Future VIROLOGY

SPECIAL REPORT

For reprint orders, please contact: reprints@futuremedicine.com

The road to HIV-1 integrase inhibitors: the case for supporting basic research

Robert Craigie*

ABSTRACT AIDS has been transformed from a death sentence to a manageable disease for many patients with access to combination antiviral therapy. It is informative to look back on some of the key advances that have led to this transformation. The arsenal of tools currently available to clinicians now includes inhibitors of the viral reverse transcriptase, protease and integrase enzymes. The author discusses some of the key advances that have led to this transformation with an emphasis on the role of basic science in developing integrase inhibitors. Many of the stepping-stones could not easily have been foreseen to lead to medical advances. Treatments for diseases that are yet to emerge will likely depend on the progress made in basic science today.

KEYWORDS
• AIDS • basic science • HIV-1 • integrase inhibitors

Retroviruses & disease

At the time HIV-1 was identified as the causative agent of AIDS in 1983 [1], retroviruses had recently been found to be associated with lymphoma [2] and Sézary syndrome [3] in humans. However, these discoveries were preceded by studies of diseases in animals that are caused by retroviruses. The first examples, leukosis and sarcoma in chickens, were discovered in the early 20th century [4,5]. Today these viruses are collectively termed as avian sarcoma/leukosis viruses (ASLVs). Subsequently, viruses associated with cancers were found in many other species of vertebrates. As molecular tools to study viral replication were developed in the 1960s, viruses such as ASLV presented a puzzle. No double-stranded RNA replication intermediates were found as with other RNA viruses. Furthermore, replication was sensitive to inhibitors of DNA synthesis and DNA-dependent RNA synthesis. These observations led Howard Temin to propose the radical provirus hypothesis that viral RNA introduced into cells upon infection is converted to DNA that is integrated into DNA of the host genome, thus explaining how the transformed state of cells infected with ASLV was maintained in the absence of viral replication [6,7]. The proposal was met with skepticism at the time because the dogma was that genetic information flows in one direction from DNA to RNA, to protein. However, the provirus hypothesis was subsequently vindicated by the discovery of reverse transcriptase [8,9], direct detection of integrated viral DNA in the genome of infected cells [10] and characterization of the structure of the integrated provirus [11].

Two paradigms for DNA integration from studies in prokaryotes

How might the viral DNA be integrated into cellular DNA? Studies with bacteriophage lambda suggested a possible mechanism. Lambda is a DNA 'virus' of bacteria that integrates into the genome as a lysogen that is analogous to the prophage of retroviruses. Integration occurs by the mechanism first postulated by Alan Campbell in 1962 [12]. Briefly, the bacteriophage DNA is circularized after infection and integration is accomplished by recombination between sites on the circular lambda DNA and chromosomal DNA. The finding of circular forms of retroviral DNA after infection gave credence to such a mechanism and for a time circular molecules that bear two copies of the viral long terminal repeat sequence were thought to be the direct precursor of the integrated provirus. However,

*Laboratory of Molecular Biology, National Institute of Diabetes & Digestive & Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA; bobc@helix.nih.gov

Future Medicine part of fsg

the hypothesis was disproved and the precursor to integrated retroviral DNA was found to be the linear form [13,14].

Retroviral DNA integration occurs by a mechanism that is shared by bacteriophage Mu and many DNA transposons in prokaryotes. Biochemical studies of these systems greatly facilitated the later biochemical studies of DNA integration of HIV-1 and other retroviruses. Mu is a bacteriophage that uses the DNA transposition mechanism to form the replication forks that initiate replication of its genome. In the process of multiple rounds of replication the host genome is destroyed and progeny phages are released. Although transposons of this class use essentially the same biochemical mechanism to move from one location in the genome to another, they have typically evolved to do so very inefficiently so as not to kill the host cell during this process. The inherently high efficiency of the Mu system, which results from its lifestyle as a phage, jump-started biochemical studies of DNA transposition [15]. It is an example where picking a system that is highly amenable to study can lead to rapid advances that can be applied later to study more intractable systems.

