The EMBO Journal Vol.18 No.4 pp.1038–1048, 1999

Structural basis for the specificity of the initiation of HIV-1 reverse transcription

Catherine Isel[1,2], Eric Westhof[1], Christian Massire[1], Stuart F.J.Le Grice[3], Bernard Ehresmann[1], Chantal Ehresmann[1] and Roland Marquet[1,4]

[1]Unité Propre de Recherche No. 9002 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg cedex, France and [3]Division of Infectious Diseases, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

[2]Present address: MRC, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

[4]Corresponding author
e-mail: marquet@ibmc.u-strasbg.fr

Initiation of human immunodeficiency virus type 1 (HIV-1) reverse transcription requires specific recognition of the viral genome, $tRNA_3^{Lys}$, which acts as primer, and reverse transcriptase (RT). The specificity of this ternary complex is mediated by intricate interactions between HIV-1 RNA and $tRNA_3^{Lys}$, but remains poorly understood at the three-dimensional level. We used chemical probing to gain insight into the three-dimensional structure of the viral RNA–$tRNA_3^{Lys}$ complex, and enzymatic footprinting to delineate regions interacting with RT. These and previous experimental data were used to derive a three-dimensional model of the initiation complex. The viral RNA and $tRNA_3^{Lys}$ form a compact structure in which the two RNAs fold into distinct structural domains. The extended interactions between these molecules are not directly recognized by RT. Rather, they favor RT binding by preventing steric clashes between the nucleic acids and the polymerase and inducing a viral RNA–$tRNA_3^{Lys}$ conformation which fits perfectly into the nucleic acid binding cleft of RT. Recognition of the 3′ end of $tRNA_3^{Lys}$ and of the first template nucleotides by RT is favored by a kink in the template strand promoted by the short junctions present in the previously established secondary structure.

Keywords: AIDS/DNA polymerase/RNA/reverse transcriptase/structure

Introduction

Reverse transcriptase (RT) is a key enzyme of the retroviral cycle that possesses RNA- and DNA-dependent DNA polymerase and RNase H activities (Baltimore, 1970; Temin and Mizutani, 1970). It converts the single-stranded RNA genome into double-stranded DNA via a complex series of steps including two template switches and duplication of the long terminal repeats (Gilboa *et al*., 1979). In retroviruses, plant pararetroviruses and most retrotransposons, initiation of reverse transcription is primed by a cellular tRNA whose 3′ end is complementary to a so-called 'primer binding site' (PBS) (for reviews see Marquet *et al*., 1995; Mak and Kleiman, 1997; Arts and Le Grice, 1998). $tRNA_3^{Lys}$ is the natural primer of all immunodeficiency viruses, including the type 1 human immunodeficiency virus (HIV-1), while $tRNA^{Trp}$ and $tRNA^{Pro}$ are used by most avian and murine retroviruses, respectively (Marquet *et al*., 1995; Mak and Kleiman, 1997; Arts and Le Grice, 1998).

In HIV-1, the PBS is not the sole determinant of primer usage. When the PBS is mutated to be complementary to the 3′ end of tRNAs other than $tRNA_3^{Lys}$, replication of the virus is dramatically reduced and rapid reversion to the wild-type PBS is observed (Li *et al*., 1994; Das *et al*., 1995; Wakefield *et al*., 1995). Similarly, avian leukosis viruses whose PBS identity is altered are unstable and revert to a PBS complementary to $tRNA^{Trp}$ (Whitcomb *et al*., 1995). In contrast, murine retroviruses are able to stably maintain a PBS complementary to an unnatural primer tRNA (Colicelli and Goff, 1987; Lund *et al*., 1993, 1997). The selective primer usage of HIV-1 correlates with the ability of HIV-1 RT to discriminate against non-self tRNA primers *in vitro* (Li *et al*., 1994; Oude Essink *et al*., 1996).

The specificity of the initiation of reverse transcription appears directly linked to virus-specific interactions between the primer tRNA and the genomic RNA, in addition to the general PBS–tRNA interaction. Mutagenesis and genetic studies on Rous sarcoma virus (RSV) support the existence of a 7-bp interaction between the TΨC stem–loop of $tRNA^{Trp}$ and the avian retroviral RNA that is required for optimal reverse transcription *in vitro* and *in vivo* (Cordell *et al*., 1979; Cobrinik *et al*., 1988; Aiyar *et al*., 1992, 1994). In contrast to avian retroviruses, structural probing suggested that the interaction between $tRNA^{Pro}$ and the genomic RNA of Moloney murine leukaemia virus is limited to the 'classical' interaction between the PBS and the 3′ end of the primer (Fossé *et al*., 1998).

In HIV-1, chemical and enzymatic probing revealed complex interactions between the anticodon loop, the 3′ portion of the anticodon stem and part of the variable loop of $tRNA_3^{Lys}$ with viral sequences located upstream from the PBS (Isel *et al*., 1993, 1995) (Figure 1). These interactions, the stability of which requires the post-transcriptional modifications of $tRNA_3^{Lys}$ (Isel *et al*., 1993), are supported by site-directed mutagenesis (Isel *et al*., 1996, 1998). Post-transcriptional modifications of $tRNA_3^{Lys}$ and extended primer–template interactions are also required for efficient initiation and the transition to elongation of HIV-1 reverse transcription (Isel *et al*., 1996). The primer–template complex is specifically recognized by the homologous enzyme: most non-homologous RT, including lentiviral RTs, which naturally utilize $tRNA_3^{Lys}$

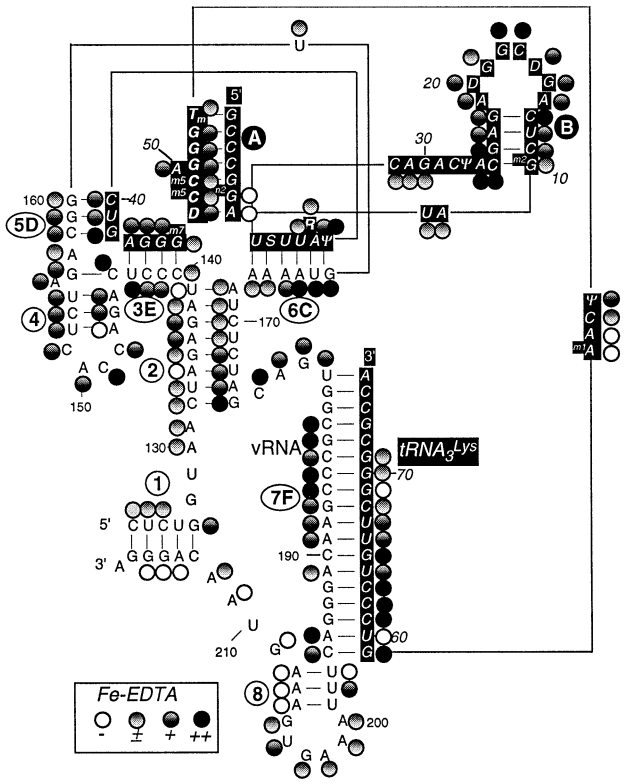


Fig. 1. Probing of the HIV-1 (nt 1–311) vRNA–$tRNA_3^{Lys}$ complex with hydroxyl radicals. The intensities of cleavage by hydroxyl radicals at each position of the primer and of the template were quantified using a Fuji Bas2000 BioImager. The experimental values were classified into very weak, weak, medium and strong reactivities and represented on the secondary structure model of the primer–template complex (Isel *et al*., 1995). The tRNA sequence is in white on a black background. The region of vRNA corresponding to nt 123–217 of the HIV-1 genome (MAL isolate) is represented. Helices are numbered according to Isel *et al*. (1995). The absence of symbols indicates the residues whose reactivity could not be determined.

as primer, are unable to efficiently extend the HIV-1 binary complex (Arts *et al*., 1996; Isel *et al*., 1996).

The interaction between the anticodon loop of $tRNA_3^{Lys}$ and the viral A-rich loop recently gained support from *in vivo* studies. First, a $tRNA_3^{Lys}$ with a mutated anticodon is incorporated into the HIV-1 particles, but is not used as primer for reverse transcription (Huang *et al*., 1996). Secondly, deletion of the A-rich loop results in viruses with diminished levels of infectivity and reduced DNA synthesis, and this loop is progressively restored upon prolonged cell culture (Liang *et al*., 1997). Furthermore, HIV-1 can stably use $tRNA^{His}$ and $tRNA^{Met}$ as primers, provided that both the PBS and the A-rich loop are simultaneously mutated to be complementary to these tRNA species (Wakefield *et al*., 1996; Kang *et al*., 1997; Zhang *et al*., 1998).

Knowledge of the three-dimensional structure of the initiation complex is required to define the molecular basis for the specificity of HIV-1 reverse transcription. Here, we used chemical probing to gain insight into the tertiary structure of the HIV-1 RNA–$tRNA_3^{Lys}$ complex and enzymatic footprinting to determine regions of the primer–template interacting with HIV-1 RT. These and previous experimental data (Isel *et al*., 1995, 1998) were used to construct three-dimensional models of the viral RNA–$tRNA_3^{Lys}$ and viral RNA–$tRNA_3^{Lys}$:RT complexes. Our models highlight the role of the extended primer–template interactions in primer–template folding and in RT binding, revealing viral RNA–$tRNA_3^{Lys}$:RT interactions specific to the initiation phase.

