis, the dried gel was exposed to X-ray film. Panel A was exposed with an intensifying screen at -70°C for 16 h, and panels B and C were exposed without a screen for 72 h.

Fig. 4. Quantitation of catalytic activities of folded pNY HIV-1 IN. Shown is a quantitative comparison of activities of folded and control IN. The substrates employed for dinucleotide cleavage and disintegration are described in the legend of Figure 3. The strand transfer only reaction employed a 5' end labeled U5 19-mer hybridized to a complementary 21-mer, producing a recessed substrate. **A:** The 3' OH recessed DNA substrate (5 pmol) was incubated with control IN (●) and folded IN (▲) at the indicated concentrations. Reactions were stopped after 10 min within the linear range of the reaction. The input substrates and strand transfer products were separated on denaturing gels and quantitated using a Molecular Dynamics Phosphorimager. The percentage of products formed from the input substrate is shown on the left. **B:** Conditions were identical to A except the 18-mer processed strand from the 20-bp U5 substrate (10.8 pmol) (Fig. 3) was quantitated. **C:** The Y substrate (5 pmol) was incubated for 10 min with each enzyme, as described in A. The conversion of the 5'-labeled 16-mer oligonucleotide to the 30-mer product as shown in Figure 3 was quantitated. **D:** The conversion of the 5'-labeled 38-mer dumbbell substrate (4.8 pmol) to the 14-mer cleaved product as shown in Figure 3 was quantitated. The reaction time was 10 min.

scattering measured at 340 nm would suggest that IN was forming aggregates. For enzymatically active preparations, the 280/340 nm ratio of folded IN between 3 and 10 μ M was generally around 15–18 or slightly higher. Centrifugation of preparations that possessed this ratio at $23,000 \times g$ for 30 min did not improve the ratio, suggesting few physically large-sized aggregates in these samples. In addition, there was no loss of HIV-1 IN enzymatic activities in the supernatant upon centrifugation (data not shown). Increasing the temperature at the dilution step

to 22 °C instead of 2 °C, with IN at a final concentration of 5 μ M, resulted in a marked decrease of the 280/340 nm ratio to 2.3, with visible aggregates present. An additional parameter affecting the folding of HIV-1 IN was dialysis. After dilution of IN to 1 M Gdn·HCl using the standard protocol, the sample was dialyzed by an inverted microfuge tube technique. The exposed surface of the dialysis membrane on the microfuge tube cap was a 7-mm circle. Dialysis of the 1 M Gdn·HCl diluted sample in dialysis tubing was not possible due to an approximate 10-fold higher dialysis membrane surface. The more rapid exchange of solutions occurring with the dialysis tubing versus the microfuge tube resulted in the formation of visible protein aggregates.

Enzymatic analysis of folded HIV-1 IN

A stringent criterion for correct protein folding is enzymatic activity. Expressed HIV-1 IN has 3 enzymatic activities *in vitro*. They are LTR 3' OH end processing, strand transfer, and disintegration. A blunt-ended 20-bp U5 LTR substrate was used for the first 2 reactions (Fig. 3A) or, later, a 19/21-bp 3' OH recessed U5 substrate was used to measure strand transfer only (Fig. 4A). For the disintegration assay, a 5' end labeled 38-mer dumbbell substrate (Fig. 3B) or a 5' end labeled 16-mer complexed to 2 other oligonucleotides producing a Y substrate (Fig. 3C) was used. We compared various folded preparations of HIV-1 IN both qualitatively (Fig. 3) and quantitatively (Fig. 4) with the same HIV-1 IN (control) purified directly from bacteria (Engelman & Craigie, 1992).

First, optimum assay conditions were established with folded HIV-1 IN for measuring 3' OH end processing and strand transfer using Mn^{2+} as the divalent metal ion (data not shown). Optimum assay conditions were very similar to published assay conditions for HIV-1 IN purified directly from bacteria. These results indicate that Gdn·HCl treatment did not adversely alter the enzymatic functions of IN.

Qualitatively, folded IN was capable of specific dinucleotide cleavage of a blunt-ended U5 LTR end and subsequent strand transfer of the processed 18-mer into integration products (Fig. 3A, lanes, 2, 3), similar to what was observed with control IN (Fig. 3A, lanes 4, 5). Incubation was for 30 min and therefore not in the linear range of the reaction. However, control IN performed better than folded IN for 3' end trimming and strand transfer (see Fig. 4 also).

