A model for reverse transcription by a dimeric enzyme

Peter R. Cook

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, U.K.

A speculative model for reverse transcription of a viral RNA template into proviral dsDNA is presented. It has two essential features that are not included in current models: (i) the functional complex is dimeric, with two polymerization/RNase H sites and (ii) two templates are initially attached to the complex at their 3' ends. The model also has the optional features that (iii) the

Introduction

The process whereby a retrovirus generates proviral dsDNA from a viral RNA template is extremely complicated and still incompletely understood (reviewed by Fields, 1990; Jacobo-Molina & Arnold, 1991). It is carried out by a remarkable enzyme contained within the virion, reverse transcriptase (RT). RT provides the sole catalytic activity required; it has both a polymerizing activity (copying an RNA or a DNA template) and an RNAse H activity (able to digest the RNA strand of an RNA-DNA duplex into short RNA oligomers).

Two particular problems must be solved during retrotranscription. Simple end-to-end copying of the linear template is precluded because RT, like other DNA polymerases, can extend only from a 3' end, raising the difficulty of how the extreme 5' end might be copied. Secondly, the viral RNA on which RT acts is made by cellular RNA polymerase II using signals lying upstream of the site of transcriptional initiation, so these signals cannot be copied directly into the transcript. Obviously copies of these signals must be encoded within the genome and some way must be found to copy and rearrange them so that they end up 5' to the site of initiation of viral transcription on proviral DNA. In the current model, these problems are solved by careful choreography of RT's movements (Gilboa et al., 1979; Fields, 1990).

However, several aspects of retro-transcription are explained with difficulty by the conventional model. First, the virion contains two similar RNA molecules (Bender *et al.*, 1978), but there remains no satisfactory explanation for this redundancy. [For recent discussions, see Bieth *et al.* (1990), Prats *et al.* (1990).] Secondly, complex is rotationally symmetrical and (iv) attached to the virion core. The model attempts to explain why the viral genome is dimeric, the specificity of the 'jumps' between the ends of templates and how recombination occurs so readily. It also suggests novel targets for drug therapy during retroviral infection, for example, in AIDS.

genetic information from only one of the two molecules is transferred to the integrated provirus, even though RT acts on both molecules (Panganiban & Fiore, 1988; Hu & Temin, 1990*a*,*b*; Stuhlmann & Berg, 1992). Thirdly, and for the reasons described above, the sequence of the dsDNA product is not collinear with that of the ssRNA template. It must be derived by precisely choreographed 'jumps' between the ends of templates. In the first, RT carries the DNA product it has just made from the end of one template to the 3' end of one or other genome in the virion (i.e. the jump is inter- or intra-molecular); in the second, the jump is intra-molecular (Panganiban & Fiore, 1988; Hu and Temin, 1990a,b; Stuhlmann & Berg, 1992). It is difficult to see how a mobile and diffusible enzyme might be choreographed to jump so precisely, without some mechanism forcing the appropriate ends together.

We might expect inter-molecular jumping to be much less frequent than intra-molecular jumping on statistical grounds since at low concentrations, ends of the same molecule are together more frequently than ends of different molecules. However, initial observations using spleen necrosis virus suggested that first jumps were exclusively inter-molecular (Panganiban & Fiore, 1988). Subsequent analysis showed that inter- and intramolecular jumps occur at roughly the same frequency (Hu & Temin, 1990*b*), implying that the (jumping) polymerization site lies close to the 3' ends of both genomes at the appropriate time. Any model for reverse transcription should account for these jumps, especially the inter-molecular ones.

Another intriguing aspect involves the role played by the viral core. Both genetic and biochemical experiments show that a core constituent, the nucleocapsid protein, is



