AD_____

(2)

CLONING AND BIOCHEMICAL CHARACTERIZATION OF HIV INTEGRASE

FINAL REPORT

ELLEN MURPHY

JULY 1, 1991

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21702-5012

Contract No. DAMD17-88-C-8126

The Public Health Research Institute
455 First Avenue
New York, NY 10016

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT: Approved for public release; distribution unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: The Public Health Research Institute; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code): 455 First Avenue, New York, NY 10016	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION: U.S. Army Medical Research & Development Command; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: DAMD17-88-C-8126
8c. ADDRESS (City, State, and ZIP Code): Fort Detrick, Frederick, Maryland 21702-5012	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
63105A	3M2-63105DH29	AC	WUDA314849

11. TITLE (Include Security Classification): (U) Cloning and Biochemical Characterization of the HIV Integrase

12. PERSONAL AUTHOR(S): Ellen Murphy

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 6/16/88 TO 6/15/91	1991 July 1	18

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	RA1, Human immunodeficiency virus, HIV AIDS, Integration, Integrase Viral DNA, LTR
06	03		
06	13		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Our long range goal is to understand the macromolecular interactions responsble for the integration of HIV DNA into human chromosomes. The primary goal of this project was to develop biochemical assays for the interaction of the HIV-1 integrase protein with its DNA target, the viral LTRs. The gene encoding IN was subcloned from infectious viral DNA and was expressed in both *E. coli* and in insect cells. The lack of any additional carboxy-terminal processing of IN was demonstrated. Several different solubilization procedures were developed for the purification of integrase. DNA substrates for IN have been constructed; these include synthetic oligonucleotide substrates corresponding to the LTRs and a circular substrate containing the ligated junction of the two LTRs, obtained by polymerase chain amplification from HIV-infected cells. Recombinant protein purified from *E. coli* is active in a trimming assay in which 2 base pairs are removed from the 3' end of an oligonucleotide substrate. Integration of one labeled substrate into another was also observed. *E. coli* strains expressing integrase were found to be resistant to infection by single-stranded DNA and RNA bacteriophages. This will be exploited in developing assays for screening inhibitors of integrase and for selection of integrase mutants. Ribozymes that prevent the expression of integrase in *E. coli* have been developed, supporting the idea that ribozymes have therapeutic potential.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Mary Frances Bostian	301-663-7325	SGRD-RMI-S

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

_____ Where copyrighted material is quoted, permission has been obtained to use such material.

_____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

_____ Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

_____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

_____ For the protection of human subjects, the investigator(s) have adhered to policies of applicable Federal Law 45CFR46.

EM In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

Ellen Murphy 7/1/91

PI Signature Date

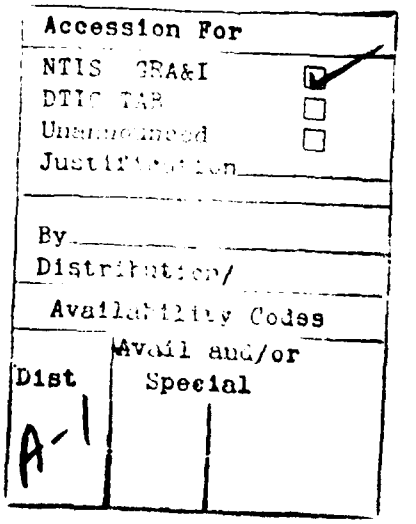

TABLE OF CONTENTS

FIGURES

TABLE

INTRODUCTION

Our long range goal is to understand the macromolecular interactions responsible for the integration of HIV DNA into human chromosomes with the specific purpose of identifying compounds that block integration. Since integration is an early step in the retroviral life cycle, blocking integration prevents cells from becoming infected. Moreover, integration is not a part of normal cell physiology; consequently, compounds that block integration should be good candidates as therapeutic agents for AIDS. What follows in this section is a brief review of relevant biochemical aspects of retroviral integration and an outline of our experimental rationale.

Integration of viral DNA into the host chromosome is both an essential and unique step of the retroviral life cycle. Since integration is an early step, understanding the integration event may lead to the development of inhibitors that can block it and prevent cells from becoming infected. Moreover, since integration is not a part of normal cell physiology, compounds that block integration should be good candidates as therapeutic agents for AIDS. In combination with agents that inhibit reverse transcriptase (RT), drugs inhibiting integration should make it possible not only to inhibit viral replication to a much greater extent than is now possible but also to lower the therapeutic doses of both inhibitors. Since the commonly used RT inhibitors have significant toxicity, this could lead to a substantial improvement in AIDS therapy.

Genetic analyses have identified three regions of the viral genome required for integration (reviewed in [36]). One is the 3′ end of *pol*, which encodes the viral integrase (9, 21, 25, 29). The other two regions are the viral targets of integrase, the ends of viral DNA (5, 6, 20). Integrase carries out a two-step reaction, first trimming two base pairs from the 3′ ends of the viral DNA and then inserting this trimmed viral genome into the host chromosome (2, 13, 14, 27). Both reactions appear to take place while the viral DNA is part of a protein-DNA complex commonly called a core (1, 11). This complex is stable and can be purified from infected cells by gel filtration or sucrose gradient centrifugation. After extensive purification, integrase is the only viral protein still associated with the viral DNA; the purified complex is integration-competent (12). The trimming and integrating activities of integrase are also detected using exogenous DNA and partially purified preparations of avian (17), murine (8), or HIV-1 integrase (3, 4, 30).