The partially purified and folded HIV-1 IN preparations have exonuclease activity producing a 19-nucleotide product. The contribution of the 19-nucleotide product to the formation of the 18-mer product by further exonuclease digestion was probably minimal. As shown in Figure 3B (lanes 7 and 8), the 5' end labeled 14-nucleotide product released by IN from the dumbbell substrate is also attacked by the exonuclease, but the formation of a -2 product (12-mer) is also minimal. These results suggest that the majority of the 18-nucleotide product observed is due to dinucleotide cleavage of the U5 substrate by IN (Fig. 3A).

Folded HIV-1 IN was capable of using 2 different substrates for the disintegration reaction (Fig. 3B, lanes 7 and 8; 3C, lanes 12 and 13). Both folded and control IN enzymes were able to effectively convert a larger proportion of their disintegration substrates to products than was observed for 3' end processing and strand transfer reactions (Fig. 3A). Nearly equivalent quantities of products were formed with both substrates with either folded or control IN. With the dumbbell substrate, the 3' OH-

terminus was used by IN as a nucleophile to attack the viral/target junction, releasing the labeled 14-mer nucleotide containing viral sequences (Fig. 3B, lanes 6–10). With the Y substrate, the 3' OH-terminus of the labeled 16-mer also acted as a nucleophile, but the ligated product (30-mer) was observed (Fig. 3C, lanes 11–15). In this latter experiment, the products were not completely denatured, resulting in a band migrating slightly more slowly than the 30-nucleotide product (see also Fig. 5C) (Vincent et al., 1993). These data established that folded HIV-1 IN was capable of catalyzing the same enzymatic reactions as IN purified directly from bacteria.

The enzymatic activities of these folded IN preparations were stable upon multiple freeze-thawings and for at least 4 months of storage at -70°C . Prior heat treatment (85°C for 10 min) of folded IN or addition of 1% SDS in the reaction mixture eliminated enzymatic activity using the 20-bp LTR U5 substrate. As substrate controls, folded HIV-1 IN was not capable of performing 3' end processing or DNA strand transfer using a blunt-ended 15-bp avian retrovirus LTR U3 oligonucleotide or an 18/20-bp 3' OH recessed LTR U3 substrate.

Quantitative enzymatic assay conditions were established to compare the specific activities of folded IN to those of control

Fig. 5. Analysis of folded pNY HIV-1 IN by glycerol gradient sedimentation. Conditions for sedimentation are described in the Materials and methods. **A:** Folded HIV-1 IN (128 μg), avian myeloblastosis virus $\alpha\beta$ DNA polymerase ($\alpha\beta$), bovine serum albumin (BSA), or ovalbumin (OV) was sedimented in individual tubes. Fractions (120 μL) were collected from the bottom. Catalytic activities of HIV-1 IN and $\alpha\beta$ DNA polymerase were measured, and the $\text{OD}_{280\text{nm}}$ profiles of BSA and OV were determined. The arrows indicate the location of the molecular weight markers expressed as kilodaltons. Aliquots (5 μL) of each even fraction were assayed as described in the legend of Figure 3, except the total reaction volume was 80 μL to lower the NaCl concentration to 43 mM. The strand transfer and the disintegration assays were incubated for 1 h and 15 min, respectively. The products produced by IN using the Y substrate, the dumbbell, and the recessed U5 substrate are expressed as pixels derived by Phosphorimager analysis. The protein concentration was determined by quantitating the amount of silver-stained (Bio-Rad) IN in various fractions, shown in D, using a Zeineh video scanning densitometer. On a separate gel, various amounts of BSA were mixed with the above IN fractions and were subjected to electrophoresis, silver staining, and densitometry for quantitation. Both IN and BSA were in the linear range for silver staining. **B:** The log molecular weights of $\alpha\beta$ DNA polymerase, BSA, and OV are plotted against fraction number. **C:** Aliquots (5 μL) of the indicated fractions were assayed with the Y substrate. Twelve percent of the reaction was analyzed on a DNA sequencing gel. The gel was dried and exposed for 30 h at -70°C with an intensifying screen. **D:** Aliquots (10 μL) of the indicated fractions were analyzed by SDS-PAGE and silver stained.

IN. The reactions were linear for at least 10 min at 37 °C. It was previously shown that 3' end processing and DNA strand transfer reactions show a distinctly nonlinear dependence on IN concentration in the approximately 0.08 μ M or lower range (Engelman & Craigie, 1992). Therefore, we avoided these lower protein concentrations in our analysis of folded IN. A typical result is shown in Figure 4. The ability of control IN to perform strand transfer and dinucleotide cleavage using the appropriate substrates was approximately 2-fold higher than for folded IN (Fig. 4A,B). However, the abilities of both enzyme preparations to perform the disintegration reaction using both substrates were equivalent (Fig. 4C,D). Similar results were obtained with 3 other folded HIV-1 IN preparations. The reaction for catalysis of the dumbbell substrate was not proportional to enzyme concentration (Fig. 4D), due either to higher salt concentration from enzyme addition (Vincent et al., 1993) or possibly to other unknown factors.