Results and discussion

Probing the tertiary structure of the HIV-1 RNA–$tRNA_3^{Lys}$ complex

The secondary structure model shown in Figure 1 is supported by extensive chemical and enzymatic probing (Isel *et al*., 1993, 1995) and by site-directed mutagenesis (Isel *et al*., 1998). In order to provide more information on the tertiary structure of the HIV-1 RNA–$tRNA_3^{Lys}$ complex, we tested accessibility of the N7 position of purines and of the C4′ position of riboses. The A-N7 and G-N7 positions were tested with diethyl pyrocarbonate (DEPC) and dimethyl sulfate (DMS), respectively (data not shown). The A-N7 positions are usually protected from DEPC modification when they are located within a regular RNA helix (Moine *et al*., 1997). Thus, the absence of modification of A23 in the $tRNA_3^{Lys}$ and of A134, -138, -145, -165, -166, -188, -189 and -214 in the viral RNA (vRNA) supports the secondary structure model. The absence of reactivity of A-N7 at helix termini (i.e. A7, -43, -147, -167 and -208) and in single-stranded regions (i.e. A26, -57 and -157) may reflect base stacking, hydrogen bonding or metal-ion binding, or burying inside the tertiary structure. Reactivity of the N7 position of A206 and -207 reflects limited stability of helix 8.

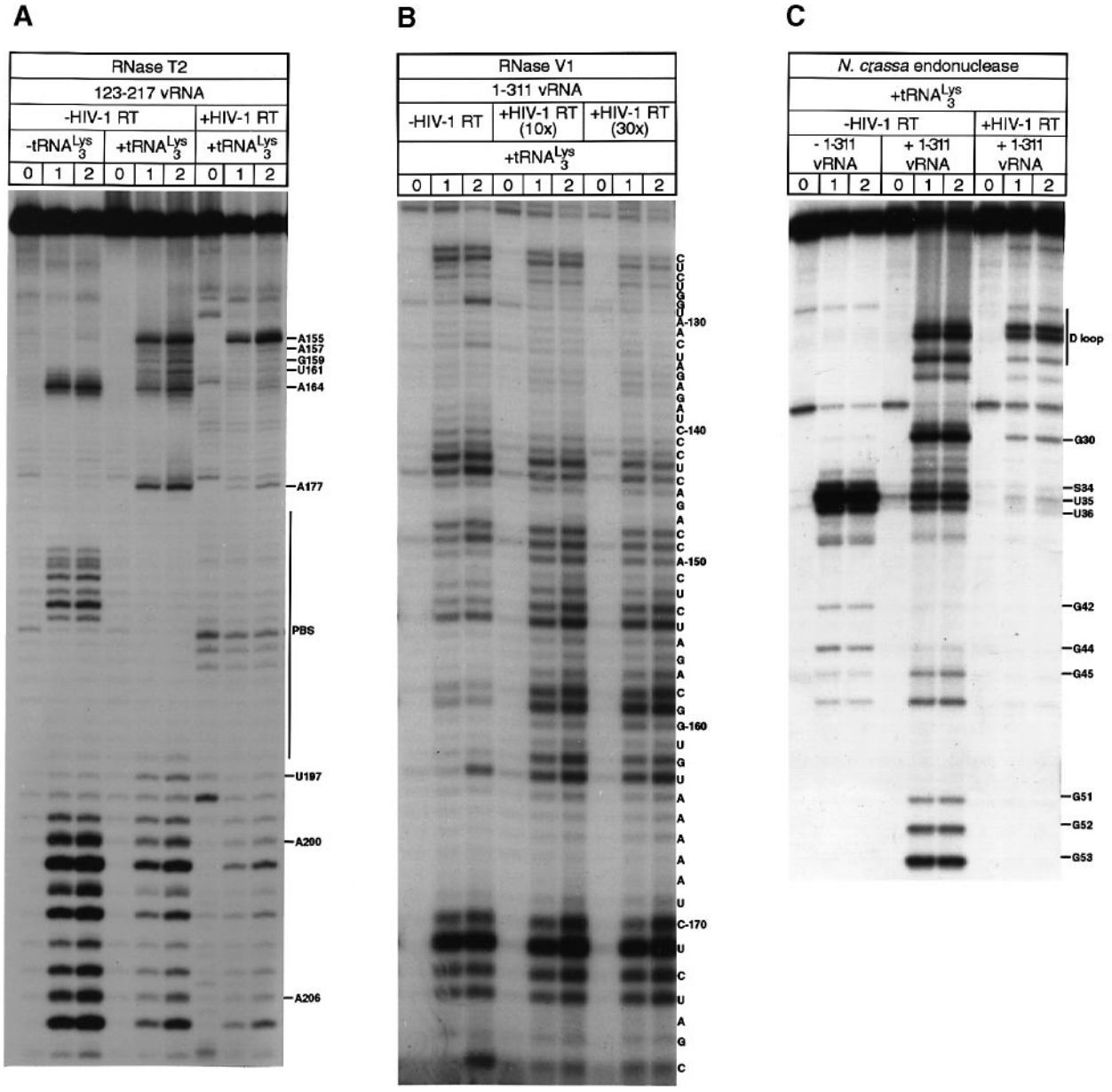
Contrary to that of A-N7, the N7 reactivity of paired Gs cannot be predicted from the secondary structure (Moine *et al*., 1997). However, the absence of reactivity of G-N7 in single-stranded regions would indicate hydrogen bonds or metal ion chelation involving these positions (Moine *et al*., 1997). No such site was identified in the vRNA–$tRNA_3^{Lys}$ complex. The N7 position of G160 was protected from modification, either because it is stacked on helix 5D or because this nucleotide (nt) is buried inside the structure (data not shown).

We used hydroxyl radicals generated by Fe(II)-EDTA to test accessibility of the riboses of the vRNA and primer tRNA. This probe can distinguish the interior and exterior parts of large RNA molecules (Latham and Cech, 1989; Celander and Cech, 1991), and can detect non-canonical interactions (Murphy and Cech, 1994) and metal-ion binding sites involving the sugar–phosphate backbone (Serganov *et al*., 1996). In some short helices (e.g. 1 and 8), only one strand was accessible to hydroxyl radical attack (Figure 1). In longer, regular helices (2 and 7F), reactivity varied along the helix axis and was shifted on one strand compared with the other. The entire stem–loops B and 4 were accessible, suggesting there are located at the periphery of the primer–template tertiary structure (Figure 1). We also tested susceptibility of vRNA encompassing nt 123–217 of the HIV-1 genome towards Fe(II)-EDTA in the binary complex with $tRNA_3^{Lys}$ and obtained similar results (data not shown). Not surprisingly, almost all previously reported enzymatic cleavages (Isel *et al*., 1995) are located in regions cleaved by Fe(II)-EDTA.

Footprinting of the HIV-1 RT on the vRNA–tRNA$_3^{Lys}$ complex

Hydroxyl radicals (Metzger *et al*., 1993) and nucleases (Wöhrl *et al*., 1995) have been used to map HIV-1 RT on DNA–DNA primer–template. However, despite numerous attempts in different laboratories, no RT footprint has so far been published using an RNA template. When using either Fe(II)-EDTA or base-modifying reagents, we were unable to detect any footprint of RT on the vRNA–tRNA$_3^{Lys}$ complex. Nevertheless, the absence of reactivity changes of the vRNA–tRNA$_3^{Lys}$ complex in the presence of RT suggests that the secondary structure of the primer–template complex was not significantly modified. In particular, there was no indication that the intermolecular interactions were affected by RT binding.

In contrast, enzymatic probes were successful in monitoring binding of HIV-1 RT to the vRNA–tRNA$_3^{Lys}$ complex, using either 1–311 or 123–217 vRNA templates (Figure 2; data not shown). In order to optimize binding of RT to the vRNA–tRNA$_3^{Lys}$ complex, the potassium chloride concentration was reduced from 300 mM, when probing the structure of the primer–template complex, to 40 mM. When free vRNA 123–217 was submitted to digestion with RNase T2, an endonuclease that cleaves single-stranded regions with a preference for adenines, major cleavage sites were observed in the A-rich loop (at A164), in the PBS (nt 182–188) and in stem–loop 8, downstream of the PBS (nt 197–208) (Figure 2A). Upon formation of the vRNA–tRNA$_3^{Lys}$ complex, PBS susceptibility totally disappeared, those at nt 164 and 197–208 strongly decreased, and new strong cleavages were observed at A155 and A177 (Figure 2A), two bases that are unpaired in the binary complex (Figure 1). Reduced cleavage at A164 results from base-pairing of the anticodon loop of tRNA$_3^{Lys}$ with the viral A-rich loop (helix 6C), which is a dynamic interaction that does not totally protect RNA from nuclease cleavage and chemical modification (Isel *et al*., 1995, 1998). The strongest cleavages in stem–loop 8 were located in the loop (Figure 2A). However, cleavage was also observed in the stem, indicating that



the stability of this helix is limited, in agreement with modification of A205 and 206 by DEPC (data not shown). Upon binding of RT, cleavage at A155 was enhanced, while A164, A177 and nt 197–208 were efficiently protected from hydrolysis by RNase T2 (Figure 2A and E). Reduced cleavage at single-stranded nt 177 and 200–205 most probably reflects a direct protection of the vRNA by RT. For nt 164, 197–199 and 206–208, which are involved in unstable helices, reduced cleavage might be a consequence of the stabilization of these helices by RT. The enhanced cleavage at A155 probably reflects a slight RT-induced change in the relative orientation of helices 4 and 5D.