Following reverse transcription, infected cells contain three forms of viral DNA, one linear form, containing a complete LTR at each end, and two circular forms, containing one LTR or two directly repeated, tandem LTRs. Thus the topological form of the viral DNA substrate for integration was not initially obvious. Early *in vivo* studies with an avian virus had suggested that the 2-LTR circle was the precursor for integration (22). However, it is now clear that this is not a general phenomenon (10, 19) and that integrase carries out its reaction with a linear DNA substrate (2, 13, 14, 27). The first step in the reaction is the removal of two base pairs at the 3′ ends of the LTRs to produce a recombinogenic intermediate with a 2 bp 5′ overhang, CA, a sequence highly conserved among all retroviruses (35). The 3′ hydroxyl ends of the LTRs are then transferred to the 5′ ends of the target DNA (13). Cleavage of the target, as for procaryotic transposable elements, involves a pair of staggered nicks that are eventually repaired to generate a target duplication whose length is characteristic of the element; for HIV-1 it is 5 bp (38). A low-specificity endonucleolytic activity of integrase (30) could be responsible for making the required staggered cuts in the host DNA.

The recombination reaction catalyzed by retroviral integrase resembles that of bacteriophage Mu (13). The polarity of the initial strand transfer is the same for both reactions, as is the lack of a requirement for an external energy source or for superhelicity of one of the recombining partners. For a number of other site-specific recombination reactions, the energy derived from cleavage of the element's ends is conserved via a transient covalent protein-DNA intermediate. For Mu, as for integrase, cleavage of the viral ends can be completely uncoupled from the strand transfer reaction; both reactions will occur with pre-cleaved donor molecules (1, 7, 13). Thus, the energy required for formation of the new phosphodiester bonds cannot be derived from cleavage of the viral ends, and has therefore been suggested to be derived from cleavage of the target DNA (13). This could proceed via a covalent intermediate between the protein and the target DNA; this question has not yet been settled.

We have focused on developing *in vitro* assays for integrase that would circumvent the problems inherent in handling large amounts of infectious material. To this end, we have developed vectors to overexpress the IN protein and a DNA substrate, as described in our midterm report. We have succeeded in purifying IN protein from these vectors and in establishing *in vitro* trimming and integration assays, as discussed below. Recently, we have begun work on a bacterial assay for integrase that we believe holds great promise for screening of inhibitors and for selection of integrase mutants. As an outgrowth of this work we also initiated studies on anti-integrase ribozymes, RNA molecules that seek and specifically destroy other RNA molecules. Our work showed that intracellular ribozymes are very effective in bacteria. This opens the way to using bacterial systems to refine anti-HIV ribozymes.

BODY

A. Absence of C-terminal processing of HIV-1 integrase. We showed that, unlike Rous sarcoma virus, HIV-1 integrase is not processed at its carboxy terminus *in vivo*. This result is important in designing cloning strategies in which the goal is to produce bonafide integrase. These results were detailed in the midterm report.

B. Expression of HIV-1 integrase in baculovivus vectors. In preparation for biochemical studies of integrase we have been working with two expression systems, one in insect cells and one in *Escherichia coli*. Two expression strategies were used. In one, an artificial AUG was used to initiate translation of integrase, placing an extra methionine at the amino terminus of the protein. In the other, the natural form of integrase was processed in insect cells from the *pol* polyprotein by HIV-1 protease. In the midterm report we described in detail the constructions and the expression of integrase from both constructs. The major conclusions were 1) the *int* clone, vMPD50, produced much greater amounts of integrase than did the *pol* clone, vMPD32, and 2) cells infected with the *pol* construct lose viability at an earlier time than those infected with either *int* or the wild-type baculovirus, as determined by trypan blue permeability at various times after infection. Thus, expression of HIV-1 protease or reverse transcriptase or both may reduce the viability of these cells.

C. N-terminal sequencing. We have confirmed the sequence of the first 14 residues of the integrase produced by the *int* clone by N-terminal sequencing; confirmation of the N-terminal sequence of the integrase produced by the *pol* clone has not been possible

because of the much lower level of protein production by this clone.

D. Expression of recombinant integrase from E. coli. We are also expressing integrase in *E. coli* from a clone, pLJS10, obtained from J. Groarke at Sterling Research Group. Integrase is produced by this clone under control of a T7 promoter; it was constructed by inserting a PCR fragment corresponding to *int* (nt 4230–5097) into the T7 promoter vector pET8c (26). The coding sequence contains two additional amino acids, Met-Gly, at the N-terminus; N-terminal sequencing indicates that the methionine residue is removed (J. Groake, personal communication). A time course of induction of integrase expression from pLJS10 is shown in Fig. 1 (lanes 1–3). Also shown is the effect of expression of T7 lysozyme, which inhibits T7 RNA polymerase (33), on the level of integrase expression (Fig. 1, lanes 4–8).

Integrase is currently obtained from *E. coli* by preparation of insoluble inclusion bodies, which are solublized with 4M urea in high salt buffer. This is followed by FPLC using butyl sepharose. The resulting protein is about 90% pure, as judged by Coomassie staining of polyacrylamide gels, and is substantially free of the contaminating nucleases that prevent the protein from being assayed at an earlier stage.