In the simplest type of DNA transposition, a transposase encoded by the transposon excises the transposon from its original site in DNA such that free hydroxyl groups exist at the 3′ termini of linear transposon DNA. The transposase then uses these 3′-OH groups to attack a pair of phosphodiester bonds in the target DNA, resulting in the covalent joining of the transposon to the target DNA. Finally cellular enzymes complete the process. The sites of joining on the target DNA strands are typically staggered by several nucleotides resulting in the duplication of a short sequence of target DNA flanking the newly integrated transposon. Some of the many variations on this theme have been previously reviewed in [16]. These studies revealed the integral role of highly stable nucleoprotein complexes in DNA transposition. The transposon ends are first paired in the Stable Synaptic Complex, within which the 3′ ends of the transposon are cleaved to form the Cleaved Donor Complex. Subsequently, the Cleaved Donor Complex captures a target DNA and DNA strand transfer occurs to form the Strand Transfer Complex. Each successive complex on the reaction pathway is progressively more stable. These complexes have collectively been called transpososomes [17]. It is noteworthy that retroviral DNA integration recapitulates such a series of stable nucleoprotein complexes. The retroviral term ‘intasomes’ is analogous to the prokaryotic transpososomes.

Early biochemical studies of retroviral DNA integration

Biochemical studies of retroviral DNA integration were pioneered by studies of avian and murine systems, notably ASLV and Moloney murine leukemia virus (MLV), which in turn were stimulated by earlier work with prokaryotic DNA transposition systems. Genetic studies identified mutations in two regions of the viral DNA that abolished integration while allowing infection and reverse transcription to proceed normally. One site mapped to the ends of the linear viral DNA while the second site mapped to the 3′ end of the viral *pol* gene. This region of *pol* encodes the protein we now call integrase that is cleaved from the Gag-Pol polyprotein by the viral protease. The first biochemical studies of this protein identified it as a relatively non-specific endonuclease present in ASLV virions [18]. Although this activity was proposed to be important for integration, the mechanism was unknown as was the proper precursor DNA substrate for integration.

The finding that viral DNA within extracts made from cells infected with MLV efficiently integrated into exogenous DNA *in vitro* invigorated biochemical studies of retroviral DNA integration [19]. Using this system, the precursor viral DNA for integration was unambiguously determined to be the linear form [13,14]; the circular forms are dead-end products that do not go on to integrate. The next advance was the development of cell-free systems using detergent-disrupted MLV virions as the source of integration activity [20]. Although these *in vitro* reaction systems provided powerful tools to study the mechanism of DNA integration, the assay systems were laborious and not suitable for large-scale screening for inhibitors of integrase activity.

Simplified *in vitro* assay systems & inhibitors

The discovery that integrase is necessary and sufficient for catalysis of 3′ end processing and DNA strand transfer paved the way for high-throughput screening for inhibitors. The structure of the integration intermediate proved that the precursor for integration is linear and that the reaction must proceed in two steps. The

terminal two nucleotides are missing from each end of the integrated viral DNA, so the first step must involve removal of two nucleotides from each 3′ end of the viral DNA–3′ end processing. These 3′ ends are then integrated into target DNA–DNA strand transfer (joining). Cellular enzymes then repair the resulting integration intermediate structure. The necessary steps are removal of two overhanging nucleotides at each 5′ end of the viral DNA, filling in of the single strand gaps between viral and target DNA, and ligation. It was shown with ASLV [21], MLV [22] and HIV-1 [23] that the integrase protein is sufficient for 3′ end processing with short oligonucleotide substrates that mimic the ends of the respective blunt ended linear viral DNA. The same integrases are also necessary and sufficient for DNA strand transfer with oligonucleotide DNA substrates [22,24,25].

Recapitulation of the chemical steps of DNA integration with purified integrase and simple DNA substrates laid the foundation for high-throughput screening for compounds that inhibit these reactions. For a detailed account of the steps leading to the development of the first integrase inhibitor approved by the US FDA, the Merck drug raltegravir, see [26]. Screening for inhibitors of DNA strand transfer rather than 3′ end processing was one of the keys to the success of the Merck effort; indeed, all the integrase inhibitors to date that target the active site of integrase are selective for the DNA strand transfer step. Still the path from initial screening to a therapeutic drug was laborious and spanned a period from the early 1990s to final approval of raltegravir in 2007. The first promising compounds were the diketo acids L-731988 and L-708906. These compounds inhibited HIV replication in cell culture and passed the important tests that steps prior to integration, such as reverse transcription, were not affected and serial passage of HIV in cell culture in the presence of the inhibitor resulted in mutations that mapped to integrase [27]. Diketo acids have undesirable properties that make them unsuitable for drug development, but they served as the starting point for medicinal chemists at Merck and Shionogi to develop compounds with more favorable properties that inhibit integrase by the same mechanism.