When the primer–template complex was subjected to treatment with RNase V1, an endonuclease that specifically cleaves double-stranded and stacked regions, cleavage

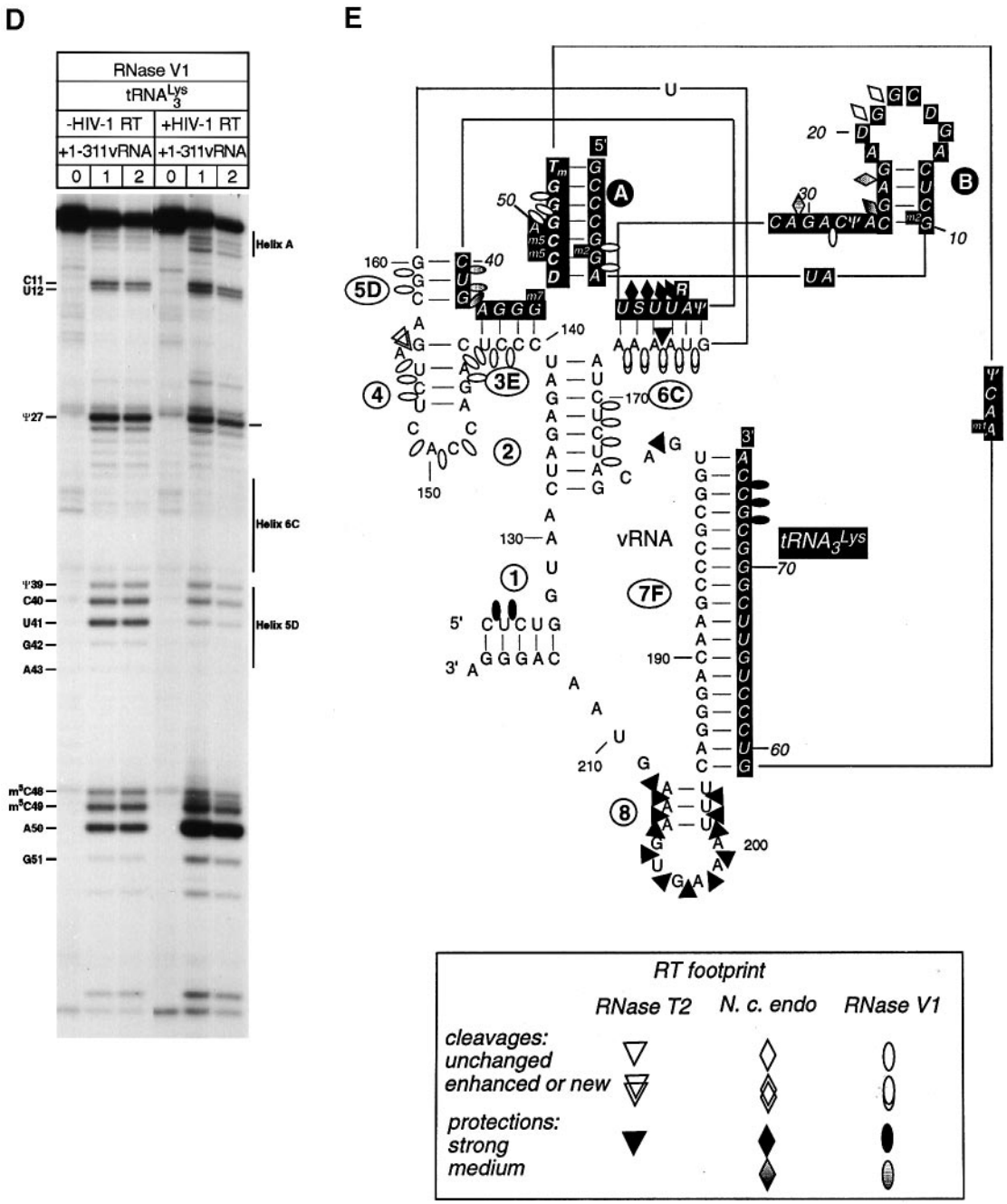


Fig. 2. Footprinting of the HIV-1 RT on the vRNA–$tRNA_3^{Lys}$ complex. (**A**) 3′ end-labeled vRNA 123–217 either free or bound to $tRNA_3^{Lys}$ in the presence or absence of RT was probed with RNase T2. Lanes 0, 1 and 2 correspond to controls incubated in the absence of nuclease and reactions for 5 and 10 min, respectively. Note that RT induced cleavage at position C190 of the vRNA, even in the absence of RNase T2. (**B**) vRNA 1–311 involved in the primer–template complex was probed with RNase V1 in the absence and presence of 10- or 30-fold excess of RT. Lanes 0, 1 and 2 correspond to controls incubated in the absence of nuclease and reactions for 8 min with 0.15 and 0.22 units RNase V1, respectively. (**C**) 3′ end-labeled $tRNA_3^{Lys}$, either free or annealed to vRNA in the presence and absence of RT was probed with the *N.crassa* endonuclease. Lanes 0, 1 and 2 correspond to controls incubated in the absence of nuclease and reactions for 5 and 10 min, respectively. (**D**) Bound $tRNA_3^{Lys}$ was probed with RNase V1 in the absence and presence of RT. Lanes 0, 1 and 2 correspond to controls incubated in the absence of nuclease and reactions for 5 and 10 min, respectively. (**E**) Summary of the footprinting data.

sites in vRNA were located in helix 1 (nt 123–126), helix 3E (nt 141–143), helix 4 (nt 144–145 and 152–154), helix 5D (nt 158–159), helix 6C (weak cleavage at nt 162–164) and helix 2 (nt 170–173) (Figure 2B and E). Cleavage was also observed in the loop of stem–loop 4 (nt 148–150), probably reflecting base-stacking of these residues (Isel *et al*., 1995). Upon addition of increasing amounts of RT, RNase V1 cleavages in helices 6C and 5D were strongly enhanced and nt A164–166 became susceptible to hydrolysis. Weak reactivity changes were observed in stem–loop 4 and helix 3E, while cleavage in helix 2 was unaffected. These data indicate that extended intermolecular interactions formed by helices 3E, 5D and 6C are not disrupted upon RT binding. On the contrary, enhanced cleavages in helices 5D and 6C suggest these interactions are stabilized by RT, even though there is no direct contact of these helices with the protein. Enhanced cleavage in these helices also indicates that protection of A164 from RNase T2 cleavage by RT was not due to a direct contact with RT, but to the stabilization of helix 6C. The only protection from RNase V1 cleavage by RT was observed at nt 123–124 (Figure 2B and E), suggesting contacts between the polymerase and helix 1, or at least spatial proximity.

Neurospora crassa endonuclease, which cleaves single-stranded regions, and RNase V1 were used to identify contacts between RT and $tRNA_3^{Lys}$ engaged in the primer–template complex. In the absence of RT, the cleavage pattern of $tRNA_3^{Lys}$ by *N.crassa* endonuclease in the footprint buffer was very similar to that observed previously at higher salt concentrations (Isel *et al*., 1995, 1998) (Figure 2C). However, protection of the anticodon loop of the tRNA in the primer–template was weaker in the footprint buffer and new cuts were observed in the dihydrouridine stem (helix B) (Figure 2C), indicating the limited stability of helices 6C and B at low ionic strength. When forming the ternary primer–template:RT complex, hydrolysis of the B loop was still observed, while other sites were partially (G22, G24 and G30) or totally (33USUU36) protected (Figure 2C and E). As already observed, protection of the anticodon loop of $tRNA_3^{Lys}$ probably reflects stabilization of helix 6C rather than a direct protection by RT. The same may hold true for helix B. In contrast, protection of G30 probably results from a direct protection by the retroviral polymerase.

In the primer–template complex, $tRNA_3^{Lys}$ was cleaved by RNase V1 in both strands of helix A, at C28, in helix 5D and at the 3′ end of helix 7F (Figure 2D and E). In the ternary complex, protections were observed in helices 7F and 5D, while increased cleavage was observed in the 3′ strand of helix A (Figure 2D and E). Protection of the 3′ end of $tRNA_3^{Lys}$ hybridized to the PBS was not complete, probably due to the dynamic nature of the tertiary complex (Lanchy *et al*., 1996). Only the primer strand of helix 5D was protected by RT, since as indicated above, the viral strand remained susceptible to RNase V1 cleavage in the tertiary complex (Figure 2B).

Taken together, our footprinting data suggest that, in the ternary complex, RT protects helix 1, helix 7F, stem–loop 8, and the single-stranded nt G30 in the primer and A177 in the template. At this stage, it remains unclear whether partial protection of the viral strand of helix 5D is direct or indirect (see below).

Modeling the three-dimensional structure of the HIV-1 RNA–$tRNA_3^{Lys}$ complex

In the first modeling step, the elements of tertiary structure were constructed and assembled according to Figure 1. For simplicity, the post-transcriptional modifications of $tRNA_3^{Lys}$ were not represented. As expected, the multiple short connections between helices strongly constrain the three-dimensional structure of the primer–template complex and limit the number of possible arrangements of the helices. Assembly required only minor modifications of the secondary structure: residue R37 of $tRNA_3^{Lys}$, rather than A38, was base-paired to U163 in the vRNA, residues A38 and G162 formed a non-canonical base pair and Ψ39 was unpaired. These slight modifications of the secondary structure of the vRNA–$tRNA_3^{Lys}$ complex are compatible with the probing data (Isel *et al*., 1995). The modeling of the binary complex was performed independently by two of us (E.W. and C.M.) and converged towards the same tertiary structure, within the approximations linked to the method. In addition, we were able to model intermediate states that could constitute the successive steps of tRNA annealing to the vRNA *in vivo* (C.Massire and E.Westhof, unpublished). This series of steps led to the three-dimensional model presented here, indicating that formation of this structure is topologically feasible. The constraints derived from the probing of R-N7 positions and ribose residues were then introduced in the model.