E. Solubilization of integrase. HIV-1 integrase has a strong tendency to aggregate, whether expressed in a bacterial (15, 30) or baculovirus (3) system. Thus, most of the integrase expressed from vMPD50 is found in the nuclear pellet following low-speed centrifugation (850 x g for 5 min). Sonic disruption to fragment nuclei in the cell lysate fails to prevent rapid sedimentation of integrase, indicating that integrase is part of a large aggregate and is not simply a soluble protein trapped inside nuclei.

Although 4M urea has been reported to solubilize baculovirus-expressed integrase (4, 8); in our hands it had a marginal effect on solubilization. Higher concentrations (up to 6 M) were more effective. Once integrase has been solubilized, most of it remains soluble upon dilution into 3.5 M urea, but this fraction gradually decreases as the urea concentration is lowered. At concentrations below 1.0 M urea, little integrase remains in solution.

Others (30) have reported solubilization of integrase expressed in *E. coli* using high concentrations of NaCl. We found that 0.8 M NaCl solubilized baculovirus-expressed integrase, but only in the presence of both deoxycholate and dithiothreitol. In this situation integrase rapidly precipitated following dilution or dialysis. We also found that integrase could be solubilized by treatment with 0.5 M $NaHCO_3/Na_2CO_3$ at pH 10. This method requires dithiothreitol but not deoxycholate.

F. Analysis of HIV-1 circle junction sequences from infected with cells: substrates for integration. When we began this work, the circular form of the viral DNA was still believed to be the precursor for integration. In addition, neither the precise sequence of the circle junction nor the number of bp removed during integration were known for certain, because HIV-1 differs from all other retroviruses in the placement of the minus strand tRNA primer relative to the end of the LTR (1 bp away rather than the usual 2 bp [32]). We (16) and others (18, 23, 31, 39) have now demonstrated, using PCR amplification of HIV-1-infected cells, that HIV-1 viral DNA contains 2 bp at each end that are absent from the proviral sequence. Our results differ from some of the other reports in that we detected a high proportion of larger deletion and insertion mutations at the circle junction. This was due to the fact that we did not pre-select our clones for size, and it indicates that a significant proportion, about 50%, of unintegrated circular HIV-1 DNA contains mutant circle junction sequences. A reprint describing this work is attached (Hong *et al.*, Appendix 2).

From this work, we also developed DNA substrates, both linear and circular, that can be used in assaying biochemical functions of integrase. The wild-type circle junction sequence contains a ScaI site precisely at the circle junction; thus cleavage of circle junction-containing plasmids with ScaI generates a linear minivirus that carries LTR sequences at both ends. Small fragments containing only U5 or U3 sequences can also be obtained from these clones.

G. 3′ trimming activity of recombinant integrase. Integrase prepared from *E. coli* carrying pLJS10, solublized by treatment with 4M urea and purified by butyl sepharose chromatography, was assayed for 3′ trimming activity as shown in Fig. 2. (4, 8, 17, 30), using a 5′–end labeled, double-stranded oligonucleotide substrate corresponding to the end of U5. 3′ trimming of 2 nucleotides is clearly seen in lane 2. In addition, a strand transfer reaction, representing integration of one oligonucleotide substrate into another, occurs, resulting in reaction products of higher molecular weight than the oligonucleotide substrate (lane 2). The reaction is dependent upon $MnCl_2$ (lane 3) and $MgCl_2$ cannot substitute (not shown). Integration also occurs with a pre-trimmed substrate (not shown). The appearance of several products of discrete length suggests that there are preferred sites for integration within the oligonucleotide; this specificity is under investigation.

H. In vivo integration assay. We are developing an *in vivo* integration assay similar to those described by Brown *et al.* (1) and Craigie *et al.* (3). In the published assays, integration of a linear minivirus molecule into lambda DNA is detected by virtue of a selectable marker on the minivirus. Following the *in vitro* integration reaction, the lambda DNA is packaged using packaging extracts, and it is then transfected into *E. coli*. To omit the costly and difficult lambda packaging step, we are developing a system that uses instead a plasmid target that cannot replicate in *E. coli* unless the integration reaction occurs. The integration substrate carries a Km^r marker, selectable in *E. coli*, adjacent to a fragment containing the circle junction. These DNA elements are cloned to a plasmid of *Staphylococcus aureus*, pE12 (24), that cannot replicate in *E. coli*. This plasmid was constructed by subcloning a 243 bp KpnI–HindIII fragment containing the circle junction to pBluescript KSII+, cloning a kanamycin resistance gene (Pharmacia GenBlock) adjacent to the circle junction sequences at the EcoRI site, and then subcloning a 1.9 kb PvuII fragment containing the Km^r gene, the circle junction sequences, and 223 bp derived from pBluescript, to the unique AvrII site in pE12 (Fig. 3).