Mechanism

Diketo acids exhibit a strong inhibitory selectively for DNA strand transfer but have little effect on 3′ end processing. This is contrary to what would be expected for simple binding to integrase near the active site blocking access to DNA substrate. An important step in understanding the mechanism was the finding that binding of radiolabeled diketo acids to integrase is approximately 1000-fold less in the absence of viral DNA and a divalent metal ion [28]. The inhibitor therefore binds to integrase after engagement of the viral DNA ends. To date, efforts to obtain high-resolution structures of HIV-1 integrase in complex with viral DNA has not been successful. However structures of the closely related prototype foamy virus intasomes have been determined [29]. These structures reveal the detailed mechanism of inhibitors such as raltegravir. Binding of the inhibitor displaces the reactive 3′ end of the viral DNA from the active site, thus preventing catalysis of DNA strand transfer. This is an example where studying a mechanistically related system that is not involved in human disease has provided invaluable insights that could not otherwise have been obtained.

Future perspective

The mobility of human populations in the modern age enables newly emerging diseases to quickly spread worldwide as happened with AIDS. It is easy to envisage worse scenarios in which a new disease is readily transmitted by airborne routes and it seems likely that we will face such problems in the future. In the case of HIV, earlier basic research on polymerases and proteases in other systems laid the foundation for rapid development of antiviral drugs targeting these enzymes. The development of drugs targeting integrase was much slower because the mechanism of integration was not well understood at the time HIV emerged as a serious public health problem. High-throughput screening for integrase inhibitors had to await the development of suitable *in vitro* assay systems. Combating future emerging diseases will undoubtedly depend on knowledge acquired from basic research that is ongoing today. Since the nature of the challenges we will face in the future is unknown, our best defense is advancing broad-based basic research in the biological sciences.

Financial & competing interests disclosure

This work was supported by the Intramural Program of NIDDK, the NIH and the AIDS Targeted Antiviral Program of the Office of the Director of the NIH. We thank

Alan Engelman for comments on the manuscript. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

EXECUTIVE SUMMARY

Retroviruses & diseases

- First examples of diseases in animals that we now know to be caused by retroviruses were discovered over 100 years ago.
- AIDS is one of the first human diseases known to be caused by a retrovirus.
- DNA integration is essential for retroviral replication; the integrated proviral DNA is stably replicated along with cellular DNA.

Two paradigms for DNA integration

- A popular early model was that retroviral DNA integration occurs by a mechanism similar to that of the extensively studied bacteriophage lambda. This hypothesis was later disproved.
- Retroviral DNA integration uses a mechanism that is very similar to how a class of DNA transposons hop from one location to another in DNA.
- Biochemical studies of DNA transposition in prokaryotes laid the groundwork for later biochemical studies of retroviral DNA integration.

Early biochemical studies of retroviral DNA integration

- Pioneered by studies of avian and murine retroviruses.
- The discovery that extracts of cells infected with a murine retrovirus support DNA integration *in vitro* was a pivotal advance.

Simplified *in vitro* assay systems & inhibitors

- Discovery that integrase is necessary and sufficient for catalysis of the chemical steps of DNA integration enabled the development of simple *in vitro* assay systems.
- Provided a powerful tool to study reaction mechanism.
- Were adapted into high-throughput screening format to search for inhibitors.
- The first integrase inhibitor was approved by the US FDA in 2007.
- Structural studies with prototype foamy virus integrase have been crucial to understanding the mechanism of action of this class of inhibitor.

Future perspective

- Basic research on DNA transposition in prokaryotes and retroviruses that are non-pathogenic to humans laid the foundation for the development of HIV-1 integrase inhibitors.
- Combating future emerging diseases will undoubtedly depend on knowledge acquired from basic research that is ongoing today.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Barre-Sinoussi F, Chermann JC, Rey F *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune-deficiency syndrome (AIDS). *Science* 220(4599), 868–871 (1983).

2 Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC. Detection and isolation of type-C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc. Natl Acad. Sci. USA* 77(12), 7415–7419 (1980).

3 Poiesz BJ, Ruscetti FW, Reitz MS, Kalyanaraman VS, Gallo RC. Isolation of a new type C retrovirus (HTLV) in primary uncultured cells of a patient with Sezary T-cell leukemia. *Nature* 294(5838), 268–271 (1981).

4 Ellermann V, Bang O. Experimentelle Leukämie bei Hühnern. Zentralbl. Bakteriol. Parasitenkd. Infectionskr. *Hyg. Abt. Orig.* 46, 595–609 (1908).