There is good correlation between the reactivity of $tRNA_3^{Lys}$ and vRNA riboses towards Fe-EDTA and accessibility of the ribose C4′ positions calculated from our three-dimensional structure model of the vRNA–$tRNA_3^{Lys}$ complex (Figure 3). The protected regions of $tRNA_3^{Lys}$ (helix A and the junction between helices A and B, the junction between helices B and 6C and helix 6C, the junction between A and 7F and the 3′ portion of helix 7F) are located in domains of low calculated accessibility. On the other hand, the 5′ portion of helix 7F is highly reactive, in agreement with our model (Figure 3). The protected regions of vRNA correspond to domains of low calculated accessibility (the junction between helices 1 and 2, part of the 5′ strand of helix 2, the junction between helices 5D and 6C, the helix 6C–helix 2 junction, the 3′ part of helix 8 and the 3′ strand of helix 1) or to a local minimum in the calculated curve (the junction of helices 2 and 3E). As expected, the most exposed regions of the model are highly reactive towards hydroxyl radicals (Figure 3). The only discrepancies between the experimental and calculated accessibility are located at the junction of helices 4 and 5D and helix 7F that are predicted regions of low accessibility, but contain several moderately reactive riboses. Indeed, calculations made on a rigid structure might not correctly account for the dynamic structure of these regions of the vRNA–$tRNA_3^{Lys}$ complex. Similar discrepancies were observed between the backbone accessibilities of the P4–P6 domain of a group I intron that were determined biochemically and calculated from the crystallographic structure (Cate *et al*., 1996).

The three-dimensional model also correctly accounts for the protection observed at N7 positions. Most N7 positions are protected by base-stacking between the 5′ base-paired nucleotide and the 3′ nucleotide, which is either base-paired (A145 and A167) or single-stranded (G30, A7, A147 and A208). Protection of A26-N7 and

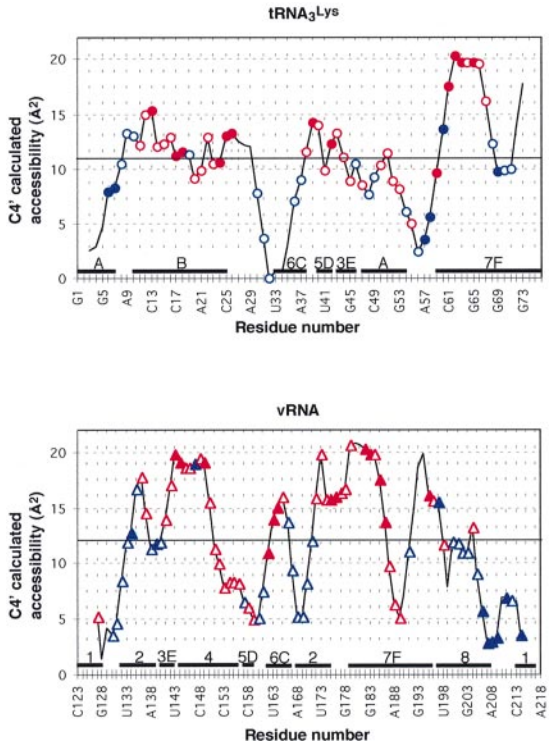


Fig. 3. Calculated and experimental accessibilities of $tRNA_3^{Lys}$ (top) and vRNA (bottom) in the binary complex. The theoretical accessibility (solid line) of the C4′ atoms was calculated from the three-dimensional model of the vRNA–$tRNA_3^{Lys}$ complex with the program ACCESS (Richmond, 1984) using a 2.8 Å radius probe. Filled blue, open blue, open red and filled red symbols correspond to very weak, weak, medium and strong cleavage by Fe(II)-EDTA, respectively.

A57-N7 are due to base-stacking with single-stranded nucleotides, while that of A43 and A157 involves spatial proximity of nucleotides not located immediately upstream or downstream in the primary structure (not shown).

Despite intermolecular interactions involving several distinct helices and the complex central junction of the secondary structure model, the tertiary structure model is remarkably simple. Indeed, except for the interaction of the 3′ end of $tRNA_3^{Lys}$ with the PBS, the two RNA molecules form distinct and compact domains (Figure 4). Intermolecular helices 3E, 5D and 6C form a single interface explaining why helices 3E and 5D are unstable when helix 6C is mutated (Figure 4A) (Isel *et al*., 1998). At the central four-helical junction (Figure 1), helices 2 and 6C are stacked and nearly coaxial, while helices 3E and A are not. Helix 7F is exposed at the periphery of the structure, thus allowing recognition by RT. The three single-stranded template nucleotides upstream of the PBS were modeled in continuity with this helix and there is a sharp turn between the third nucleotide and helix 2. The sharp kink between helices 7F and 2 is a direct consequence of the secondary structure of the complex. Modeling of these single-stranded nucleotides was based on analogy with the MoMLV primer–template complex, in which the corresponding three nucleotides are more constrained (Fossé *et al*., 1998). Helix 8 extends helix 7F, even though these are not colinear. Helices 2, 3E, 4, 5D and 6C form a flat surface parallel to helix 7F, and helices A and B protrude on the other side of the complex (Figure 4B).

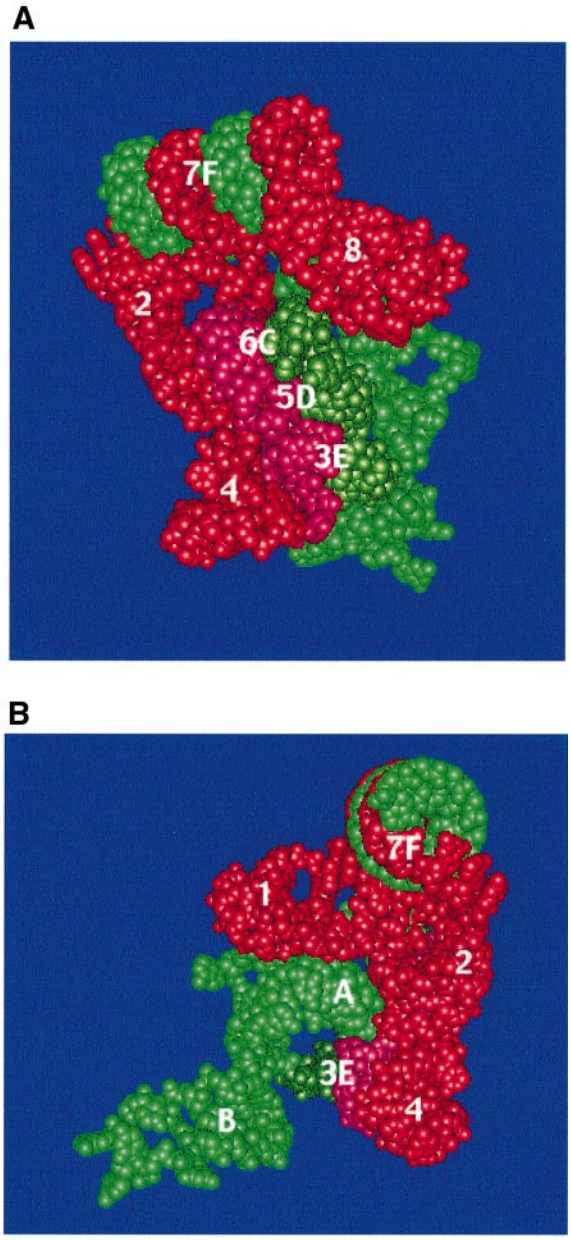


Fig. 4. Space-filling three-dimensional structure model of the vRNA–$tRNA_3^{Lys}$ complex. vRNA is in red and the tRNA in green, except for the regions of those RNAs involved in the extended intermolecular interactions (helices 3E, 5D and 6C, which are in purple and dark green, respectively). In view (**A**), the 3′ end of $tRNA_3^{Lys}$ is at the left end of helix 7F, while it is pointing towards the reader in view (**B**), in which the model was turned by ~120° around the vertical axis as compared with (A). Helices are labeled according to Figure 1.