We have constructed two plasmids. One contains the wild-type circle junction fragment, which has a unique ScaI site precisely at the LTR-LTR boundary. Cleavage by ScaI thus produces a blunt-ended substrate with the correct ends. In the second substrate the ScaI site has been replaced by a unique NdeI site; cleavage by NdeI produces a linear DNA with a 2-bp, 3′ overhang (i.e. a pre-trimmed substrate). This substrate was described by Craigie *et al.* (3). Although the 2 bp overhang is of the correct polarity, the nucleotide sequence is not exactly correct, which may account for the aberrant integration junctions sometimes observed with this substrate (3). The NdeI substrate should be useful for distinguishing integrase mutants which may be unable to trim the 3′ end of the substrate but which may retain integration activity with a pre-trimmed substrate. We have constructed both donor molecules (ScaI and NdeI varieties) consisting of the Km^r cassette and the circle junction cloned into pE12. The target molecule is a 6.2 pGEM1 clone containing an insert of part of the MuLV *env* gene. Kanamycin resistant colonies arising after the *in vitro* integration reaction and transformation into *E. coli* will contain integrants of the pE12 con-

struct into the target. Tests of this bioassay for integrase activity are now underway.

I. Inhibition of bacteriophage growth by expression of HIV-1 integrase in E. coli. We have discovered that expression of integrase under control of the T7 promoter inhibits growth of the single stranded DNA and RNA viruses M13 and R17, respectively. The effect is proportional to the level of integrase expression. The basal level of integrase produced by an uninduced culture of *E. coli* HMS174 F′ (DE3) carrying pLJS10 was sufficient to totally inhibit plaque formation by M13 (>10^{10}fold reduction in PFU). When the level of constitutive integrase expression was reduced by including in the host the T7-lysozyme-producing plasmid pLysS to inhibit T7 RNA polymerase, the number of plaques was normal (3–8×10^{11}/ml). However, the plaques were unusually turbid. Even lower levels of integrase were produced in a host carrying pLysE from which higher levels of lysozyme are produced. In this case normal numbers of clear plaques were formed. Inducing integrase expression in the pLysE strain with 1 mM IPTG resulted in the formation of tiny, turbid plaques at 50-fold lower frequency. The results of one experiment are summarized in Table 1; integrase production by these cells is shown in Fig. 1, lanes 9–14. Similar results were obtained with the single-stranded, male-specific RNA phages R17 and Qβ, while the double-stranded DNA phages P1 and T4 were not inhibited. Two common features of the bacteriophages inhibited by integrase expression are that their genomes are single stranded nucleic acids and that they depend on F pili for adsorption. Because the proteins of the filamentous and RNA bacteriophages are so different, the mechanism of inhibition is unlikely to involve direct protein-protein interactions between integrase and bacteriophage proteins. Rather, an indirect effect on bacteriophage replication mediated by binding to the phage genome, or an effect on the host's F pili, would be more likely explanations.

We have attempted to rule out the second possibility by analyzing the conjugal ability of *E. coli* strains expressing integrase. Two integrase-producing strains carrying different F′ factors were tested in crosses with F⁻ strains; both had normal transfer frequencies. Generally, mutations in *traA* (the gene encoding pilin, the structural protein of the F pilus) that affect adsorption of both the DNA and RNA phages also reduce conjugal transfer. This strongly suggests that the F pili in integrase-expressing cells are normal, although there are some uncharacterized mutations in *traA* that generate resistance to the RNA phages without affecting conjugation (40).

Regardless of the mechanism of inhibition, it is not exerted upon all single-stranded DNA entering the cell, as shown by the fact that conjugal transfer efficiencies of F′ plasmids to F⁻, integrase-expressing recipients are normal. Thus, incoming, single-stranded F′ DNA is in some way protected from the effect of integrase.

Another, trivial, explanation was that the formation of inclusion bodies by integrase might somehow inhibit bacteriophage infection by a mechanism totally unrelated to its normal biochemical functions. To test this, we examined the effects of a plasmid expressing a truncated form of integrase (pBLE11), which also forms inclusion bodies. M13 plated normally on hosts expressing the truncated protein (Table 1; Fig. 1, lanes 13 and 14), suggesting that the presence of inclusion bodies *per se* is not the determining factor for the inhibition. This experiment also reveals that the N-terminal domain of integrase lacks this activity. IPTG was included in this experiment to increase the expression of truncated integrase to a level similar to that of the uninduced wild-type protein, since the basal level of expression of the truncated protein is considerably lower than that of wild-type integrase. Although induction with 4 μM IPTG caused some inhibition of growth of the bacterial lawn, plaques were formed at all dilutions of the phage.

Experiments with M13K07 (37), an M13 derivative carrying a kanamycin resistance gene, suggest that integrase may exert its effect early in the phage life cycle. Cells infected with M13K07 form colonies on plates containing kanamycin. Colony formation following infection of integrase-expressing cells with M13K07 was inhibited about 100-fold compared to infection of non-integrase-expressing hosts. The level of phage production from isolated Km^r colonies was then determined. Phage production from colony to colony varied over a wide range, but there was no significant difference between integrase-expressing and non-expressing hosts. Thus, establishment of infection appears to be the critical step: once a Km^r (infected) colony forms, phage is produced at a normal rate. This conclusion is supported by a preliminary pulse-labeling experiment in which RFI synthesis in the integrase-expressing strain was decreased at least 5-fold compared to the control, and no single-stranded DNA was synthesized. Other preliminary experiments show that at least some of the inhibitory effect is on conversion of single stranded DNA to RF DNA: integrase reduces the yield of phage arising from transfection with single stranded phage DNA more than it reduces the yield from RF DNA.