Fig. 1. A model for reverse transcription. Each panel illustrates two symmetrical and fixed half-complexes in the virion, each containing a polymerizing/RNase H activity and a binding site for a 3' end. RNase H acts in concert with, and just behind, each polymerization site. Events at only the lower of the two sites are described in (a) to (e). (a) Two genomic RNAs are associated with the complex [5' caps and 3' poly(A) are not shown]. Each genome is associated with both half-complexes, at the polymerization site of one half-complex

absolutely required for reverse transcription during a natural infection; in its absence, human immunodeficiency virus type 1 (HIV-1) RT cannot retrotranscribe viral RNA from the tRNA primer (Barat et al., 1989). Moreover, some aspect of core structure is required for the first jump, as it seems to be performed inefficiently by pure RT. [High template concentrations increase this efficiency (Luo & Taylor, 1992).] Indeed, cores of murine leukaemia virus are involved in the integration of the product of reverse transcription into the host chromosome (Bowerman et al., 1989: but see also Craigie et al., 1990; Katz et al., 1990). Whilst the virion is surrounded by an impermeable lipid bilayer. cores of yeast retro-transposons are porous, allowing ingress of nucleoside triphosphates and even small proteins the size of RNase A (Burns et al., 1992). As a result, the idea is growing that the core both provides a protective exoskeleton for the genome and organizes its functions until well after retro-transcription is complete (Burns et al., 1992). It is only natural to assume that specific attachments of templates to the core position the templates such that jumps can be made precisely.

A model

I now suggest a speculative modification of the current model that attempts to explain these intriguing aspects (Fig. 1). Discussion concentrates on the HIV-1 enzyme, but reference is made to other activities. The model is not meant to be definitive but to highlight current assumptions, provoke experimentation and suggest novel targets for drug therapy. It has the following features: (i) the functional RT complex is dimeric, with two polymerization/RNase H sites, (ii) two templates are initially attached to the complex at their 3' ends (as well as through the primer-binding site), (iii) the complex is rotationally symmetrical and (iv) attached to the virion core. Analogous models with the essential features (i) and (ii), but excluding features (iii) and/or (iv), have most of the advantages of the model drawn here.

Although a great deal is known about the structureactivity relationships of RT expressed in bacteria (Kohlstaedt et al., 1992), little is known about the natural enzyme found in the virion core (Jacobo-Molina & Arnold, 1991), and even less about the core itself (but see Burns et al., 1992). We do not know, for example, how many RT molecules there are in an infectious virus particle: a value of 80 is often quoted (Layne et al., 1992), but as less than 1 in 10^4 of the particles in the sample analysed were infectious, this is inevitably an estimate. Nevertheless, it is known that the virion enzyme is a heterodimer (Veronese et al., 1986), formed by an initial dimerization of two p66 subunits; one subunit is then cleaved proteolytically into p51 and p15. Although both p66 and p51 encode a polymerizing domain, only one is active in the bacterially expressed p66/p51 heterodimer and therefore this is probably also true of the virion heterodimer. [It is not known whether the cleaved p15 remains associated with p66/p51 in the virion. Therefore, given that the virion and bacterially expressed enzymes differ (e.g. in jumping efficiency), it remains possible, but unlikely, that one p66/p51-p15 complex in the virion could contain two polymerase/ RNase H sites.] Since RT must act on each of the two genomes in the virion, it is natural to assume that this is performed by a dimeric complex (i.e. one with two active

⁽through a tRNA primer) and at the 3' binding site of the other. (An alternative involving attachment of each genome to one, or other, of the two half-complexes is also possible but not illustrated.) The tRNA is shown priming synthesis as the template slides (arrow) through the polymerization/RNase H site. (b) As the template slides to the right, nascent negative strand DNA is extruded into a loop (tRNA remains bound). Simultaneously, the template is cut 5' to the primer-binding site (PB) and then U5 is degraded as it passes the RNase H site. (c) U5 and R have been copied and degraded, completing synthesis of 'strong stop' (negative strand) DNA. The 3' R on the other genome 'jumps' (arrow) into the now-empty template-binding region of the polymerization site, annealing with negative strand R'. This gives an inter-molecular first jump. [Alternatively, the 3' end and PB of each genome might be attached to the same half of the complex in (a) to give an intra-molecular first jump (not shown). (d) The new template slides rightwards and U3 is both copied and degraded as more negative strand is extruded. (e) As the polypurine tract (PP) slid past RNase H, it was cut at each end, remaining annealed to nascent negative strand. The resulting PP-RNA/DNA duplex immediately competes for the upper site. (f). The lower loop of negative strand DNA in (e) rearranged as the PP-RNA/DNA duplex occupied the upper site, displacing the resident RNA/DNA duplex which is now not shown and which takes no further part. The tRNA is shown detached from the lower site for clarity. DNA synthesis and RNA degradation continue at both sites. (g) At the lower site, PB has been copied and degraded. The upper site, having copied beyond the end of the negative strand, has stopped synthesis at a methyl-adenine in the tRNA; RNase H has just cleaved the tRNA-DNA junction 18 nucleotides from the end. After tRNA dissociation, both 3' ends remain attached to their respective polymerization sites whilst complementary sequences at each end anneal, driving the 'second jump'. Additional synthesis then directly gives (j). However, this transition is difficult to draw on paper, so intermediates (h) and (i) are included for illustration only. (h) Interstitial DNA is redrawn below and PB has detached from the upper site and annealed with PB' at the lower site. (i) Negative strand DNA is extruded by strand-displacement synthesis at the lower site. (j) The 3' end of the positive strand has reattached to the upper site; synthesis continues at both sites. Annealing and synthesis (without detachment and reattachment) generates this structure directly from (g). (k) Synthesis at the lower site is complete. The 5' end of the negative strand is shown rearranged; positive strand synthesis continues. (1) Synthesis is complete.