5 Rous P. A sarcoma of the fowl transmissible by an agent from the tumor cells. *J. Exp. Med.* 13(4), 397–411 (1911).

6 Temin HM. Participation of DNA in Rous sarcoma virus production. *Virology* 23(4), 486-& (1964).

7 Temin HM. DNA provirus hypothesis. *Science* 192(4244), 1075–1080 (1976).

8 Temin HM, Mizutani S. RNA dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226(5252), 1211–1213 (1970).

9 Baltimore D. RNA-dependent DNA polymerase in virions of RNA tumor viruses. *Nature* 226(5252), 1209–1211 (1970).

10 Varmus HE, Vogt PK, Bishop JM. Integration of deoxyribonucleic acid specificfor Rous-sarcoma virus after Iinfection of permissive and nonpermissive hosts. *Proc. Natl Acad. Sci. USA* 70(11), 3067–3071 (1973).

11 Hughes SH, Shank PR, Spector DH *et al.* Proviruses of avian-sarcoma virus are terminally redundant, co-extensive with unintegrated linear DNA and integrated at many sites. *Cell* 15(4), 1397–1410 (1978).

12 Campbell AM. EPISOMES. *Adv. Genet.* 11, 101–145 (1962).

13 Brown PO, Bowerman B, Varmus HE, Bishop JM. Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc. Natl Acad. Sci. USA* 86, 2525–2529 (1989).

14 Fujiwara T, Mizuuchi K. Retroviral DNA integration: structure of an integration intermediate. *Cell* 54, 497–504 (1988).

15 Mizuuchi K. *In vitro* transposition of bacteriophage Mu – a biochemical approach to a novel replication reaction. *Cell* 35(3), 785–794 (1983).

•• **The first *in vitro* system for DNA transposition.**

16 Bouuaert CC, Chalmers RM. Gene therapy vectors: the prospects and potentials of the cut-and-paste transposons. *Genetica* 138(5), 473–484 (2010).

17 Chaconas G. Studies on a "jumping gene machine": higher-order nucleoprotein complexes in Mu DNA transposition. *Biochem. Cell Biol.* 77(6), 487–492 (1999).

18 Grandgenett DP, Vora AC, Schiff RD. A 32,000-Dalton nucleic acid-binding protein from avian retrovirus cores possesses DNA endonuclease activity. *Virology* 89, 119–132 (1978).

•• **The first reported biochemical activity of a retroviral integrase.**

19 Brown PO, Bowerman B, Varmus HE, Bishop JM. Correct integration of retroviral DNA *in vitro*. *Cell* 49, 347–356 (1987).

•• **The first *in vitro* system for integration of retroviral DNA.**

20 Fujiwara T, Craigie R. Integration of mini-retroviral DNA: a cell-free reaction for biochemical analysis of retroviral integration. *Proc. Natl Acad. Sci. USA* 86, 3065–3069 (1989).

21 Katzman M, Katz RA, Skalka AM, Leis J. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the *in vivo* sites of integration. *J. Virol.* 63, 5319–5327 (1989).

22 Craigie R, Fujiwara T, Bushman F. The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration *in vitro*. *Cell* 62, 829–837 (1990).

•• **A retroviral integrase is necessary and sufficient for the chemical steps of DNA integration.**

23 Sherman PA, Fyfe JA. Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleaving activity. *Proc. Natl Acad. Sci. USA* 87, 5119–5123 (1990).

24 Katz RA, Merkel G, Kulkosky J, Leis J, Skalka AM. The avian retroviral IN protein is both necessary and sufficient for integrative recombination *in vitro*. *Cell* 63, 87–95 (1990).

•• **A retroviral integrase is necessary and sufficient for the chemical steps of DNA integration.**

25 Engelman A, Mizuuchi K, Craigie R. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell* 67, 1211–1221 (1991).

26 Egbertson MS, Anthony NJ. HIV integrase inhibitors: from diketo acids to heterocyclic templates: history of HIV integrase medicinal chemistry at Merck West Point and Merck Rome (IRBM) leading to discovery of raltegravir. In: *HIV-1 Integrase: Mechanism and Inhibitor Design*. Neamati N (Ed.). John Wiley & Sons, Inc., NJ, USA 197–229 (2011).

• **A detailed account of the development of HIV-1 integrase strand transfer inhibitors.**

27 Hazuda DJ, Felock P, Witmer M *et al.* Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science* 287, 646–650 (2000).

•• **Proof of concept that integrase inhibitors can effectively block HIV-1 replication.**

28 Espeseth AS, Felock P, Wolfe A *et al.* HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. *Proc. Natl Acad. Sci. USA* 97(21), 11244–11249 (2000).

29 Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* 464, 232–236 (2010).

•• **The first high-resolution structure of a retroviral intasome. Elucidated how integrase strand transfer inhibitors work at the molecular level.**