The finding that plaque formation is so much more stringently inhibited than is colony formation (>10 logs vs. 2 logs) is not unexpected; once infected with M13K07, each infected cell should give rise to a colony, whereas a plaque can form only if the surrounding cells are also susceptible to infection. That the background of Km^r colonies represents leakiness of inhibition rather than modification of the phage is shown by the fact that the M13K07 phage produced by these colonies do not form plaques on pLJS10-containing strains. We feel that this bacteriophage system is a promising approach for selection of integrase mutations, for screening of inhibitors of integrase, and for identification of new biochemical functions of integrase.

J. Anti-integrase Ribozymes. Ribozymes are RNA molecules that catalyze RNA cleavage. We designed a ribozyme directed against the HIV-1 integrase RNA at a position corresponding to nt 4027-4029 of the BH10 strain of HIV-1. Ribozyme α was synthesized *in vitro* by transcription from a bacteriophage T7 promoter. Ribozyme β is identical to ribozyme α except for the addition of a T7 transcription terminator at the 3′ end; it was synthesized *in vivo* from a plasmid introduced into *E. coli*. Ribozyme Δβ contains a single guanosine residue in place of the 22 nt catalytic domain of ribozyme β; it was used as a control for antisense effects.

Ribozymes α and β, both prepared *in vitro*, cleaved a 2,000 nt long target RNA containing the integrase gene into fragments of 1,500 and 500 nt. The reaction was much more extensive at 50°C than at 37°C. When ribozyme β and and integrase mRNA were simultaneously induced *in vivo*, the ribozyme cut its target. This led to destruction of the target RNA and blocked production of the integrase protein. Cleavage products were not recovered, probably due to their rapid degradation by ribonucleases. When the substrate mRNA was preinduced to allow its accumulation prior to expression of the ribozyme, the smaller of the cleavage products could be detected. Ribozyme Δβ, lacking the catalytic domain, caused only a slight reduction in the amount of *int* mRNA present compared to the controls. Thus, ribozyme β is extremely efficient inside *E. coli*, at 37°C, and the effect is not due to antisense properties. These results are detailed in the accompanying manuscript (Sioud and Drlica, Appendix 2) which has been accepted for publication in the *Proceedings of the National Academy of Sciences*.

CONCLUSIONS

We set out to develop *in vitro* assays for HIV integrase. We succeeded at this, and in addition we have discovered an *in vivo* screen for integrase in *E. coli*. Moreover, we have been able to demonstrate that a ribozyme aimed at an HIV gene is effective in *E. coli*, opening the way to to use this organism in the refinement of other ribozymes. Specific conclusions from our work are listed below.

A. Carboxy terminal processing does not occur with HIV-1 integrase, justifying the expression strategies used by ourselves and others.

B. HIV-1 circle junctions contain two additional base pairs compared to the integrated provirus. Like other retroviruses, HIV-1 integration results in the removal of 2 bp from the ends of the linear viral DNA. A significant fraction of the circular DNA present in infected cells contains abnormal circle junction sequences, which are most easily explained as arising via mispriming errors by reverse transcriptase. This work resulted in a publication (Hong *et al.*, J. Virology vol. 65, pp. 551-555; Appendix 2).

C. Integrase expression. We have expressed integrase in two ways in eukaryotic cells. Cells expressing only integrase produce large amounts of protein while cells expressing the entire *pol* gene produce little protein and are less viable. We confirmed that active integrase can be expressed in *E. coli* cells.

D. Solubilization of integrase. We have developed alternate solublization methods for integrase which produce higher yields of soluble protein than published methods; whether integrase activity is affected by these methods is not yet known.

E. Assays for integrase activity. In a Mn^{++}–dependent reaction, integrase expressed in *E. coli* trims 2 bp from the 3′ end of double-stranded oligomeric substrates corresponding to the ends of the HIV-1 LTR (U5 or U3). With different preparations, this reaction is between 20 and 80% complete. A strand transfer reaction in which one oligonucleotide is inserted into another is also observed, and there appears to be some site-specificity for this reaction. Plasmid constructions for an *in vivo* assay have also been prepared.

F. Bacteriophage inhibition assay. We have discovered that integrase blocks infection of *E. coli* by single-stranded DNA and RNA bacteriophages. Although we have not yet determined the mechanism(s) by which integrase blocks phage growth, the magnitude of the effect (10 orders of magnitude) makes this a promising approach for screening inhibitors of integrase. The method is simple, reliable, and can be automated. Since live cells are required for the assay to work, any inhibitor that kills cells will be eliminated from consideration. This may be advantageous, since such compounds may be expected to have harmful side effects in humans. The assay can also be used to select integrase mutants that are unable to inhibit phage replication; studies of such mutants will provide further information of functional domains of the integrase protein.

G. Anti-integrase ribozymes. Ribozymes that prevent the expression of HIV-1 integrase inside *E. coli* have been developed. These results support the concept that ribozymes have therapeutic potential. Furthermore, the demonstration that an HIV-1 RNA target can be eliminated in bacteria provides an important option for refining antiviral ribozymes, since powerful genetic selection can be brought to bear on developing more effective ribozymes.