polymerase/RNase H sites). This requirement would be met by dimerization of two heterodimers to give $(p66/p51)_2$. Analysis of the crystal structure of the bacterially expressed enzyme suggests that the p66/p51 complex is unlikely to dimerize directly (Kohlstaedt *et al.*, 1992), but it might well do so through some other core component (e.g. the essential nucleocapsid protein). It is a general rule that dimeric subunits are related by a twofold axis of rotation so this is assumed here, but an analogous model involving an asymmetrical (e.g. parallel) arrangement is also possible. [Asymmetrical p66/p51 is an exception to this rule; the sequence encoding the catalytic site in p66 is accessible, whereas the same sequence in p51 is buried (Kohlstaedt *et al.*, 1992).]

Templates are probably initially attached through primer-binding sites to tRNA at the polymerization sites (Fig. 1*a*). The 3' ends of the two templates are also attached (either directly to the complex or via the core) to facilitate jumping. Two possibilities can be envisaged: the primer-binding site and the 3' end of one genome could be attached to the same, or different, halves of the complex. Only the latter structure is illustrated in Fig. 1(a) as it facilitates the inter-molecular first jump that is explained with difficulty by current models. The other structure is probably more likely to be assembled into the virion; it can be visualized by cutting the middle of each template in Fig. 1(a), then joining the two lower (cut) ends below the lower tRNA and the two upper (cut) ends above the upper tRNA. It leads to an intra-molecular first jump and then on to dsDNA, much as described in Fig. 1.

In Fig. 1, 5' ends are drawn unattached for clarity; however, all ends probably remain attached to prevent entanglement and nucleolytic degradation. Thus in Fig. 1(b), the 5' primer-binding site (PB), which lacks a cap (unlike the 5' R of intact positive strand RNA) and would be particularly sensitive to degradation unless attached, probably remains annealed to tRNA and therefore attached.

If RT has two symmetrically disposed polymerase/ RNase H sites and is a mobile enzyme, then, since the two templates necessarily point in opposite directions, RT must first track along one template in one direction and then along the second in the other. In other words, the two sites work discontinuously. Since positive strand DNA synthesis begins before completion of negative strand DNA in avian retroviruses (Varmus *et al.*, 1978; Boone & Skalka, 1981), this suggests, but does not prove, that both sites are active together. A second possibility is that two asymmetrical polymerase/RNase H sites act together as they track along two parallel templates.

A third possibility, the particular one illustrated here, is that the functional dimer is symmetrical and fixed to the virion core, with reverse transcription occurring as two templates move past their respective polymerization sites, probably (but not necessarily) with the two sites active simultaneously. Attaching the enzyme to the core accounts for the core's necessary role. It also incorporates the way RT operates into a general view of how polymerases might work. All polymerases (including both DNA and RNA polymerases) have some very weak, but nevertheless convincing, sequence similarities in local regions (e.g. Argos, 1988) and this homology is much stronger at the structural level (reviewed by Kohlstaedt et al., 1992). For example, the threedimensional structure of only two DNA polymerases is known, RT and the Klenow fragment of the bacterial DNA polymerase I. Despite almost non-existent sequence homology, two of the four polymerase subdomains of RT are structurally related to subdomains of the Klenow fragment, including the catalytic one (Kohlstaedt et al., 1992). As recent models for DNAdependent RNA and DNA synthesis involve reeling templates through fixed polymerization sites (Cook, 1989, 1991), it becomes attractive to suppose that all polymerases work in the same basic way. The idea that polymerases run along their templates stems from a perception of relative size, i.e. that the smallest object moves. But if RT is attached to the virion core, then the template is relatively small and would be expected to move instead.

