

Strained template under the thumbs

How reverse transcriptase of human immunodeficiency virus type 1 moves along its template

Thomas HERMANN and Hermann HEUMANN

Max-Planck-Institut für Biochemie, Martinsried, Germany

(Received 17 June 1996) – EJB 96 0888/2

In retroviruses, such as human immunodeficiency virus type 1 (HIV-1), the reverse transcriptase (RT) copies single-stranded viral RNA into complementary DNA, which is then used as a template for synthesis of the second DNA strand. The resulting double-stranded DNA is integrated into the host genome. How RT translocates on the different templates is the subject of this study. We have developed a theoretical model for RT translocation during processive DNA synthesis. The model is based on the assumption that there are two template-binding sites, namely the helix clamps, located in the thumb subdomains of RT subunits p66 and p51. Flexibility of the p66 thumb provides undisrupted template-binding during polymerase translocation. Coordinated association and dissociation of the template at the thumbs, triggered by nucleotide incorporation, is assumed, which ensures template contact with at least one subdomain throughout translocation. We suggest that coordination between the sites is effected by stress in the template region located between the thumbs. Translocation of HIV-1 RT proceeds continuously but with different processivities on RNA and DNA templates. These findings are explained in detail by the proposed model.

Keywords: human immunodeficiency virus type 1 (HIV-1) reverse transcriptase; helix-clamp motif; non-nucleosidic inhibitor; polymerase processivity; thumb subdomain.

Whether there is a unifying concept that underlies all template-dependent nucleic acid polymerases is a question that is interesting from the evolutionary as well as biotechnological, point of view. Development of polymerase inhibitors, e.g. against human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT), requires information not only about the function of the HIV-1 enzyme but also about other cellular polymerases, to avoid undesirable side effects associated with polymerase inhibition.

Have all template-dependent nucleic acid polymerases descended from a common ancestor? Studies aimed at answering this question have included comparisons of polymerase sequences and of polymerase structures. With one exception, sequence comparisons have revealed little overall similarity in polymerases of different origin [1, 2]. There is a motif consisting of two or three aspartate residues, which all polymerases have in common [3, 4]. Mutational analysis [5] and crystallographic studies [6] have shown that this motif is part of the polymerase active site. Structural comparison of polymerases is limited, since only a few three-dimensional structures with sufficient resolution are available, namely those of the Klenow fragment of *Escherichia coli* polymerase I [6, 7] and the similar DNA polymerase from *Thermus aquaticus* [8, 9], HIV-1 RT [10–12], T7 phage RNA polymerase [13], and rat DNA polymerase β [14], all of which were determined by X-ray crystallography. Furthermore, 3-D models exist for the RNA polymerases of *E. coli* [15] and of *Saccharomyces cerevisiae* [16], obtained from two-

dimensional crystal electron diffraction. The structures of all these polymerases have three discernible subdomains, termed 'fingers', 'palm' and 'thumb' [11], which can accommodate the nucleic acid template [17]. By sequence comparison and secondary-structure analyses, an amino acid motif was found, which is shared by many polymerases. This motif has been termed the 'helix clamp' and is probably involved in template-binding [18].

It is still unclear whether the common features of polymerases mentioned above can be explained by evolutionary descent, assuming a common ancestor, or by convergence due to a common function. In addition to sequence and structure comparisons, analysis of polymerase function is another approach to identify a potential unifying concept in template-dependent nucleic acid polymerases. Here we concentrate on an important feature of polymerase function, namely translocation. The polymerase moves along the template in concert with the growing nucleic acid, or, expressed differently, the nucleic acid template is transported through the enzyme. Elongation of the nascent nucleic acid can take place in a distributive or a processive manner. The synthesis is termed distributive or processive, depending on whether the polymerase dissociates from the template after each step of synthesis or remains bound. The degree of processivity is measured by the number of nucleotides incorporated/polymerase-binding event. The translocation step is critical for processivity, since, during polymerase movement, contacts between the protein and template are transiently disrupted. Processivity varies widely among the polymerases, which might reflect different functional needs for processivity, depending on the polymerase's biological function (Table 1). For example, DNA polymerases involved in chromosomal replication must

Correspondence to H. Heumann, Max-Planck-Institut für Biochemie, D-82152 Martinsried, Germany

Abbreviations. HIV-1, human immunodeficiency virus type 1; PCNA, proliferating-cell nuclear antigen; RT, reverse transcriptase.