Modeling the three-dimensional structure of the HIV-1 vRNA–$tRNA_3^{Lys}$:RT complex

When the enzymatic footprint data of RT were superimposed on the three-dimensional model of the vRNA–$tRNA_3^{Lys}$ complex, the protected and accessible regions clearly formed distinct domains consistent with binding of the 3′ terminus of the primer in the polymerase active site (Figure 5A). Thus, these data can be used to elaborate a three-dimensional model of the HIV-1 vRNA–$tRNA_3^{Lys}$:RT complex. This tertiary complex was built by directly docking the primer–template model on the crystal structure of HIV-1 RT solved by Arnold and co-workers

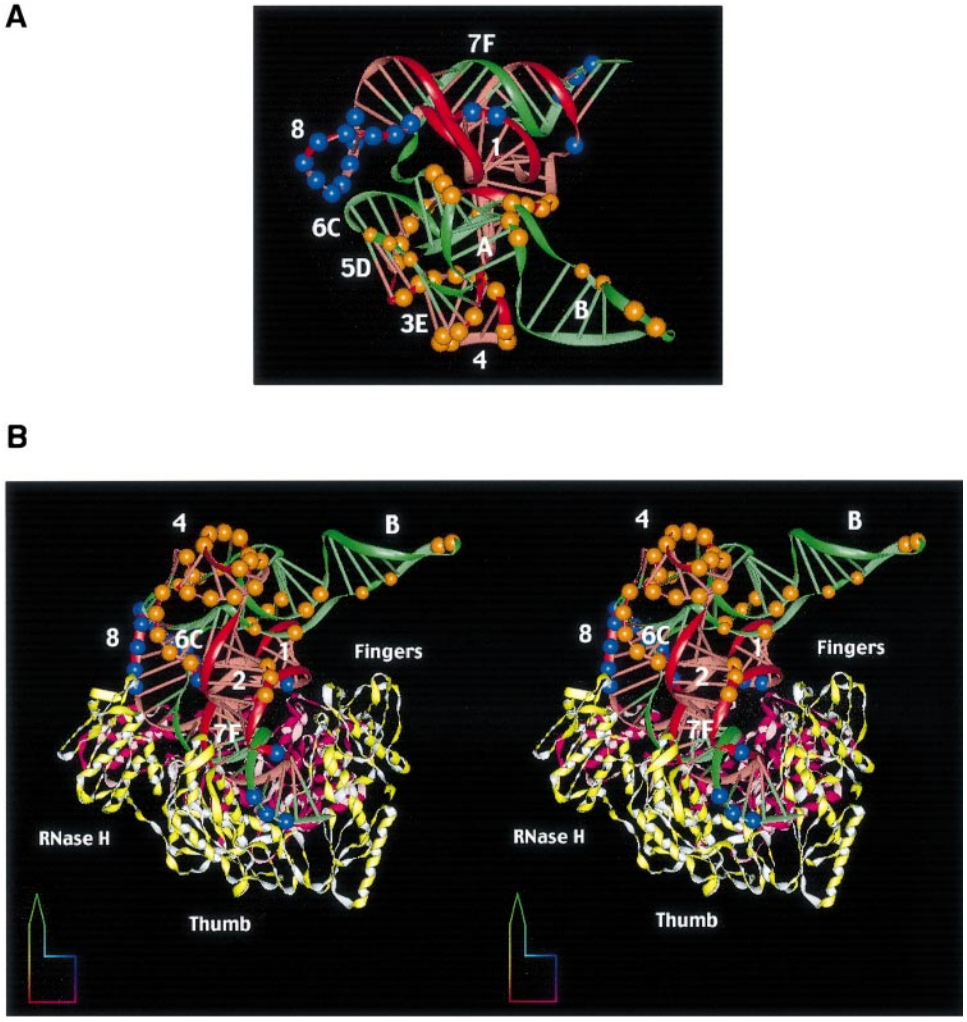


Fig. 5. Modeling the three-dimensional structure of the vRNA–$tRNA_3^{Lys}$:RT complex. (**A**) Representation of the RT footprint on the vRNA–$tRNA_3^{Lys}$ model. Red and green ribbons correspond to the vRNA and $tRNA_3^{Lys}$, respectively. The 3′ end of $tRNA_3^{Lys}$ is at the right end of helix 7F. Sites strongly protected against cleavage by RT are represented by blue spheres. Unaffected and enhanced cleavage sites are indicated by golden spheres and weak protections by small golden spheres. (**B**) Stereo view of the vRNA–$tRNA_3^{Lys}$:RT complex. The p66 subunit of RT is in yellow and p51 in purple. The subdomains of the p66 subunit of HIV-1 RT are labeled according to Kohlstaedt *et al.* (1992).

(Jacobo-Molina *et al.*, 1993; Ding *et al.*, 1997). The phosphate 5′ of the 3′ terminal A of $tRNA_3^{Lys}$ was superimposed with the corresponding phosphate of the DNA primer present in the crystallographic structure. The orientation of the vRNA–$tRNA_3^{Lys}$ complex was then modified to avoid steric clashes with the polymerase. During these processes, the primer–template complex and the polymerase were both considered as rigid bodies and no modification of their structures was allowed. Chemical and enzymatic footprinting data suggested that the secondary structure of the vRNA–$tRNA_3^{Lys}$ complex was unaffected by RT binding and that the tertiary structure underwent no dramatic change. The possibility that the RT structure might slightly differ in the initiation complex and the crystal structure solved by Arnold and coworkers cannot be ruled out. Rotation of the thumb (Esnouf *et al.*, 1995) or a swivel motion of RNase H domain (Jager *et al.*, 1994) would locally modify the interactions between lateral chains of RT and the primer–template complex; however, they would not change the general architecture of the initiation complex and the consequences discussed here.

The vRNA–$tRNA_3^{Lys}$ complex fits perfectly into the binding cleft of RT and the docking easily accounts for most of the footprinting data (Figure 5). The 3′ adenosine of $tRNA_3^{Lys}$ is in the polymerase catalytic site and C196 on the template, which is located 18 nt downstream from the initiation site of reverse transcription, is close to the catalytic residues of the RNase H domain of HIV-1 RT (Figure 5B). The interaction between the RNase H domain of HIV-1 RT and the vRNA–$tRNA_3^{Lys}$ complex extends to stem–loop 8 of the template, explaining protection of nt 197–208 from cleavage by RNase T2. Moreover, an interaction of helices αH and αI in the thumb subdomain of the p66 subunit in the shallow groove of the vRNA–$tRNA_3^{Lys}$ complex is very similar to that observed in the crystallographic structure with duplex DNA (Jacobo-Molina *et al.*, 1993; Ding *et al.*, 1997). This interaction protects nt 72–74 of $tRNA_3^{Lys}$ from cleavage by RNase V1.

The three single-stranded nucleotides upstream of the PBS and helix 2 do not bind to the cleft of RT by passing between the thumb and the finger subdomains towards the front of RT (Figures 1 and 5B). Rather, they fold backwards and are maintained simultaneously by the

fingers and thumb, explaining protection of A177 from cleavage by RNase T2 (Figure 5). Protection of nt 123–124 from cleavage by RNase V1 is due to a direct interaction of helix 1 in the vRNA–$tRNA_3^{Lys}$ complex simultaneously with the finger subdomains of the p66 and p51 subunits of HIV-1 RT. G30, which is partially protected by RT from cleavage by *N.crassa* endonuclease, is located on the floor of a pocket, the walls of which consist of helices 1, 2, 4, 3E, B and the finger subdomain of p66 (Figure 5B). C28, which is located in the same pocket, is cleaved by RNase V1 even in the presence of RT (Figure 2E). This result is explained by the low mol. wt of RNase V1 (15.9 kDa), as compared with the *N.crassa* endonuclease (55 kDa). The partial protection of nt 40–42 from nuclease attack by RT is probably indirect. It might result from a movement of stem–loop 8 towards helix 5D upon binding of HIV-1 RT (Figure 5).

Biological implications of the vRNA–$tRNA_3^{Lys}$:HIV-1 RT model—comparison with other retroviruses and retrotransposons

The specificity of initiation of reverse transcription in HIV-1 appears to rely on extended interactions between the vRNA and $tRNA_3^{Lys}$, in particular between the anticodon loop of $tRNA_3^{Lys}$ and the viral A-rich loop (helix 6C in our secondary structure model, Figure 1). Our three-dimensional model of the vRNA–$tRNA_3^{Lys}$ complex reveals that these extended primer–template interactions constitute a single interface (Figure 4). Unexpectedly, nuclease footprinting experiments, as well as docking of HIV-1 RT on the vRNA–$tRNA_3^{Lys}$ model, suggests that there is no direct contact between this interface and the HIV-1 polymerase. Similarly, there are no contacts between RT and $tRNA_3^{Lys}$ outside helix 7F. These results are in keeping with those of Morrow and co-workers showing that HIV-1 can stably use tRNAs other than $tRNA_3^{Lys}$ as primers provided both the PBS and A-rich loop are mutated to complement these tRNAs (Wakefield *et al*., 1996; Zhang *et al*., 1996; Kang *et al*., 1997; Li *et al*., 1997), and thus imply that there is no sequence-specific recognition of the extended vRNA–tRNA interactions by HIV-1 RT. Our three-dimensional model of the ternary complex suggests indirect roles for the extended interactions in the recognition of the vRNA–$tRNA_3^{Lys}$ complex by RT. They impose a highly ordered and rigid structure to nt 141–167. In their absence, these nucleotides could adopt a loose flexible structure that could interfere with RT binding. In addition, these interactions constrain the structure of the entire vRNA–$tRNA_3^{Lys}$ complex, ensuring the proper orientation of helices 1, 2, 7F and 8 that interact directly with RT.

Site-directed crosslinking experiments suggested interactions between the anticodon loop of $tRNA_3^{Lys}$ and the thumb subdomain of p66 and/or p51 (Mishima and Steitz, 1995) which are not present in our model. However, synthetic $tRNA_3^{Lys}$ lacking the natural posttranscriptional modifications were used in these experiments. This synthetic primer, unlike its natural counterpart, is a very poor primer for HIV-1 RT (Isel *et al*., 1996), and its anticodon loop does not interact with the viral A-rich loop (Isel *et al*., 1993, 1995). Thus, these crosslinking data are not directly relevant to the initiation complex. However, they might reflect a later stage of the process, when the A-rich loop–anticodon–loop interaction is disrupted due to DNA synthesis, as suggested by mutation of K249, R307 and R311 in the p66 subunit of HIV-1 RT (Arts *et al*., 1998). These mutants retained their capacity to initiate DNA synthesis but failed to elongate the nascent transcripts.