In Fig. 1, the two polymerization/RNase H sites in the symmetrical structure remain stationary as the template slides past them. RNase H sites necessarily lie a fixed distance away from polymerizing sites and act in concert with them; there is now considerable evidence for such concerted action of RNase H approximately 15 to 18 bp behind the polymerizing site (Oyama et al., 1989; Huber & Richardson, 1990; Wohrl & Moelling, 1990; Furfine & Reardon, 1991; Fu & Taylor, 1992). Specific cuts are also shown 18 bp behind the polymerization site when both the polypurine tract (PP) and tRNA primer are released (i.e. at stages represented in Fig. 1e and g). [HIV-1 may also have an RNase D activity that cuts an RNA/RNA hybrid 10 and 14 bp from the polymerization site; this might also be involved in releasing tRNA (Ben-Artzi et al., 1992).] Each half of the dimer initially binds one or other of the templates at two (and probably more) places (Fig. 1a). The PB of one of the two templates is attached to one polymerization site (through tRNA), whilst the 3' end of the same template lies close to, but not quite at, the other polymerizing site, ready for an inter-molecular first jump. [As discussed above, the alternative arrangement in which the primer-binding site and 3' end of each genome are attached to the same half-complex, which is not shown here, leads to an intra-molecular first jump.] After synthesis of negative strand DNA and the

concerted removal of template by RNase H (Fig. 1c), the 3' end jumps the short distance to the now empty template-binding region of the polymerization site and anneals with its complement in the nascent negative strand (Fig. 1d). The proximity of the 3' end to the polymerization site and the affinity for a 3' end ensure efficient jumping.

Only one of the two RNA templates is completely copied into DNA (Panganiban & Fiore, 1988; Hu & Temin, 1990*a*,*b*; Stuhlmann & Berg, 1992). In Fig. 1(*e*) and (f), one of the two templates is displaced from the complex and usually takes no further part in the reaction. The 3' end of the PP is released by RNase H cleavage and might have a higher affinity for the upper site than the DNA that currently occupies it (Fig. 1e). The lower PP is arbitrarily shown here as successfully displacing the template from the upper site (Fig. 1 f); it might simply depend on which PP becomes available first. Alternatively, an innate asymmetry in the complex might influence which template displaces the other: for example, one site might be more efficient at polymerization or have a higher affinity for the 3' end of the PP than the other. Even if no template is displaced at this stage, there is a second chance when the 3' end of another, more central, PP in the HIV-1 genome is released by RNase H cleavage (Charneau et al., 1992); this increases the probability that one of the two templates is displaced and only one is copied completely.

When the PP primes synthesis, it becomes covalently attached to nascent positive strand DNA and must be excised. This presents a problem for any model involving concerted cutting and polymerization, because normally the other (template) strand is cut: either the non-template strand must be cut or, and this seems less likely, cutting is uncoupled from synthesis. If the former, then the polypurine sequence must twist the non-template strand so the RNA-DNA junction passes through the upper RNase H site with the stereochemistry appropriate for cutting (just after the stage shown in Fig. 1f). If the latter, the lower site could cut the junction much later, when positive strand DNA itself becomes a template (i.e. just after the stage shown in Fig. 1k).

The intra-molecular second jump is simply explained by annealing complementary PB sequences at each end (Fig. 1g, h). The template probably remains bound at both polymerization sites during annealing but it is difficult to illustrate this with two sites drawn as far apart as here. Therefore, to illustrate how structure (g) directly gives (j) while both 3' ends remain attached, the 3' end in the upper complex is shown firstly detaching (Fig. 1h) and then acting as a template (Fig. 1i) before reattaching (Fig. 1j). Strand-displacement synthesis at the lower site and DNA-directed DNA synthesis at the upper site then complete the duplex (Fig. 1k, l).