Subunit composition of folded HIV-1 IN preparations

The possibility existed that the standard folding conditions employed were conducive to folding of a specific domain, like the domain responsible for catalysis (which maps between residues 50 and 186 of HIV-1 IN) but maybe not for regions (NH₂- and COOH-termini) necessary for strand transfer (Fig. 1). In addition, the folding kinetics of multidomain proteins like IN with subsequent formation of specific dimers or multimers would be very complex (Garel, 1992). Using our standard folding conditions, we analyzed 5 different folded HIV-1 IN preparations by velocity sedimentation on glycerol gradients and obtained results similar to those shown in Figure 5. Expressed HIV-1 IN purified from bacteria appears to be dimeric upon velocity sedimentation analysis (Vincent et al., 1993).

Glycerol gradient sedimentation of folded HIV-1 IN revealed a series of integrase species sedimenting as monomers to presumed tetramers (Fig. 5A). Approximately 50% of IN sedimented as dimers or higher order aggregates. The major peak of IN sedimented near or with ovalbumin, suggesting that these species were monomeric. The sedimentation profiles of IN were similar between the NaCl concentrations of 0.5 and 1.0 M in the glycerol gradient buffers. SDS-PAGE analysis demonstrated that the enzyme was nearly homogeneous, but some IN was not full length (288 amino acids) (Fig. 5D). Western blot analysis using rabbit antisera directed against peptides specific to the amino- or carboxyl-terminus of IN revealed that approximately 70% of the molecules were slightly truncated only at their carboxyl-terminus (data not shown). The total quantity of the truncated carboxyl-terminal polypeptide varied between approximately 50 and 70% for various IN preparations.

We wanted to assess what catalytic activities were associated with the various IN species separated by size. The majority of 3' end processing (data not shown), strand transfer, and disintegration activities sedimented with apparent IN dimers or higher order oligomers (Fig. 5A,C). The ratios between strand transfer and disintegration activities with protein concentration appear relatively constant between fractions 6 and 18. Between fractions 18 and 21, both the strand transfer activity and the disintegration activity with the dumbbell substrate decreased to a minimal level (Fig. 5A). The disintegration activity with the Y substrate was still easily visualized in fraction 21 with only a

15-min incubation (Fig. 5A,C). Residual disintegration activity with the Y substrate was still detectable in fraction 26.

These data demonstrated that IN preparations folded under standard conditions consist of monomers and several multimers with different abilities to catalyze various enzymatic reactions.

Formation of enzymatically active monomers

We wanted to modify our folding conditions slightly to possibly minimize protein aggregation and to improve the catalytic activities of IN monomers. Because IN of various HIV strains are highly conserved at the amino acid level, we also wanted to test whether our folding technique for pNY IN could be applied to IN from a different virus strain. Besides the pNY clone (Figs. 1–5), we also examined another integrase derived from a different infectious clone of HIV-1 (HXB-2). From DNA sequence analysis, the HXB-2 clone has only 6 amino acid changes in comparison to the pNY clone.

To promote correct folding conditions, we chose a protein concentration for folding to minimize protein aggregation, which strongly affects renaturation (Garel, 1992). Two folding conditions were modified. First, the incubation time for denaturation of HXB-2 or pNY IN in the presence of 6 M Gdn·HCl–5 mM DTT at 22 °C was increased from 90 min to 220 min. Some proteins require longer periods of time in denaturants to permit complete unfolding. Second, the HXB-2 or pNY 6 M Gdn·HCl treated enzymes were cooled on ice for 4 min prior to the addition of ice-cold 1 M NaCl buffer. Early experiments suggested that higher temperatures may accelerate aggregation of IN. Table 1 shows that when HXB-2 IN was folded under these modified temperature conditions, the preparation had an equal or slightly higher specific activity for strand transfer than the control enzyme. The disintegration activities of both enzymes in Table 1 were also alike. Similar data were obtained with pNY IN when it was also folded under these modified temperature conditions (data not shown). These results suggest that the temperature used for the initial dilution step is critical for correct

Table 1. Specific activities of IN preparations containing active monomers^a

	μ M	% Conversion		Specific activity	
		Y substrate	19/21 substrate	Y substrate	19/21 substrate
Folded: HXB-2	0.109	4.9	8.5	45	78
	0.218	6.2	15.0	29	69
	0.328	10.1	19.5	31	59
Control: pNY	0.068	3.7	4.0	54	58
	0.136	4.9	4.9	36	36
	0.204	5.5	8.1	27	40

^a Quantitative assay conditions were used as described in the legend of Figure 4 except that the quantities of the Y and 19/21 substrates were increased to 20 and 50 pmol, respectively. The reaction time for the Y substrate was 5 min, and for the 19/21 substrate it was 15 min. The specific activities were determined by dividing the percent product formed by the micromolar concentration of IN in the reaction. The folded IN preparation was 8.6 μ M. Control IN was previously described (Fig. 4).

folding of either HXB-2 or pNY integrase and that this folding procedure can be applied to IN from various strains of HIV.