At all stages of reverse transcription, the duplex formed by the template and the newly synthesized DNA probably interacts within the cleft of HIV-1 RT as previously shown in the co-crystal of RT with a duplex DNA (Jacobo-Molina *et al*., 1993; Ding *et al*., 1997). During initiation of HIV-1 reverse transcription, helix 7F probably interacts similarly with RT (Figure 5B). However, additional interactions we propose between HIV-1 RT and the vRNA–$tRNA_3^{Lys}$ complex, namely with helices 1, 2 and 8 cannot be maintained throughout elongation of reverse transcription and thus are specific for the initiation phase. The fact that the vRNA–$tRNA_3^{Lys}$ complex is recognized efficiently only by HIV-1 RT (Arts *et al*., 1996; Isel *et al*., 1996) suggests that some of the structural features of this complex are unique to HIV-1. However, given the conservation of the general architecture of polymerase in general and RTs in particular, it is possible that homologous, although not identical, interactions exist in the initiation complex of reverse transcription of other retroelements.

An extensive probing study has been conducted on the primer–template complex formed by MoMLV RNA and $tRNA^{Pro}$, and a three-dimensional model of this complex was proposed (Fossé *et al*., 1998). The MoMLV primer–template model reveals structural equivalents to helices 2, 7F and 8, but no equivalent to helix 1. The absence of an equivalent to helix 1 in MoMLV might be related to the fact that this helix interacts with both subunits of HIV-1 RT, while the MoMLV RT is a monomer. In this retroviral RNA genome, there is no equivalent to helices 3E, 4, 5D and 6C, and helix 2 is simply closed by a tetraloop. Thus, the absence of extended MoMLV RNA–$tRNA^{Pro}$ interactions are consistent with our proposal that one of their roles in HIV-1 is to prevent interference of nt 140–167 with RT binding.

Secondary structure models of the RSV RNA–$tRNA^{Trp}$ (Aiyar *et al*., 1992, 1994) and Ty1 RNA–$tRNA_i^{Met}$ (Friant *et al*., 1996) complexes were proposed on the basis of structural probing and genetic data, respectively. In Ty1 retrotransposon, the 10-bp PBS helix is prolonged by a 7-bp helix resulting from extended primer–template interactions. With colinear stacking, these two helices might be analogous to helix 7F. In addition, viral helices that may be structural equivalents to helices 2 and 8 could be easily identified. Finally, extended Ty1 RNA–$tRNA_i^{Met}$ interactions form an additional helix that might play a role similar to helix 1. In RSV, the U5-IR and U5-leader stems are probably structurally equivalent to helices 2 and 8 in HIV-1. Furthermore, the helix resulting from interaction of the TΨC arm of $tRNA^{Trp}$ with a viral sequence upstream of the PBS might be structurally related to helix 1. Thus, structural equivalents of helices 2 and 8 can probably be found in the initiation complex of most retroviruses and retrotransposon, while more diversity is found regarding helix 1, which might play an important role in the specificity of RT binding to the tRNA–vRNA complex.

A remarkable feature of our three-dimensional model

of the HIV-1 initiation complex is the kink between helices 7F and 2, which is imposed by the short junctions between helices 1 and 2, as well as between helices 8 and 1, and by the extended vRNA–$tRNA_3^{Lys}$ interactions. A similar kink is observed in the three-dimensional model of the MoMLV RNA–$tRNA^{Pro}$ complex (Fossé *et al*., 1998) and can be predicted from the secondary structure models of the Ty1 RNA–$tRNA_i^{Met}$ (Friant *et al*., 1996) and RSV RNA–$tRNA^{Trp}$ (Aiyar *et al*., 1992, 1994) complexes. This kink fully exposes the 3′ end of the primer tRNA and the single-stranded nucleotides linking helices 7F and 2, which are the first to be reverse transcribed. This kink might be of crucial importance in preventing stacking between these helices, which would otherwise mask the 3′-terminus of the primer. The kink between helices 7F and 2 is reminiscent of the situation observed in the crystal structure of human DNA polymerase β complexed with a gapped DNA–DNA primer–template, where the 3′ end of the primer is exposed by a 90° kink in the gapped region separating the upstream and downstream helices (Sawaya *et al*., 1997).

As a consequence of the kink, the position of the template relative to RT in our initiation complex model differs from that previously proposed in elongation complex models (Kohlstaedt *et al*., 1992; Wöhrl *et al*., 1995). In the elongation complex, the single-stranded template region upstream of the active site was modeled in continuity with the double-stranded primer–template region (equivalent to helix 7F) bound in the RT cleft. Thus, in the elongation complex models, the single-stranded template nucleotides pass through the cleft formed by the thumb and finger subdomains of the p66 subunit of HIV-1 RT (Kohlstaedt *et al*., 1992; Wöhrl *et al*., 1995). In our initiation complex model, helix 2 folds backward with respect to helix 7F, making an angle of ~110°, and thus the template nucleotides upstream of the active site do not pass through the cleft formed by the thumb and fingers. Rather, they interact with the top and rear regions of these subdomains (Figure 5B). Whether the position of the template nucleotides upstream of the PBS differs during the initiation and elongation stages or whether the template was not adequately modeled in the elongation complex due to the lack of constraints is unclear. We showed previously that RT dissociates from the primer–template complex during the transition from the initiation to elongation (Isel *et al*., 1996). Dissociation and rebinding of RT could allow flipping of the template after disruption of the extended vRNA–$tRNA_3^{Lys}$ interactions. On the other hand, recent X-ray structures of T7 phage (Doublié *et al*., 1998) and *Bacillus* (Kiefer *et al*., 1998) DNA polymerases complexed with DNA primer–DNA template hybrids revealed that the template-strand overhang makes a sharp turn (~90°) relative to the axis of the DNA duplex bound in the RT cleft. Thus, it is possible that even during the elongation phase, the template nucleotides upstream of the active site lie on the surface of the fingers and/or thumb rather than passing through the cleft formed by these subdomains. This hypothesis is supported by the structure of a covalently trapped catalytic elongation complex with a 3 nt overhang in the template that was published after submission of this manuscript (Huang *et al*., 1998). This crystal structure does not follow the original suggestion that the template strand might pass through the cleft between the fingers and the thumb, since contacts between the first template base upstream of the catalytic site and the fingers prevent the polynucleotide chain from continuing in a helical path (Huang *et al*., 1998).

In summary, our probing and modeling data provide a structural explanation for the specificity of HIV-1 reverse transcription, and an initiation complex which constitutes an attractive target for selective antiviral compounds.

Materials and methods

Templates, primer and RT

Wild-type RNA templates corresponding to nt 1–311 or 123–217 of HIV-1 (Mal isolate) were synthesized by *in vitro* transcription with T7 RNA polymerase as previously described (Paillart *et al*., 1994; Skripkin *et al*., 1996). Purification and labeling of natural $tRNA_3^{Lys}$ was as previously described (Isel *et al*., 1993). The HIV-1 RT used in this study was an RNase H(–) enzyme bearing the E478Q point mutation purified as described by Le Grice and Grueninger-Leitch (1990). This mutant polymerase was used in order to prevent cleavage of the RNA template by the dsRNase activity associated with the HIV-1 RNase H (Götte *et al*., 1995).

Chemical and enzymatic probing of vRNA

For G-N7 modification of the vRNA, the binary primer–template complex was formed as described previously (Isel *et al*., 1993, 1995) and cooled at room temperature for 15 min. One microliter of DMS diluted 20-fold in ethanol was added to 20 µl samples and the reaction was allowed to proceed for 6 and 12 min at 20°C. Chain scission at methylated G-N7 positions was performed by aniline treatment (Peattie and Gilbert, 1980). Probing of the A-N7 positions of HIV-1 RNA involved in the primer–template with DEPC was essentially as described for $tRNA_3^{Lys}$ (Isel *et al*., 1995), except that aniline treatment was omitted.

For hydroxyl radical probing, wild-type 1–311 or 123–217 RNAs (8 pmol) were incubated alone or with a 3-fold molar excess of $tRNA_3^{Lys}$ in water for 2 min at 90°C, cooled on ice and incubated at 70°C for 20 min in 50 mM sodium cacodylate (pH 7.5) and 300 mM KCl. After hybridization, samples were incubated in the same buffer supplemented with 5 mM $MgCl_2$ at 20°C for 15 min. After addition of 2 µg of total yeast tRNA, Fe-EDTA probing experiments were performed on 21 µl samples by adding separately on the wall of the Eppendorf tubes 1 µl of 100 mM EDTA, 1 µl of 250 mM DTE, 1 µl of diluted H_2O_2 (H_2O_2 final concentration: 0.025, 0.05 or 0.1%) and 1 µl of a freshly prepared $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ solution at 20 mg/ml (Murphy and Cech, 1994). Reactions were started by centrifuging, incubated for 5 min at 37°C, and were stopped by addition of 25 µl of 10 mM thiourea and precipitated with 0.3 M sodium acetate and 3 vol ethanol for at least 20 min at 0°C.

Modified bases and riboses and positions of cleavage were detected by primer extension as previously described (Baudin *et al*., 1993).

Chemical and enzymatic probing of $tRNA_3^{Lys}$

For chemical modifications of $tRNA_3^{Lys}$, viral RNA (12 pmol) and 50 000 c.p.m. of 3′ or 5′ end-labeled $tRNA_3^{Lys}$ (~0.4 pmol) were hybridized as described above and the reaction mixture was cooled at 37°C for 10 min. Chemical modification at G-N7 was with 2 µl of DMS diluted 20-fold in ethanol, for 5 and 10 min, at 37°C. Chain scission at methylated G-N7 was induced by aniline treatment. Hydroxyl radical footprinting of $tRNA_3^{Lys}$ was essentially as described above for vRNA. Probing with *N.crassa* endonuclease and RNase V1 of $tRNA_3^{Lys}$ involved in the binary complex with either 1–311 or 123–217 RNAs was as described elsewhere (Isel *et al*., 1993, 1995).