Table 1. Processivity of DNA polymerases. Data from Kornberg and Baker [46] and as indicated. SV 40, simian virus 40; EBV, Epstein Barr virus; AMV, avian myeloblastosis virus; gp, gene product; pol, polymerase; n.d., not determined; –, not applicable.

Polymerase	Subunits		Processivity		Processivity device
	core	holoenzyme	core	holoenzyme	
<i>E. coli</i> pol III	3 (α , ϵ , θ)	≥ 10	10–15	>5000	β subunit sliding clamp
Phage T4 pol	1 (gp43)	4	11–13	>20000	gp45 sliding clamp
Yeast/human pol δ	1 (δ)	≥ 7	2	high	PCNA sliding clamp
SV 40 pol α	1 (α)	4	n.d.	high	PCNA sliding clamp
Phage T7 pol	1 (gp5)	2	50 [22]	>10000 [22]	<i>E. coli</i> thioredoxin host factor
<i>E. coli</i> pol II	–	1	–	5 [47]	n.d.
<i>E. coli</i> pol I	–	1	–	11–200 (DNA template), 3–4 (RNA template) [48]	n.d.
Human pol α	–	1	–	10 [49]	n.d.
Vaccinia virus pol	–	1	–	≥ 2000 [50]	n.d.
EBV pol	–	1	–	≥ 7200 [51]	n.d.
Phage T5 pol	–	1	–	155–170	n.d.
AMV pol	–	1	–	22–30	n.d.
HIV-1 RT	2 (p66/p51)	2 (p66/51)	5 (DNA template), 340 (RNA template) [32]	5, 340	n.d.

replicate long DNA templates, and these enzymes thus display a high degree of processivity. Polymerases needed for DNA repair show low processivity, since these enzymes need to synthesize only short stretches of nucleic acid. Processivity is realized in different systems quite differently. Some polymerases with very high processivity bind to accessory proteins that form closed circular clamps around the template DNA, thus preventing polymerase dissociation. The three-dimensional structures of two such sliding processivity clamps have been determined, namely the β subunit of *E. coli* polymerase III [19] and the proliferating-cell nuclear antigen (PCNA) that associates to polymerase δ in the human and yeast systems [20]. Although the structures and functions of these two clamp proteins are similar, they share no significant sequence identity [21]. The sliding clamps are found in eucaryotes, bacteria, and viruses (Table 1). The polymerase of bacteriophage T7 gains high processivity in DNA synthesis by recruiting the host protein thioredoxin, the processivity-promoting mechanism of which is unknown [22]. The variety of different mechanisms that enhance processivity suggests that these mechanisms have developed largely independently in the different polymerase systems, and emerged some time after development of the polymerase activity itself.

Processivity is closely linked to translocation when protein/nucleic-acid contacts are disrupted and re-established during polymerase movement. Therefore, any proposed mechanism for polymerase translocation needs to explain the observed processivity of the respective enzyme. There is a debate on how template-dependent nucleic acid polymerases move along their templates. It was suggested that contraction and extension of the substrate nucleic acid by a transition of the substrate nucleic acid from the A to B form contributes to polymerase translocation [12]. This model is plausible for polymerases with DNA substrates, because DNA can convert from the B form to the A form upon binding to proteins. We exclude the possibility of this model for HIV-1 RT, since this polymerase must accept substrate nucleic acids which are already in the A form such as RNA/DNA and RNA/RNA duplexes. For *E. coli* RNA polymerase, an inchworm-like movement of the protein was suggested [23, 24], assuming contraction and stretching of the polymerase. A ratchet-type mechanism was proposed for translocation of HIV-1 RT [25]. Inchworm and ratchet-type models are conceptually similar, as they require at least two binding sites for the template that act separately but in concert, indicating some kind

of communication between the two sites. One of the binding sites must be flexible to follow the movement of the template relative to the polymerase active site. Both models apply to continuous or to discontinuous translocation. The translocation process is termed continuous if movement of the polymerase occurs after each step of synthesis, and is termed discontinuous if the movement occurs only after incorporation of several nucleotides. While recent studies on stalled transcription complexes of *E. coli* RNA polymerase indicate that translocation encompasses continuous and discontinuous steps, such as jumping of polymerase [26], a simple continuous translocation model is assumed to apply for HIV-1 RT. This model is suggested from footprinting data of HIV-1 RT arrested in specific registers of DNA synthesis, which show no changes in the protection pattern in the different registers [27].