REFERENCES

1. Brown, P , B. Bowerman, H. E. Varmus, and J. Bishop. (1987) Cell 49:347-356.
2. Brown, P., B. Bowerman, H. E. Varmus, and J. M. Bishop. (1989) Proc. Natl. Acad. Sci. USA 86:2525-2529.
3. Bushman, F. D., T. Fujiwara, and R. Craigie. (1990) Science 249:1555-1558.
4. Bushman, F. D., and R. Craigie. (1991) Proc. Natl. Acad. Sci. USA 88:1339-1343.
5. Cobrinik, D., R. Katz, R. Terry, A. M. Skalka, and J. Leis. (1987) J. Virol. 61:1999-2008.
6. Colicelli, J., and S. P. Goff. (1985) Cell 42:573-580.
7. Craigie, R., and K. Mizuuchi. (1987) Cell 51:493-501.
8. Craigie, R., T. Fujiwara, and F. Bushman. (1990) Cell 62:829-837.
9. Donehower, L. A., and H. E. Varmus. (1984) Proc. Natl. Acad. Sci. USA 81:6461-6465.
10. Ellis, J., and A. Bernstein. (1989) J. Virol. 63:2844-2846.
11. Farnet, C., and W. Haseltine. (1990) Proc. Natl. Acad. Sci. USA 87:4164-4168.
12. Farnet, C. M., and W. A. Haseltine. (1991) J. Virol. 65:1910-1915.
13. Fujiwara, T., and K. Mizuuchi. (1988) Cell 54:497-504.
14. Fujiwara, T., and R. Craigie. (1989) Proc. Natl. Acad. Sci. USA 86:3065-3069.
15. Hizi, A., and S. Hughes. (1988) Virology 167:634-638.
16. Hong, T., K. Drlica, A. Pinter, and E. Murphy. (1991) J.Virol. 65:551-555.
17. Katzman, M., R. A. Katz, A. M. Skalka, and J. Leis. (1989) J. Virol. 63:5319-5327.
18. Kulkostky, J., R. A. Katz, and A. M. Skalka. (1990) J. Acquired Immune Deficiency Syndromes 3:852-858.
19. Lobel, L., J. Murphy, and S. Goff. (1989) J. Virol. 63:2629-2637.
20. Panganiban, A. T., and H. M. Temin. (1983) Nature 306:155-160.
21. Panganiban, A. T., and H. M. Temin. (1984) Proc. Natl. Acad. Sci. USA 81:7885-7889.
22. Panganiban, A. T., and H. M. Temin. (1984) Cell 36:673-679.
23. Pauza, C. D. (1990) Virology 179:886-889.
24. Projan, S. J., M. Monod, C. S. Narayanan, and D. Dubnau. (1987) J. Bacteriol. 169:5131-5139.
25. Quinn, T., and D. Grandgenett. (1988) J. Virol. 62:2307-2312.
26. Rosenberg, A., B. Lade, D. Chui, S. Lin, J. Dunn, and W. Studier. (1987) Gene 56:125-135.
27. Roth, M., P. Schwartzberg, and S. Goff. (1989) Cell 58:47-54.
28. Sambrook, J., E. F. Fritsch, and T. Maniatis. (1989). *in* Molecular Cloning: A Laboratory Manual (2nd edition). Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
29. Schwartzberg, P., J. Colicelli, and S. P. Goff. (1984) Cell 37:1043-1052.

30. Sherman, P. A., and J. A. Fyfe. (1990) Proc. Natl. Acad. Sci. USA 87:5119-5123.

31. Smith, J. S., S. Kim, and M. Roth. (1990) J. Virol. 64:6286-6290.

32. Starcich, B., L. Ratner, S. F. Josephs, T. Okamoto, R. C. Gallo, and F. Wong-Staal. (1985) Science 227:538-540.

33. Studier, F., and B. Moffatt. (1987) Cell 49:221-227.

34. Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff. (1990) Meth. Enzymol. 185:60-89.

35. Varmus, H. E., and R. Swanstrom. (1985), p. 75-134. *in* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA Tumor Viruses, vol 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

36. Varmus, H., and P. Brown. (1989), p. 53-108. *in* D. E. Berg, and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.