Recombination

A high rate of legitimate recombination between the two viral templates is another intriguing feature of retrotranscription (Coffin, 1990). Most recombination probably takes place before positive strand DNA synthesis begins (i.e. before the stage represented in Fig. 1f; Goodrich & Duesberg, 1990; Stuhlmann & Berg, 1992). Three features of the model might lead to high levels of recombination by a 'copy-choice' mechanism (Coffin, 1979, 1990): (i) the proximity of the two polymerization sites ensures that templates lie together, (ii) sites are adapted to allow invasion by new templates (at stages just before those shown in Fig. 1d or f) and (iii) templates are tethered to the complex. The 5' RNA ends are shown unattached to simplify the structure at active sites, but they are probably attached to protect them from entanglement and degradation. Then a tethered intact template (or a 3' end generated by nicking) will be able to compete more effectively for a template-binding region, especially if it anneals with the nascent DNA in the polymerization site. Attachment of two different genomes to one half-complex (as in Fig. 1a) promotes their entanglement and would be expected to lead to more recombination than attachment of each to different half-complexes; then recombination during negative strand DNA synthesis would be associated with an intermolecular first jump, as is found (Hu & Temin, 1990b). Seen in this light, template switching (and so recombination) results from other templates or nicked ends, rather than the 'proper' 3' ends of R or PP, jumping into a template-binding region.

Some recombination also occurs during synthesis of positive strand DNA, a high rate correlating with intramolecular first jumps (Hu & Temin, 1990b). The alternative structure for Fig. 1(a) (in which both the primer-binding site and 3' end of each genome are attached to the same half-complex) leads to an intramolecular first jump; then the template displaced at the stage shown in Fig. 1(f), which normally takes no further part in the reaction, could occasionally recombine with the other during positive strand synthesis.

Potential drug targets

Virus-encoded RT is a natural target for drug therapy of AIDS (Mitsuya *et al.*, 1990). The above model immediately suggests several new targets. Dimeric drugs, with the two active groups connected by a linker with appropriate stereochemistry, might be more effective than monomeric compounds; such drugs might be dimeric analogues of azido-dideoxythymidine (AZT) or tetrahydro-imidazo-benzodiazepin-one. Similarly, dimeric nucleic acid-binding compounds could bind

more specifically than their monomeric counterparts to the DNA/RNA or DNA/DNA duplexes in the complex (e.g. Fig. 1b and i). Such compounds might be bisintercalating agents with rigid linkers that point the two functional groups in the appropriate directions (Annan et al., 1992). An upper limit to the length of linker required for such molecules can be derived if the second jump occurs as shown in Fig. 1(g and j); the two polymerization sites must lie sufficiently close together to allow annealing between the two primer-binding sequences, which are 18 nucleotides long. Agents inhibiting the formation of the dimeric complex, as well as those inhibiting the formation of a heterodimer (Restle et al., 1990), might also have therapeutic value. Similarly, sites of binding to the nucleocapsid or core are potential targets, as are the attachment sites of the template RNA molecules to RT or the core. Antiviral agents like hypericin that inhibit RT in the core, but not the pure enzyme (Lavie et al., 1989), could bind to such targets. Nearly all these targets are also suggested by analogous models involving dimers that are not rotationally symmetrical or immobilized by attachment to the core.

Conclusions

This speculative model adds two features to the current model, a dimeric enzyme and attachment of the 3' ends of template RNA to the complex. Two further features, the rotational symmetry of the complex and its attachment to the virion core are optional extras. At first sight the model appears topologically complicated, but it is not markedly more so than the traditional one; we are just used to seeing a static template and a fully mobile RT that performs some remarkable jumps. Here nucleic acid strands do the pirouetting, so inevitably it takes a little time to learn the new steps, especially when we know the old ones so well. However, the model explains why the genome is dimeric, the specificity of the jumps and how recombination occurs so readily. Most importantly, it leads directly to the design of some novel antiviral drugs. We might imagine a dimeric AZT derivative to be a more potent anti-HIV agent than its monomeric counterpart, allowing the use of lower concentrations with fewer side-effects. Furthermore, monomers with too low a potency but acceptable pharmaco-kinetics might become usable drugs if dimerization increased their potency sufficiently.