The HXB-2 integrase prepared under the modified temperature (0 °C) folding conditions was subjected to sedimentation analysis. We wanted to determine if IN monomers possessed increased strand transfer and disintegration activities upon assaying in comparison to those produced under standard temperature (22 °C) folding conditions (Fig. 5). As shown in Figure 6, HXB-2 IN sedimented mainly as a monomer, with some aggregate forms of IN still observed under these folding conditions. These modified temperature folding conditions enhanced the formation of HXB-2 IN monomers, which upon assaying were enzymatically active for both strand transfer and disintegration (Fig. 6). pNY IN monomers produced under these modified temperature folding conditions also had similar enzymatic activities in comparison to HXB-2 monomers (data not shown).

Discussion

We have developed a protein folding procedure for the multidomain HIV-1 IN. The enzyme from several strains of HIV-1 appears to have similar folding properties. The criterion for proper folding was the use of several functional probes to measure catalytic efficiencies. The folded enzyme was capable of 3' OH end processing, strand transfer, and disintegration using HIV-1 LTR probes. Physical interactions of IN monomers were

analyzed by velocity sedimentation of folded IN preparations on glycerol gradients.

Partial purification of HIV-1 IN from inclusion bodies by gel filtration in the presence of 6 M Gdn·HCl was achieved. Large-scale preparation of IN should be easily obtainable by this procedure. The $(\text{NH}_4)_2\text{SO}_4$ precipitation of Gdn·HCl purified IN provides a convenient storage medium for study of the folding of IN. The precipitate was stable on ice for at least 1 year. The reversible dissolution of the precipitate by 6 M Gdn·HCl-5 mM DTT appears to play a role in the complete unfolding of IN and subsequent folding. Initial studies reported here suggest that the temperature of the folding reaction may determine whether enzymatically active monomers of IN are obtained. Further studies will be needed to clarify other folding parameters and solution conditions that may influence the catalytic and noncatalytic functions of IN.

An important noncatalytic function of IN may be multimerization. In solution, IN purified from avian myeloblastosis virus (Grandgenett et al., 1978) or from bacterially expressed Rous sarcoma virus (RSV) IN (Jones et al., 1992) and HIV-1 IN is dimeric (Sherman & Fyfe, 1990). In vitro complementation of IN proteins mutated in different domains suggests that the catalytic domain mapping between amino acids 50 and 186 of HIV-1 IN is also important for dimerization. The folding of HIV-1 IN under standard conditions resulted in the formation of aggregates whose physical sizes ranged from potential tetramers to monomers (Fig. 5). We do not know if the dimerization domain (Fig. 1) is responsible for forming folded IN dimers or what other regions of IN are involved in forming the larger multimers.

For the larger folded aggregates, the relative specific activity for all the enzymatic functions appears higher than for monomers, suggesting that subunit composition affects catalytic function or that these aggregates are more stable upon sedimentation or enzymatic analysis (Fig. 5). These enzymatically active aggregates are most likely formed by nonspecific interactions between IN monomers possessing 1 or more functional domains that are folded properly. In other words, properly folded domains of IN may be acting in *trans* or *cis* relationships to enhance catalysis of other integrase molecules lacking a functional domain besides the catalytic core (Fig. 1). Complementation can occur between IN subunits possessing null mutations or deletions (Engelman et al., 1993; Vink et al., 1993). The NH_2 -terminus of HIV-1 IN acts *trans* to the catalytic core, and the COOH -terminus can function either *cis* or *trans* to the catalytic core.

Our data suggest that folded monomers are functional and probably form multimers upon incubation under enzymatic reaction conditions; dimerization of IN may be necessary for catalytic activity. RSV IN acts as a multimer in assays for LTR dinucleotide cleavage and strand transfer (Jones et al., 1992), and in vitro complementation among HIV-1 IN mutant proteins strongly suggests that these enzymatic reactions require multimers of IN (Engelman et al., 1993; van Gent et al., 1993b). Aggregation of HIV-1 IN also appears to occur upon incubation with LTR or nonspecific DNA under low-salt assay conditions (van Gent et al., 1991). Further analysis is necessary to determine what mechanisms are involved in IN multimerization both in solution and in the presence of DNA.