37. Vieira, J., and J. Messing. (1987) Meth. Enzymol. 153:3-11.

38. Vincent, K. A., D. York-Higgins, M. Quiroga, and P. O. Brown. (1990) Nucl. Acids Res. 18:6045-6047.

39. Whitcomb, J. M., R. Kumar, and S. H. Hughes. (1990) J. Virol. 64:4903-4906.

40. Willetts, N. S., P. M. Moore, and W. Paranchych. (1980) J. Gen. Microbiol. 117:455-464.

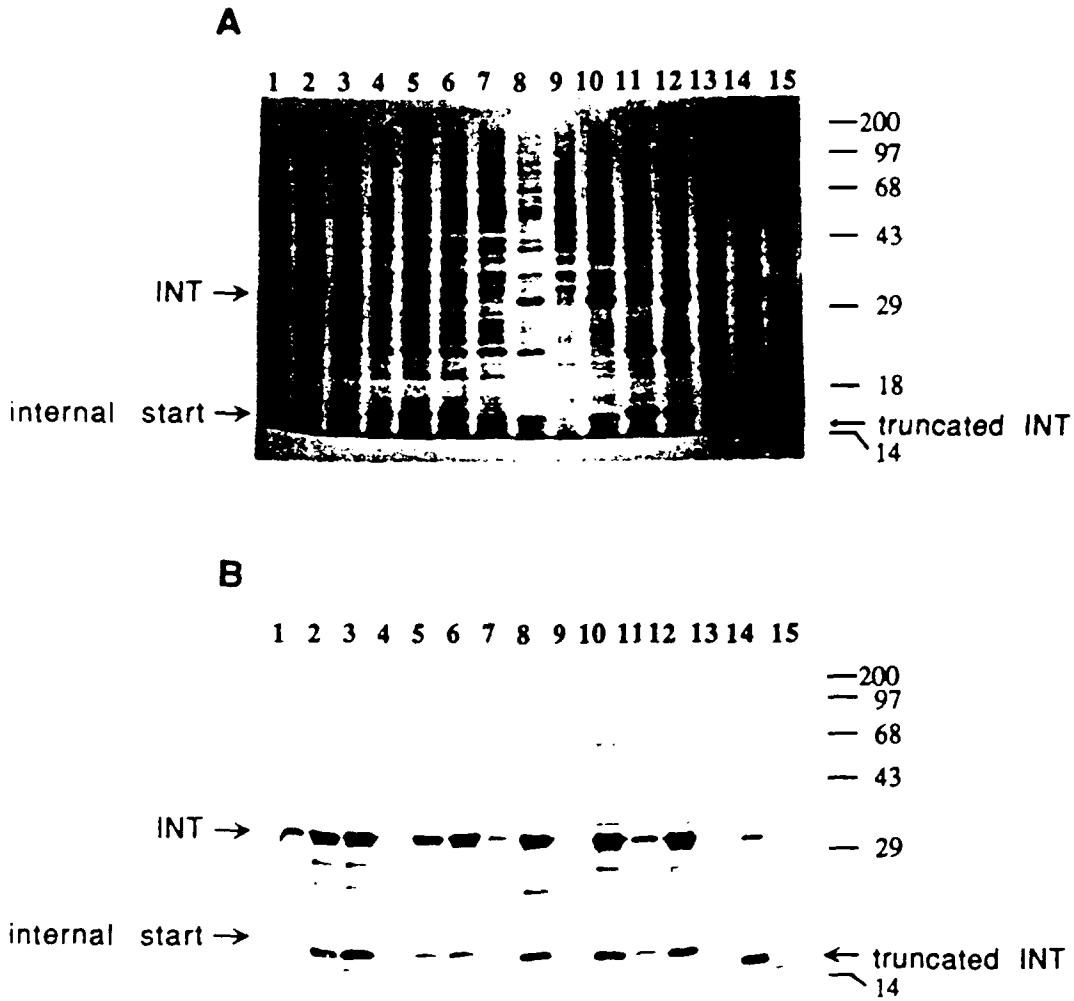

Figure 1. Expression of integrase in E. coli.

Panel A, Coomassie stained 13% PAGE gel; **Panel B**, Western blot developed with AIDS patient's serum. Molecular weight markers are indicated by short horizontal lines. Arrows show the locations of wild-type integrase (32 kDa), truncated integrase (16.4 kDa) produced by pBLE11 (lanes 13 and 14), and an 18.4 kDa protein derived from an internal start in *int* (visible in lanes 1–3, 4–6, 10; [15]). The additional prominent bands in lanes 11 and 12 are proteins derived from pLysE.

Lanes 1–8: cells (derivatives of F' HMS174 [DE3]) were induced with 1 mM IPTG in M9 medium for the times indicated. Aliquots were centrifuged and the pellets were suspended in 1 X SDS loading buffer (28). Each lane contains the protein from the equivalent of 5×10^7 cells. Lane 1, pLJS10 (integrase), time 0 (no IPTG); lane 2, pLJS10, 1.5 hr induction, lane 3, pLJS10, 3 hr induction; lane 4, pLJS10 and pLysE, no induction; lane 5, pLJS10 and pLysE, 1.5 hr induction; lane 6, pLJS10 and pLysE, 3 hr induction; lane 7, pLJS10 and pLysS, no induction; lane 8, pLJS10 and pLysS, 3 hr induction.

Lanes 9–14: cells were plated in 3 ml B top agar on LB plates for titration of M13 plaques. Plates included IPTG where indicated. After overnight incubation, cells were scraped from the lawn and lysed as above. M13 titres from these plates are shown in Table 1. Lane 9, F' HMS174 (DE3) host, no plasmids; lane 10, pLJS10 (integrase); lane 11, pLJS10 and pLysE; lane 12, pLJS10 and pLysE plus 1 mM IPTG; lane 13, pBLE11 (truncated integrase); lane 14, pBLE11 plus 4 μM IPTG; lane 15, molecular weight markers.

Figure 2. 3′ trimming and integration assay.