References

ANNAN, N. K., COOK, P. R., MULLINS, S. T. & LOWE, G. (1992). Evidence for cross-linking DNA by bis-intercalators with rigid and extended linkers is provided by knotting and catenation. *Nucleic Acids Research* 20, 983–990.

- ArGOS, P. (1988). A sequence motif in many polymerases. Nucleic Acids Research 16, 9909–9916.
- BARAT, C., LULLIEN, V., SCHATZ, O., KEITH, G., NUGEYRE, M. T., GRUNINGER-LEITCH, BARRE-SINOUSSI, F., LEGRICE, S. F. J. & DARLIX, J. L. (1989). HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. *EMBO Journal* 8, 3279–3285.
- BEN-ARTZI, H., ZEELON, E., GORECKI, M. & PANET, A. (1992). Doublestranded RNA-dependent RNAase activity associated with human immunodeficiency virus type 1 reverse transcriptase. *Proceedings of* the National Academy of Sciences, U.S.A. 89, 927–931.
- BENDER, W., CHIEN, Y.-H., CHATTOPADHYAY, S., VOGT, P. K., GARDNER, M. B. & DAVIDSON, N. (1978). High-molecular-weight RNAs of AKR, NZB, and wild mouse viruses and avian reticuloendotheliosis viruses all have similar dimer structures. *Journal* of Virology 25, 888–896.
- BIETH, E., GABUS, C. & DARLIX, J.-L. (1990). A study of the dimer formation of Rous sarcoma virus RNA and of its effect on viral protein synthesis *in vitro*. *Nucleic Acids Research* 18, 119–127.
- BOONE, L. R. & SKALKA, A. M. (1981). Viral DNA synthesis in vitro by avian retrovirus particles permeabilized with melittin: I. Kinetics of synthesis and size of minus- and plus-strand transcripts. *Journal of Virology* 37, 109–116.
- BOWERMAN, B., BROWN, P. O., BISHOP, J. M. & VARMUS, H. E. (1989). A nucleoprotein complex mediates the integration of retroviral DNA. Genes and Development 3, 469–478.
- BURNS, N. R., SAIBIL, H. R., WHITE, N. S., PARDON, J. F., TIMMINS, P. A., RICHARDSON, S. M. H., RICHARDS, B. M., ADAMS, S. E., KINGSMAN, S. M. & KINGSMAN, A. J. (1992). Symmetry, flexibility and permeability in the structure of yeast retrotransposon viruslike particles. *EMBO Journal* 11, 1155–1164.
- CHARNEAU, P. C., ALIZON, M. & CLAVEL, F. (1992). A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. *Journal of Virology* 66, 2814–2820.
- COFFIN, J. M. (1979). Structure, replication and recombination of retrovirus genomes: some unifying hypotheses. *Journal of General* Virology 42, 1–26.
- COFFIN, J. M. (1990). Genetic variation in retroviruses. Applied Virology Research 2, 11-33.
- COOK, P. R. (1989). The nucleoskeleton and the topology of transcription. European Journal of Biochemistry 185, 487-501.
- Cook, P. R. (1991). The nucleoskeleton and the topology of replication. *Cell* **66**, 627–635.
- CRAIGIE, R., FUJIWARA, T. & BUSHMAN, F. (1990). The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. Cell 62, 829–837.
- FIELDS, B. N. (1990). Virology, vol II. New York: Raven Press.
- FU, T.-B. & TAYLOR, J. (1992). When retroviral reverse transcriptases reach the end of their RNA templates. *Journal of Virology* 66, 4271–4278.
- FURFINE, E. S. & REARDON, J. E. (1991). Reverse transcriptase. RNAse H from the human immunodeficiency virus: relationship of the DNA polymerase and RNA hydrolysis activities. *Journal of Biological Chemistry* 266, 406–412.
- GILBOA, E., MITRA, S. W., GOFF, S. & BALTIMORE, D. (1979). A detailed model of reverse transcription and tests of crucial aspects. *Cell* 18, 93-100.
- GOODRICH, D. W. & DUESBERG, P. H. (1990). Retroviral recombination during reverse transcription. *Proceedings of the National Academy* of Sciences, U.S.A. 87, 2052–2056.
- HU, W.-S. & TEMIN, H. M. (1990 a). Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proceedings of the National* Academy of Sciences, U.S.A. 87, 1556–1560.
- Hu, W.-S. & TEMIN, H. M. (1990b). Retroviral recombination and reverse transcription. Science 250, 1227-1233.
- HUBER, H. E. & RICHARDSON, C. C. (1990). Processing of the primer for plus strand DNA synthesis by human immunodeficiency virus 1 reverse transcriptase. Journal of Biological Chemistry 265, 10565-10573.