The simplest model for folding of the multidomain HIV-1 IN may involve independent folding of each functional domain (Garel, 1992). The catalytically active center of folded IN (Fig. 1) obtained from both folding procedures appears to be somewhat

Fig. 6. Isolation of enzymatically active HXB-2 IN monomers. Conditions for sedimentation are described in the legend of Figure 5. Modified temperature folding conditions were used to produce IN. For HXB-2 IN, a total of 140 µg was subjected to sedimentation analysis. Aliquots (5 µL) were assayed using the recessed U5 substrate or the Y substrate; the reaction times were 1 h and 15 min, respectively. A fraction of each reaction was analyzed on a DNA sequencing gel, and the reaction products are expressed as pixels. The protein concentration of each fraction was determined by silver staining as described in the legend of Figure 5. The inset is a silver-stained gel of various IN fractions (8-µL aliquot).

equivalent as measured by specific activity, suggesting that this domain with intact IN is readily renatured under various folding conditions. Other subtle protein folding events may also greatly influence the 3' end processing and strand transfer catalytic activities of IN that require properly folded NH₂- and COOH-termini.

It must be emphasized that folded HIV-1 IN, although catalytically active in the presence of Mn²⁺, is not capable of performing the disintegration or half-site strand transfer reactions in the presence of Mg²⁺ (data not shown). HIV-1 IN in the context of a 80S preintegration complex isolated from virus-infected cells can efficiently perform LTR full-site insertion of viral DNA ends in the presence of Mg²⁺ (Farnet & Haseltine, 1990). Further investigation into the folding properties of IN are necessary to produce an enzyme comparable to virion-associated IN.

Materials and methods

Expression of HIV-1 IN

The integrase, derived from the HIV-1 pNY clone, was expressed in *E. coli* BL21(DE3) pLysS cells using a T7 expression vector. The T7 expression plasmid (pINT7) for expressing pNY IN was a gift from F. Bushman. Purified "control" IN from pINT7 has been characterized and was supplied by A. Engelman. The IN gene derived from an infectious HXB-2 clone was placed into another T7 expression vector in our laboratory and was used in the latter studies in this report.

Purification of IN

Following isopropylthiogalactopyranoside induction of cells expressing IN, the cells were harvested by centrifugation and washed twice with 20% (w/v) sucrose, 1 mM EDTA, and 20 mM Tris-HCl (pH 7.5) (Lin et al., 1990). Cells were frozen at -70 °C. Purification of the insoluble fraction from 2 L of cells was modified (Lin et al., 1990). The following modifications increased the purity of IN in the final pellet as analyzed by SDS-10% polyacrylamide gel electrophoresis. The frozen cells were rapidly lysed by adding 50 mL of an ice-cold buffer, termed buffer S, that contained 0.1% Triton X-100, 2 mM phenylmethylsulfonyl, 2 mM EDTA, aprotinin (20 µg/mL), and leupeptin (1 µg/mL). All subsequent steps prior to the unfolding experiments using Gdn·HCl were on ice except where indicated. After 20 min of shaking to promote cell lysis, the suspension was centrifuged at 23,600 × *g* for 1 h. The pellet was dissolved in 10 mL of buffer P (phosphate-buffered saline containing buffer S ingredients without Triton X-100) and dounced 20 times. The sample was divided into 4 equal aliquots that were sonicated 3 times (30 s each) at 60 W with intermediate ice storage between each sonication. The aliquots were pooled into 2 equal fractions that were processed separately to increase the efficiency of the following wash procedure. RNase T1 (5,700 units) was added and incubated at 22 °C for 15 min. The samples were passed through a 20-gauge needle 4 times prior to adding 40 mL of buffer W (phosphate-buffered saline with 25% [w/v] sucrose, 1% Triton X-100, and the protease inhibitors). Following centrifugation of the diluted samples at 23,600 × *g* for 20 min, the pellets were washed with buffer W 3 more times by passage through the needle and centrifugation. At the final wash step,

the 2 fractions were pooled into a single tube prior to centrifugation. The pellet was frozen in a dry-ice bath and stored at -70 °C or processed as described below.