Footprinting of RT on the primer–template complex

For RNase T2 probing experiments on vRNA, ^{32}P end-labeled 123–217 RNA (~50 000 c.p.m.) and $tRNA_3^{Lys}$ (24 pmol), were hybridized as described above in 50 mM Tris–HCl pH 7.5, 40 mM KCl. After hybridization, samples were incubated in the same buffer supplemented with 5 mM $MgCl_2$ and incubated at room temperature for 5 min. The vRNA–$tRNA_3^{Lys}$:HIV-1 RT complex was obtained by further incubation with 50 pmol (2.3 µM) of RT for 10 min at 37°C. RNase T2 cleavage was then performed with 0.1 U of enzyme for 5 and 10 min at 37°C.

Endonuclease V1 probing was performed on 1–311 vRNA (4 pmol) hybridized to $tRNA_3^{Lys}$ (12 pmol) in 50 mM Tris–HCl pH 8.0 (at 37°C), 40 mM KCl. Subsequent formation of the vRNA–$tRNA_3^{Lys}$:RT complexes (10- to 30-fold excess) in the same buffer supplemented with

5 mM $MgCl_2$ was for 10 min on ice, followed by a 10 min incubation at room temperature. Cleavage was at 20°C for 8 min with 0.15 and 0.22 U of RNase V1. After phenol/chloroform extraction, samples were precipitated with ethanol in presence of 0.3 M sodium acetate and total tRNA. Cleavage positions were detected by extension of primers complementary to nt 297–311 and 178–196, as previously described (Isel *et al*., 1995).

For enzymatic modifications of $tRNA_3^{Lys}$, vRNA (12 pmol) and 50 000 c.p.m. of 3′ end-labeled $tRNA_3^{Lys}$ (~0.4 pmol) were hybridized as described above. The ternary complex was formed by addition of 120 pmol of HIV-1 RT, incubation on ice for 10 min, then at 37°C for 10 min, in the presence of 5 mM $MgCl_2$. Probing was for 5 and 10 min at 37°C with 0.3 U of *N.crassa* endonuclease or 0.05 U of RNase V1, followed by a 20 min incubation with 0.5 mg/ml of proteinase K in 1% SDS, phenol/chloroform extraction and ethanol precipitation.

Molecular modeling and docking

Programs used for RNA modeling have been described previously (Westhof, 1993). Assembly and connection of the secondary elements into a three-dimensional fold, as well as the docking of the vRNA–$tRNA_3^{Lys}$ complex to RT, was accomplished with a graphic system PS300 from Evans and Sutherland using FRODO (Jones, 1978), adapted for the PS300 (Pflugrath *et al*., 1983). The D stem of $tRNA_3^{Lys}$, which is the only part of the primer that does not undergo a conformational change during complex formation, was modeled according to the crystal structure of yeast $tRNA^{Asp}$ (Westhof *et al*., 1985). Finally, models of the binary complex, but not those of the ternary complex were subjected to restrained least-squares refinement (Konnert and Hendrickson, 1980) with the programs NUCLIN and NUCLSQ (Westhof *et al*., 1985) in order to ensure geometry and stereochemistry with allowed distances between interacting atoms and to avoid steric conflicts. The color views were generated with DRAWNA (Massire *et al*., 1994).

Acknowledgements

We thank G.Keith and M.Yusupov for help with the tRNA purification, and J.-M.Lanchy for fruitful discussions. This work was supported by the ‘Agence Nationale de Recherches sur le SIDA’ (ANRS). S.F.J.LeG. was supported by PHS grant A131147.

References

Aiyar,A., Cobrinik,D., Ge,Z., Kung,H.J. and Leis,J. (1992) Interaction between retroviral U5 RNA and the TΨC loop of the $tRNA^{Trp}$ primer is required for efficient initiation of reverse transcription. *J. Virol*., **66**, 2464–2472.

Aiyar,A., Ge,Z. and Leis,J. (1994) A specific orientation of RNA secondary structure is required for initiation of reverse transcription. *J. Virol*., **68**, 611–618.

Arts,E.J. and Le Grice,S.F.J. (1998) Interaction of retroviral reverse transcriptase with template–primer duplexes during replication. *Prog. Nucleic Acids Res. Mol. Biol.*, **58**, 339–393.

Arts,E.J. *et al.* (1996) Initiation of (–) strand DNA synthesis from $tRNA_3^{Lys}$ on lentiviral RNAs: Implications of specific HIV-1 RNA–$tRNA_3^{Lys}$ interactions inhibiting primer utilization by retroviral reverse transcriptases. *Proc. Natl Acad. Sci. USA*, **93**, 10063–10068.

Arts,E.J., Miller,J.T., Ehresmann,B. and Le Grice,S.F.J. (1998) Mutating a region of HIV-1 reverse transcriptase implicated in $tRNA_3^{Lys}$ binding and the consequence for (–)-strand DNA synthesis. *J. Biol. Chem*., **273**, 14523–14532.

Baltimore,D. (1970) Viral RNA-dependent DNA polymerase. *Nature*, **226**, 1209–1211.

Baudin,F., Marquet,R., Isel,C., Darlix,J.L., Ehresmann,B. and Ehresmann,C. (1993) Functional sites in the 5′ region of human immunodeficiency virus type-1 RNA form defined structural domains. *J. Mol. Biol*., **229**, 382–397.

Cate,J.H., Gooding,A.R., Podell,E., Zhou,K.H., Golden,B.L., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science*, **273**, 1678–1685.

Celander,D.W. and Cech,T.R. (1991) Visualizing the higher order folding of a catalytic RNA molecule. *Science*, **251**, 401–407.

Cobrinik,D., Soskey,L. and Leis,J. (1988) A retroviral RNA secondary structure required for efficient initiation of reverse transcription. *J. Virol*., **62**, 3622–3630.

Colicelli,J. and Goff,S.P. (1987) Isolation of a recombinant murine leukemia virus utilizing a new primer tRNA. *J. Virol*., **57**, 37–45.

Cordell,B.R., Swanstrom,R., Goodman,H. and Bishop,J.M. (1979) $tRNA^{Trp}$ as primer for RNA-directed DNA polymerase: structural determinants of function. *J. Biol. Chem*., **254**, 1866–1874.

Das,A.T., Klaver,B. and Berkhout,B. (1995) Reduced replication of human immunodeficiency virus type 1 mutants that use reverse transcription primers other than the natural $tRNA_3^{Lys}$. *J. Virol*., **69**, 3090–3097.

Ding,J., Hughes,S.H. and Arnold,E. (1997) Protein–nucleic acid interactions and DNA conformation in a complex of human immunodeficiency virus type 1 reverse transcriptase with a double-stranded DNA template–primer. *Biopolymers*, **44**, 125–138.

Doublié,S., Tabor,S., Long,A.M., Richardson,C.C. and Ellenberger,T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature*, **391**, 251–258.

Esnouf,R., Ren,J.S., Ross,C., Jones,Y., Stammers,D. and Stuart,D. (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. *Nature Struct. Biol*., **2**, 303–308.

Fossé,P., Mougel,M., Keith,G., Westhof,E., Ehresmann,B. and Ehresmann,C. (1998) Modified nucleotides of $tRNA^{Pro}$ restrict interactions in the binary primer/template complex of M-MLV. *J. Mol. Biol*., **275**, 731–746.

Friant,S., Heyman,T., Wilhelm,M.L. and Wilhelm,F.X. (1996) Extended interactions between the primer $tRNA_i^{Met}$ and genomic RNA of the yeast Ty1 retrotransposon. *Nucleic Acids Res*., **24**, 441–449.

Gilboa,E., Mitra,S.W., Goff,S. and Baltimore,D. (1979) A detailed model of reverse transcription and tests of crucial aspects. *Cell*, **18**, 93–100.

Götte,M., Fackler,S., Hermann,T., Perola,E., Cellai,L., Gross,H.J., Le Grice,S.F.J. and Heumann,H. (1995) HIV-1 reverse transcriptase-associated RNase H cleaves RNA/RNA in arrested complexes: Implications for the mechanism by which RNase H discriminates beween RNA/RNA and RNA/DNA. *EMBO J*., **14**, 833–841.

Huang,H., Chopra,R., Verdine,G.L. and Harrison,S.C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science*, **282**, 1669–1675.

Huang,Y., Shalom,A., Li,Z., Wang,J., Mak,J., Wainberg,M.A. and Kleiman,L. (1996) Effects of modifying the $tRNA_3^{Lys}$ anticodon on the initiation of human immunodeficiency virus type 1 reverse transcription. *J. Virol*., **70**, 4700–4706.

Isel,C., Marquet,R., Keith,G., Ehresmann,C. and Ehresmann,B. (1993) Modified nucleotides of transfer-RNA_3^{Lys} modulate primer/template loop–loop interaction in the initiation complex of HIV-1 reverse transcription. *J. Biol. Chem*., **268**, 25269–25272.

Isel,C., Ehresmann,C., Keith,G., Ehresmann,B. and Marquet,R. (1995) Initiation of reverse transcription of HIV-1: secondary structure of the HIV-1 RNA/$tRNA_3^{Lys}$ (template/primer) complex. *J. Mol. Biol*., **247**, 236–250.