Substrate DNA was prepared by annealing a 5′-end labeled synthetic oligonucleotide (5′– CCTTTTAGTCAGTGTGGAAAATCTCTAGCAGT–3′) corresponding to the U5 end of HIV-1 with a complementary oligonucleotide. Integrase reactions (modified from [4, 30]) contained 1 pmol substrate DNA, approximately 30 pmol partially purified integrase protein, 20 mM MOPS buffer, pH 7.0, 50 mM potassium glutamate, 10 mM beta-mercaptoethanol, and 10 mM $MnCl_2$. Reactions were incubated at 37°C for 2 hours, terminated by the addition of 2 µl 0.5M EDTA, and evaporated to dryness. The reaction pellet was resuspended in loading buffer (95% formamide, 20 mM EDTA, 0.05% xylene cyanol, 0.05% bromophenol blue) and loaded on a 20% denaturing acrylamide gel. Lane 1, integrase omitted; lane 2, standard reaction; lane 3, $MnCl_2$ omitted. Upper band, 32-mer substrate; lower band, 30-mer trimmed substrate. Diffuse larger molecular weight bands in lane 2 represent products of the strand transfer reaction.

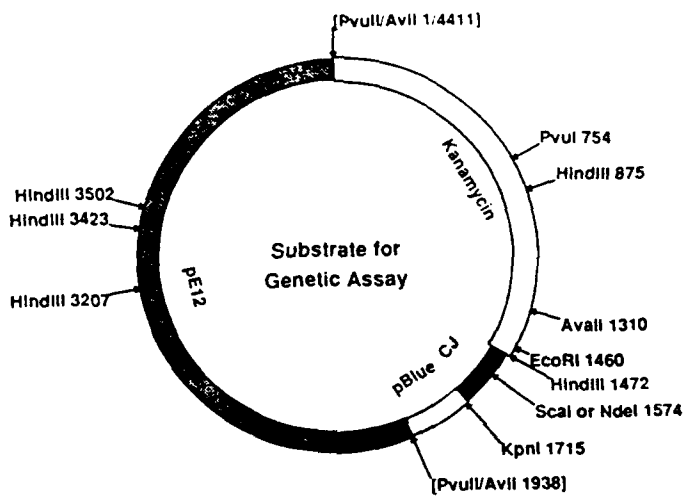

***Figure 3.* Restriction map of substrate plasmid for in vivo integration assay.** A 1.9 kb PvuII fragment containing a kanamycin resistance gene (white), a 243 bp circle junction sequence (black) and 223 bp derived from pBluescript (white) was cloned to pE12 (gray) at that plasmid's unique AvII site. The integration assay is described in the text.

TABLE 1

M13 INFECTION OF INTEGRASE-EXPRESSING CELLS

Lane No. Fig. 1	Plasmid to suppress T7 RNA polym.	Integrase plasmid	[IPTG]	Integrase expression	Relative M13 titer	Plaque morphology
9	–	–	–	NA	1.0	turbid
10	–	pLJS10	–	++++	$<10^{-10}$	–
–	pLysE	–	–	NA	1.0	clear
11	pLysE	pLJS10	–	±	1.1	clear
12	pLysE	pLJS10	1 mM	++++	2×10^{-2}	turbid, small
13	–	pBLE11	–	+	0.5	turbid
14	–	pBLE11	4 μM	++++	0.5	turbid

M13 titers are normalized to the plasmidless host HMS174 F′ [DE3] (2.5×10^{10} PFU/ml). Integrase production from cells taken from these plates is shown in Fig. 1. pLJS10 expresses wild-type integrase; pBLE11 expresses a 16.4 kDa N-terminal fragment of integrase. NA, not applicable. Note that plaque morphology comparisons are probably significant only among strains carrying the same pLys plasmid, because of the increased susceptibility of cells to lysis caused by the presence of T7 lysozyme (34).

PUBLICATIONS AND ABSTRACTS

Hong, T., K. Drlica, A. Pinter, and E. Murphy. (1991). Circular DNA of human immunodeficiency virus: Analysis of circle junction nucleotide sequences. *J. Virol.* 65, 551-555

Sioud, M. and K. Drlica (1991). Prevention of HIV-1 integrase expression in *E. coli* by a ribozyme. Proc. Natl. Acad. Sci. USA (in press).

Hong, T., Drlica K., Pinter A., and E. Murphy. Analysis of HIV-1 circular DNA: 2bp removed from each end of the viral DNA during integration. Abstract (poster) at "RNA Tumor Viruses", May 23-27, 1990, Cold Spring Harbor Laboratory.

Peinjnenburg, A., N.-Z. Xu, Drlica K., Pinter A., Murphy E., Honnen, W., and S. Kayman. Expression of HIV integrase protein in insect cells using baculovirus vectors. Abstract (poster) at "RNA Tumor Viruses", May 23-27, 1990, Cold Spring Harbor Laboratory.

Sioud, M., and K. Drlica. Prevention of HIV-1 integrase expression in *E. coli* by a ribozyme. Abstract (oral presentation) at "Catalytic RNAs as anti-HIV agents: design and delivery to cells". Oct. 21-24, 1990, San Diego.

PERSONNEL SUPPORTED BY THIS CONTRACT

Ellen Murphy, Ph.D., P.I.
Karl Drlica, Ph.D., co-P.I.
Richard Burger, Ph.D.
Samuel Kayman, Ph.D.
Nau-Zheng Xu, Ph.D.
Ad Peinjnenburg, Ph.D.
Mouldy Sioud, Ph.D.