The pellet was dissolved in 1.5 mL of 6 M Gdn·HCl, 5 mM DTT, 30 mM sodium phosphate (pH 7.0), and 1 mM EDTA (adjusted to pH 7.3 with 1 N NaOH). The Gdn·HCl treated sample was incubated at 22 °C for 90 min and then clarified at 23,600 × *g* for 20 min. The absorbance at 280 nm was approximately 37 OD/mL. The sample was adjusted to 5% (w/v) sucrose prior to chromatography on a 1.2-cm × 85-cm-long column containing Sephacryl-S200 HR (Pharmacia) and equilibrated with the above Gdn·HCl buffer with 2 mM DTT present. The flow rate was 0.4 mL/min. Fractions containing IN were identified by absorbance at 280 nm and SDS-PAGE. The appropriate fractions were pooled and dialyzed against 2.65 M (NH₄)₂SO₄, 5 mM Tris-HCl (pH 7.6), 1 mM DTT (adjusted with ammonium hydroxide to pH 7.3) at 4 °C. The (NH₄)₂SO₄ precipitate was stored on ice.

Folding of HIV-1 IN

The desired quantity of IN as an (NH₄)₂SO₄ precipitate was centrifuged at 12,000 × *g* for 5 min in a 1.5-mL microfuge tube at 4 °C. The supernatant was removed, and the pellet was dissolved in 6 M Gdn·HCl-5 mM DTT (buffered at pH 7.3) and incubated at 22 °C for 90 min. We wanted the folding reaction to take place at a lower pH than the pI, which was calculated to be 8.2. Routinely, 60 µL of 6 M Gdn·HCl treated sample was rapidly diluted 6-fold with ice-cold buffer containing 1 M NaCl, 30 mM Tris-HCl (pH 7.6), and 1 mM DTT. The diluted sample was incubated on ice for 3 h with occasional gentle shaking. Dialysis of the 1 M Gdn·HCl sample was accomplished in an upside-down microfuge tube with a dialysis membrane between the sample and the cap containing a 7-mm hole (Falson, 1992). The dialysis medium (2 L) was the 1 M NaCl buffer. Dialysis with stirring was for 17 h at 2 °C. The dialyzed samples were clarified by centrifugation at 20,000 × *g* for 20 min prior to freezing or further analysis.

All of the samples were routinely scanned between 240 and 360 nm with a Beckman DU7400 spectrophotometer. The absorbance of the solution at 340 nm was monitored to assess the degree of aggregation of IN (Baker et al., 1992). Both the denatured and folded protein had maximum absorbance between 278 and 280 nm. The enzyme was frozen in a dry ice-ethanol bath and stored at -70 °C. The concentrations of the folded protein preparations were also determined by the Bradford protein assay (Bio-Rad) using bovine serum albumin (ICN) as a standard (Bradford, 1976).

Measuring 3' OH end processing, strand transfer, and disintegration activities

Similar assay conditions for 3' end processing and strand transfer were previously described (Engelman & Craigie, 1992). Briefly, the reaction mixture (40 µL) contained 15 mM Tris-HCl (pH 7.5), 2 mM DTT, 1.5% glycerol, and 7.0 mM MnCl₂. The standard reaction times were between 30 and 60 min at 37 °C. For quantitation studies, the reaction times were reduced to between 5 and 15 min. Double-stranded oligonucleotides containing U5 end sequences were used as substrates. The substrates

were prepared by 5' end labeling of the processed LTR strand using T4 polynucleotide kinase (Life Technologies, Inc.) and [γ - 32 P]ATP (3,000 Ci/mmol; Dupont). For 3' end processing, the labeled 20-mer oligonucleotide was annealed with an equal quantity of complementary 20-mer oligonucleotide. For strand transfer alone, a labeled 19-mer oligonucleotide was annealed with an equal quantity of a 21-mer oligonucleotide, producing a 3' OH recessed substrate. All annealed substrates were subjected to analytical gel electrophoresis on native gels to ensure complete annealing. Sometimes, depending upon the purity of the labeled oligonucleotide, the oligonucleotide was purified on a 12% denaturing polyacrylamide gel prior to use. In most assays, 5–20 pmol of the annealed substrates was used, depending upon assay conditions. Aliquots (usually 10%) of each reaction were subjected to electrophoresis through a 12% polyacrylamide–8 M urea gel that was subsequently transferred to Whatman paper and dried. The reaction products were visualized by autoradiography with or without an intensifying screen. For quantitation, the dried gel was analyzed by a Molecular Dynamics Phosphorimager. To ensure that the Phosphorimager was in the linear range of detection, the instrument was standardized using various quantities of a labeled oligonucleotide and varying times of exposure.