- JACOBO-MOLINA, A. & ARNOLD, E. (1991). HIV reverse transcriptase structure-function relationships. *Biochemistry* 30, 6351–6361.
- KATZ, R. A., MERKEL, G., KULKOSKY, J., LEIS, J. & SKALKA, A. M. (1990). The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. *Cell* 63, 87–95.
- KOHLSTAEDT, L. A., WANG, J., FRIEDMAN, J. M., RICE, P. A. & STEITZ, T. A. (1992). 3.5 Angstrom resolution crystal structure of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* 256, 1783–1790.
- LAVIE, G., VALENTINE, F., LEVIN, B., MAZUR, Y., GALLO, G., LAVIE, D., WEINER, D. & MERUELO, D. (1989). Studies on the mechanisms of action of the antiretroviral agents hypericin and pseudohypericin. *Proceedings of the National Academy of Sciences*, U.S.A. 86, 5963-5967.
- LAYNE, S. P., MERGES, M. J., DEMBO, M., SPOUGE, J. L., CONLEY, S. R., MOORE, J. P., RAINA, J. L., RENZ, H., GELDERBLOM, H. R. & NARA, P. L. (1992). Factors underlying spontaneous inactivation and susceptibility to neutralization of human immunodeficiency virus. Virology 189, 695-714.
- LUO, G. & TAYLOR, J. (1992). Template switching by reverse transcriptase during DNA synthesis. *Journal of Virology* 64, 4321-4328.
- MITSUYA, H., YARCHOAN, R. & BRODER, S. (1990). Molecular targets for AIDS therapy. *Science* 249, 1533–1544.
- OYAMA, F., KIKUCHI, R., CROUCH, R. J. & UCHIDA, T. (1989). Intrinsic properties of reverse transcriptase in reverse transcription. *Journal of Biological Chemistry* 264, 18808–18817.
- PANGANIBAN, A. T. & FIORE, D. (1988). Ordered interstrand and

intrastrand DNA transfer during reverse transcription. Science 241, 1064–1069.

- PRATS, A.-C., ROY, C., WANG, P., ERARD, M., HOUSSET, V., GABUS, C., PAOLETTI, C. & DARLIX, J.-L. (1990). *cis* elements and *trans*-acting factors involved in dimer formation of murine leukemia virus RNA. *Journal of Virology* 64, 774–783.
- RESTLE, T., MULLER, B. & GOODY, R. S. (1990). Dimerization of human immunodeficiency virus type 1 reverse transcriptase: a target for chemotherapeutic intervention. *Journal of Biological Chemistry* 265, 8986–8988.
- STUHLMANN, H. & BERG, P. (1992). Homologous recombination of copackaged retrovirus RNAs during reverse transcription. *Journal of* Virology 66, 2378–2388.
- VARMUS, H. E., HEASLEY, S., KUNG, H.-J., OPPERMAN, H., SMITH, V. C., BISHOP, J. M. & SHANK, P. R. (1978). Kinetics of synthesis, structure and purification of avian sarcoma virus-specific DNA made in the cytoplasm of acutely infected cells. *Journal of Molecular Biology* 120, 55-82.
- VERONESE, F. DI M., COPELAND, T. D., DEVICO, A. L., RAHMAN, R., OROSZLAN, S., GALLO, R. C. & SARNGADHARAN, M. G. (1986). Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. Science 231, 1289–1291.
- WOHRL, B. M. & MOELLING, K. (1990). Interaction of HIV-1 ribonuclease H with polypurine tract containing RNA-DNA hybrids. *Biochemistry* 29, 10141–10147.
- (Received 22 September 1992; Accepted 24 November 1992)