Isel,C., Lanchy,J.M., Le Grice,S.F.J., Ehresmann,C., Ehresmann,B. and Marquet,R. (1996) Specific initiation and switch to elongation of human immunodeficiency virus type 1 reverse transcription require the post-transcriptional modifications of primer $tRNA_3^{Lys}$. *EMBO J*., **15**, 917–924.

Isel,C., Keith,G., Ehresmann,B., Ehresmann,C. and Marquet,R. (1998) Mutational analysis of the $tRNA_3^{Lys}$/HIV-1 RNA (primer/template) complex. *Nucleic Acids Res*., **26**, 1198–1204.

Jacobo-Molina,A. *et al.* (1993) Crystal structure of human immunodeficiency virus type-1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl Acad. Sci. USA*, **90**, 6320–6324.

Jager,J., Smerdon,S.J., Wang,J.M., Boisvert,D.C. and Steitz,T.A. (1994) Comparison of three different crystal forms shows HIV-1 reverse transcriptase displays an internal swivel motion. *Structure*, **2**, 869–876.

Jones,T.A. (1978) A graphic model building and refinement system for macromolecules. *J. Appl. Crystallogr.*, **11**, 268–272.

Kang,S.M., Zhang,Z.J. and Morrow,C.D. (1997) Identification of a sequence within U5 required for human immunodeficiency virus type 1 to stably maintain a primer binding site complementary to $tRNA^{Met}$. *J. Virol*., **71**, 207–217.

Kiefer,J.R., Mao,C., Braman,J.C. and Beese,L.S. (1998) Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature*, **391**, 304–307.

Kohlstaedt,L.A., Wang,J., Friedman,J.M., Rice,P.A. and Steitz,T.A. (1992) Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science*, **256**, 1783–1790.

Konnert,J.H. and Hendrickson,W.A. (1980) A restrained-parameter thermal-factor refinement procedure. *Acta Crystallogr*., **36**, 344–350.

Lanchy,J.M., Ehresmann,C., Le Grice,S.F.J., Ehresmann,B. and Marquet,R. (1996) Binding and kinetic properties of HIV-1 reverse transcriptase markedly differ during initiation and elongation of reverse transcription. *EMBO J.*, **15**, 7178–7187.

Latham,J.A. and Cech,T.R. (1989) Defining the inside and outside of a catalytic RNA molecule. *Science*, **245**, 276–282.

Le Grice,S.F.J. and Grueninger-Leitch,F. (1990) Rapid purification of homodimer and heterodimer HIV-1 reverse transcriptase by metal chelate affinity chromatography. *Eur. J. Biochem.*, **187**, 307–314.

Li,X.G., Mak,J., Arts,E.J., Gu,Z.X., Kleiman,L., Wainberg,M.A. and Parniak,M.A. (1994) Effects of alterations of primer-binding site sequences on human immunodeficiency virus type 1 replication. *J. Virol.*, **68**, 6198–6206.

Li,Y., Zhang,Z., Wakefield,J.K., Kang,S.M. and Morrow,C.D. (1997) Nucleotide substitutions within U5 are critical for efficient reverse transcription of human immunodeficiency virus type 1 with a primer binding site complementary to $tRNA^{His}$. *J. Virol.*, **71**, 6315–6322.

Liang,C., Li,X., Rong,L., Inouye,P., Quan,Y., Kleiman,L. and Wainberg,M.A. (1997) The importance of the A-rich loop in human immunodeficiency virus type 1 reverse transcription and infectivity. *J. Virol.*, **71**, 5750–5757.

Lund,A.H., Duch,M., Lovmand,J., Jørgensen,P. and Pedersen,F.S. (1993) Mutated primer binding sites interacting with different tRNAs allow efficient murine leukemia virus replication. *J. Virol.*, **67**, 7125–7130.

Lund,A.H., Duch,M., Lovmand,J., Jørgensen,P. and Pedersen,F.S. (1997) Complementation of a primer binding site-impaired murine leukemia virus-derived retroviral vector by a genetically engineered tRNA-like primer. *J. Virol.*, **71**, 1191–1195.

Mak,J. and Kleiman,L. (1997) Primer tRNAs for reverse transcription. *J. Virol.*, **71**, 8087–8095.

Marquet,R., Isel,C., Ehresmann,C. and Ehresmann,B. (1995) tRNAs as primer of reverse transcriptase. *Biochimie*, **77**, 113–124.

Massire,C., Gaspin,C. and Westhof,E. (1994) DRAWNA: a program for drawing schematic views of nucleic acids. *J. Mol. Graph.*, **12**, 201–207.

Metzger,W., Hermann,T., Schatz,O., Le Grice,S.F.J. and Heumann,H. (1993) Hydroxyl radical footprint analysis of human immunodeficiency virus reverse transcriptase–template·primer complexes. *Proc. Natl Acad. Sci. USA*, **90**, 5909–5913.

Mishima,Y. and Steitz,J.A. (1995) Site-specific crosslinking of 4-thiouridine-modified human $tRNA_3^{Lys}$ to reverse transcriptase from human immunodeficiency virus type 1. *EMBO J.*, **14**, 2679–2687.

Moine,H., Ehresmann,B., Ehresmann,C. and Romby,P. (1997) Probing RNA structure and function in solution. In Simons,R.W. and Grunberg-Manago,M. (eds), *RNA Structure and Function*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 77–115.

Murphy,F.L. and Cech,T.R. (1994) GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.*, **236**, 49–63.

Oude Essink,B.B., Das,A.T. and Berkhout,B. (1996) HIV-1 reverse transcriptase discriminates against non-self tRNA primers. *J. Mol. Biol.*, **264**, 243–254.

Paillart,J.C., Marquet,R., Skripkin,E., Ehresmann,B. and Ehresmann,C. (1994) Mutational analysis of the bipartite dimer linkage structure of HIV-1 genomic RNA. *J. Biol. Chem.*, **269**, 27486–27493.

Peattie,D.A. and Gilbert,W. (1980) Chemical probes for higher-order structure in RNA. *Proc. Natl Acad. Sci. USA*, **77**, 4679–4682.

Pflugrath,J.W., Saper,M.A. and Quiocho,F.A. (1983) Molecular modeling with the PS300: a new generation graphics display system. *J. Mol. Graph.*, **1**, 53–54.

Richmond,T.J. (1984) Solvent accessible surface area and excluded volume in proteins. *J. Mol. Biol.*, **178**, 63–89.

Sawaya,M.R., Prasad,R., Wilson,S.H., Kraut,J. and Pelletier,H. (1997) Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: Evidence for an induced fit mechanism. *Biochemistry*, **36**, 11205–11215.

Serganov,A.A., Masquida,B., Westhof,E., Cachia,C., Portier,C., Garber,M., Ehresmann,B. and Ehresmann,C. (1996) The 16S rRNA binding site of *Thermus thermophilus* ribosomal protein S15: Comparison with *Escherichia coli* S15, minimum site and structure. *RNA*, **2**, 1124–1138.

Skripkin,E., Isel,C., Marquet,R., Ehresmann,B. and Ehresmann,C. (1996) Psoralen crosslinking between human immunodeficiency virys type 1 RNA and primer $tRNA_3^{Lys}$. *Nucleic Acids Res.*, **24**, 509–514.

Temin,H.M. and Mizutani,S. (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature*, **226**, 1211–1213.

Wakefield,J.K., Wolf,A.G. and Morrow,C.D. (1995) Human immunodeficiency virus type 1 can use different tRNAs as primers for reverse transcription but selectively maintains a primer binding site complementary to $tRNA_3^{Lys}$. *J. Virol.*, **69**, 6021–6029.

Wakefield,J.K., Kang,S.-M. and Morrow,C.D. (1996) Construction of a type 1 human immunodeficiency virus that maintains a primer binding site complementary to $tRNA^{His}$. *J. Virol.*, **70**, 966–975.

Westhof,E. (1993) Modelling the three-dimensional structure of ribonucleic acids. *J. Mol. Struct.*, **286**, 203–210.

Westhof,E., Dumas,P. and Moras,D. (1985) Crystallographic refinement of yeast aspartic acid transfer RNA. *J. Mol. Biol.*, **184**, 119–145.

Whitcomb,J.M., Ortiz-Conde,B.A. and Hughes,S.H. (1995) Replication of avian leukosis viruses with mutations at the primer binding site: use of alternative tRNAs as primers. *J. Virol.*, **69**, 6228–6238.

Wöhrl,B.M., Tantillo,C., Arnold,E. and Le Grice,S.F.J. (1995) An expanded model of replicating human immunodeficiency virus reverse transcriptase. *Biochemistry*, **34**, 5343–5350.

Zhang,Z., Kang,S.M., LeBlanc,A., Hajduk,S.L. and Morrow,C.D. (1996) Nucleotide sequences within the U5 region of the viral RNA genome are the major determinants for an human immunodeficiency virus type 1 to maintain a primer binding site complementary to tRNA (His). *Virology*, **226**, 306–317.

Zhang,Z.J., Kang,S.M., Li,Y. and Morrow,C.D. (1998) Genetic analysis of the U5-PBS of a novel HIV-1 reveals multiple interactions between the tRNA and RNA genome required for initiation of reverse transcription. *RNA*, **4**, 394–406.

Received November 12, 1998; revised December 28, 1998; accepted December 30, 1998

Note added in proof

Elgavish *et al*. (*J. Mol. Biol.* 1999, **285**, 449–453) recently suggested that it is not possible to model the three-dimensional structure of the HIV-1RNA–$tRNA_3^{Lys}$ complex according to the previously proposed secondary structure (Isel *et al*., 1995). The present study clearly shows that this model is feasible and accounts for the experimental data.