The disintegration assay conditions were the same as described for 3' end processing. Both the 38-mer dumbbell substrate and the hybrid dumbbell or Y substrate were previously described (Vincent et al., 1993). The 38-mer substrate was 5' end labeled and purified on a 12% denaturing gel prior to use. For the Y substrate, a 5' end labeled 16-mer target sequence was annealed to both a viral and a target oligonucleotide as illustrated in Figure 3C. The annealed Y substrate was purified on a native polyacrylamide gel. The major and slowest migrating labeled substrate was excised and eluted from the gel. The sample was diluted and concentrated by several cycles of centrifugation using an Amicon Centricon 3 microconcentrator. All of the above oligonucleotides were synthesized and purified by HPLC (Operon Technologies, Inc.).

Glycerol gradient centrifugation

The subunit composition of folded HIV-1 IN preparations was analyzed by glycerol gradient centrifugation (Grandgenett et al., 1978). The samples were placed onto 5–20% glycerol gradients (4.9 mL) in a buffer containing 30 mM Tris-HCl (pH 7.6), 0.5 mM DTT, and NaCl that was varied in individual experiments between 0.5 and 1 M. The sample loading volume was 200 μ L for IN or for the molecular weight standards. Centrifugation was at 2 °C for 21 h at 47,000 rpm in an SW 50.1 rotor (Beckman). Two drops were collected from a 20-gauge needle that was inserted into the bottom of the tube. Aliquots were assayed for various enzymatic activities and by SDS-PAGE as described in the legend of Figure 5. Avian myeloblastosis virus $\alpha\beta$ DNA polymerase (Lin et al., 1991), bovine serum albumin, and ovalbumin were used as molecular weight markers.

Preparation of peptide-directed antisera and Western blot analysis

Rabbit antibodies directed against synthetic peptides corresponding to 6 separate regions of HIV-1 IN were produced

(Grandgenett et al., 1983). From the amino-terminus, the peptides used included amino acid residues 1–16, 23–34, 142–153, 179–190, 234–245, and 276–288 of HIV-1 (Ratner et al., 1985). To varying degrees, all 6 peptides produced antisera reactive to *E. coli* expressed HIV-1 IN by Western blot analysis or by immunoprecipitation of IN from 35 S-labeled HIV-1 (data not shown).

Acknowledgments

This work was supported by a grant from the National Institutes of Health (AI-31334). D.G. thanks Dr. James Lee (The University of Texas Medical Branch, Galveston) and Tony Fadell for providing support for the initial stages of this work, and particularly Xiodong Cheng for his advice on the folding experiments. D.G. also thanks A.C. Vora and rotating laboratory students D. Swope, M. McCord, and M. Brother for technical assistance, and L. Greeno for typing the manuscript.

References

- Baker D, Sohl JL, Agard DA. 1992. A protein-folding reaction under kinetic control. *Nature* 356:263–265.
- Bradford NM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254.
- Burke CJ, Sanyal G, Bruner MW, Ryan JA, La Femina RL, Robbins HL, Zeft AS, Middaugh CR, Cordingley MG. 1992. Structural implications of spectroscopic characterization of a putative zinc finger peptide from HIV-1 integrase. *J Biol Chem* 267:9639–9644.
- Bushman FD, Craigie R. 1991. Activities of human immunodeficiency virus (HIV) integration protein in vitro: Specific cleavage and integration of HIV DNA. *Proc Natl Acad Sci USA* 88:1339–1343.
- Bushman FD, Engelman A, Palmer I, Wingfield P, Craigie R. 1993. Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *Proc Natl Acad Sci USA* 90:3428–3432.
- Bushman FD, Fujiwara T, Craigie R. 1990. Retroviral DNA integration directed by HIV integration protein in vitro. *Science* 249:1555–1558.
- Chow SA, Vincent KA, Ellison V, Brown PO. 1992. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science* 255:723–725.
- Drelich M, Wilhelm R, Mous J. 1992. Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro. *Virology* 188:459–468.
- Engelman A, Bushman FD, Craigie R. 1993. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. *EMBO J* 12:3269–3275.
- Engelman A, Craigie R. 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J Virol* 66:6361–6369.
- Falson P. 1992. An efficient procedure to dialyze volumes in the range of 10–200 μ L. *Biofeedback* 13:118–120.
- Farnet CM, Haseltine WA. 1990. Integration of human immunodeficiency virus type 1 DNA in vitro. *Proc Natl Acad Sci USA* 87:4164–4168.
- Fitzgerald ML, Vora AC, Grandgenett DP. 1992. Concerted integration of viral DNA termini by purified avian myeloblastosis virus integrase. *J Virol* 66:6257–6263.
- Garel JR. 1992. Folding of large proteins: Multidomain and multisubunit proteins. In: Creighton TE, ed. *Protein folding*. New York: W.H. Freeman and Co. pp 405–454.
- Gill SC, von Hippel PH. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182:319–326.
- Grandgenett DP, Knaus RJ, Hippenmeyer PJ. 1983. Antibodies against a synthetic peptide of the avian retrovirus pp 32 protein and the β DNA polymerase subunit. *Virology* 130:257–262.
- Grandgenett DP, Vora AC, Schiff RD. 1978. A 32,000-dalton nucleic acid-binding protein from avian retrovirus cores possesses DNA endonuclease activity. *Virology* 89:119–132.
- Jones KS, Coleman J, Merkel GW, Laue TM, Skalka AM. 1992. Retroviral integrase functions as a multimer and can turn over catalytically. *J Biol Chem* 267:16037–16040.
- Kahn E, Mack JGP, Katz RA, Kulkosky J, Skalka AM. 1991. Retrovirus

- integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res* 19:851-860.
- Katz RA, Merkel G, Kulkosky J, Leis J, Skalka AM. 1990. The avian retrovirus IN protein is both necessary and sufficient for integrative recombination. *Cell* 63:87-95.
- Kulkosky J, Jones KS, Katz RA, Mack JPG, Skalka AM. 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol Cell Biol* 12:2331-2338.
- Lewis P, Hensel M, Emerman M. 1992. Human immunodeficiency virus infection of cells arrested in the cell cycle. *EMBO J* 11:3053-3058.
- Lin KH, Fukuda T, Cheng SY. 1990. Hormone and DNA binding activity of a purified human thyroid hormone nuclear receptor expressed in *Escherichia coli*. *J Biol Chem* 265:5161-5165.
- Lin TH, Grandgenett DP. 1991. Retrovirus integrase: Identification of a potential leucine zipper motif. *Protein Eng* 4:435-441.
- Lin TH, Quinn T, Walsh M, Grandgenett DP, Lee J. 1991. Avian myeloblastosis virus reverse transcriptase. *J Biol Chem* 266:1635-1640.
- Mizuuchi K. 1992. Polynucleotidyl transfer reactions in transpositional DNA recombination. *J Biol Chem* 267:21273-21276.
- Mumm SR, Grandgenett DP. 1991. Defining nucleic acid-binding properties of avian retrovirus integrase by deletion analysis. *J Virol* 65:1160-1167.
- Ratner L, Haseltine W, Patarca T, Livak KJ, Starcich B, Josephs SF, Doran ER, Rafalski JA, Whitehorn EA, Baumeister K, Ivanhoff L, Pette-way SR Jr, Pearson ML, Lanténberger JA, Papas TS, Ghrayeb J, Chang NT, Gallo RC, Wong-Staal F. 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* 313:277-284.
- Sherman PA, Fyfe JA. 1990. Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleaving activity. *Proc Natl Acad Sci USA* 87:5119-5123.
- van Gent DC, Elgersma Y, Bolk MWJ, Vink C, Plasterk RHA. 1991. DNA binding properties of the integrase proteins of human immunodeficiency viruses types 1 and 2. *Nucleic Acids Res* 19:3821-3827.
- van Gent DC, Groeneger AAMO, Plasterk RHA. 1993a. Identification of amino acids in HIV-2 integrase involved in site-specific hydrolysis and alcoholysis of viral DNA termini. *Nucleic Acids Res* 21:3373-3377.
- van Gent DC, Vink C, Groeneger AAMO, Plasterk RHA. 1993b. Complementation between HIV integrase proteins mutated in different domains. *EMBO J* 12:3261-3267.
- Vincent KA, Ellison V, Chow SA, Brown PO. 1993. Characterization of human immunodeficiency virus type 1 integrase expressed in *Escherichia coli* and analysis of variants with amino-terminal mutations. *J Virol* 67:425-437.
- Vink C, Groeneger AAMO, Plasterk RHA. 1993. Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type 1 integrase protein. *Nucleic Acids Res* 21:1419-1425.
- Vink C, van Gent DC, Elgersma Y, Plasterk RHA. 1991a. Human immunodeficiency virus integrase protein requires a subterminal position of its viral DNA recognition sequence for efficient cleavage. *J Virol* 65:4636-4644.
- Vink C, Yeheskiely E, van de Marel GA, van Boom JH, Plasterk RHA. 1991b. Site-specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini mediated by the viral integrase protein. *Nucleic Acids Res* 19:6691-6698.
- Woerner AM, Klutch M, Levin JG, Marcus-Sekura CJ. 1992. Localization of DNA binding activity of HIV-1 integrase to the C-terminal half of the protein. *AIDS Res Hum Retroviruses* 8:297-304.