In this study, we investigated HIV-1 RT. This enzyme has a prominent position among template-dependent nucleic acid polymerases, since its three-dimensional structure is known in detail from crystallographic analyses [10–12, 28]. This facilitates testing of models that explain RT translocation on a structural level. HIV-1 RT is a template-dependent DNA polymerase that catalyzes a key step during retroviral infection, namely the synthesis of a DNA copy of the viral RNA genome [29, 30]. During reverse transcription, RT has to accept RNA and DNA as a template [31]. The viral RNA is the template during minus (first)-strand DNA synthesis. This DNA serves as template during plus (second)-strand DNA synthesis. Processivity of RT is low and varies depending on whether RNA or DNA is used as a template. Reardon has shown [32] that around 340 nucleotides are incorporated/binding event in the RNA-dependent synthesis mode and only 5 nucleotides/binding event in the DNA-dependent mode.

Here we suggest a model for RT translocation which is based on two assumptions. (a) There are two template-binding sites, which are located in the two subunits of RT, in the thumb subdomains. We suggest that the helix-clamp motif, recently described as being conserved in different polymerases [18], is essentially responsible for interaction with the nucleic acid template. (b) The thumb subdomain in the p66 thumb is flexible, which provides a movable template-binding domain required for translocation. Flexibility of the thumb was suggested previously by inspection of HIV-1 RT models obtained by crystallographic studies [12, 28].

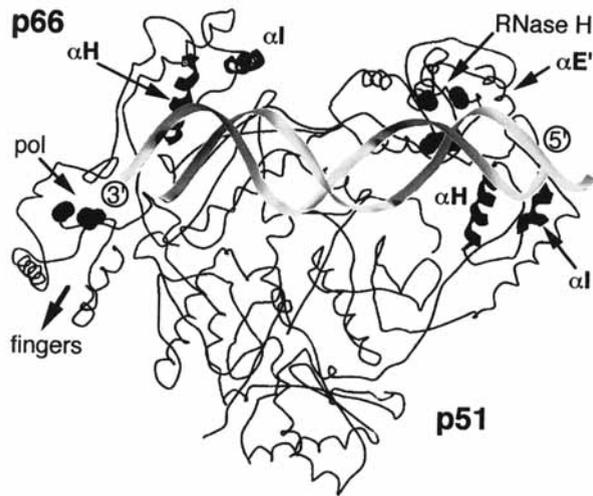


Fig. 1. Overview of a complex of HIV-1 RT and a 27-bp dsRNA substrate modeled by Hermann et al. [18]. For the sake of clarity, parts of the p66 fingers subdomain are omitted. The polymerization and RNase-H active sites are marked by spheres. Helices α H and α I within the helix-clamp motif of both thumb subdomains, which interact with the RNA substrate, are marked by black ribbons. The backbone of the 27-bp template/primer RNA is shown as grey ribbons. The ends of the primer strand are marked 3' and 5'.

RESULTS AND DISCUSSION

In HIV-1 RT, the template is fixed between two sites, namely the thumbs of the p66 and p51 subunits. HIV-1 RT is composed of two subunits, named p66 and p51. The amino acid sequence of p51 is identical to the N-terminal portion of p66. Both subunits contain four subdomains, termed 'fingers', 'palm', 'thumb' and 'connection' [11]. The p66 subunit contains an extra C-terminal domain of 15 kDa, which carries the RNase-H activity. Although the N-terminal part of the two subunits is identical, the structures and functions of the two subunits are different. While the local folding within the subdomains is similar in both subunits, the arrangement of the subdomains differs, as X-ray crystallographic analyses revealed [10–12]. As a consequence of the different folding, only the p66 subunit carries the polymerase active site [11, 12, 33].

X-ray crystallographic analysis of an RT · DNA complex and mutagenesis studies showed that the nucleic acid template binds in the cleft formed by the finger and thumb subdomains of the p66 subunit [12, 34]. The three-dimensional structure derived from analysis of an RT · DNA crystal revealed protein contacts to nucleic acid in proximity to the polymerase active site, namely in the p66 fingers, palm and thumb subdomains [12, 35]. In accordance with their function, these contacts were termed 'primer grip', 'template grip' [12] and 'helix clamp' [18].

Whether there is a second site further upstream in the p51 subunit, which contributes to template-binding has not been proven experimentally. In the three-dimensional structure of an RT · DNA complex, the 18-bp dsDNA fragment used was too short to contact the p51 subunit [12]. However, there are indications that interaction between RT and template occur upstream of the polymerase active site. Molecular-modeling studies, which use a dsRNA fragment of 27 bp as a model substrate, suggest that template contacts are possible with residues of helix α E' in the p66 RNase-H domain and with amino acids in the p51 thumb subdomain [18] (Fig. 1). DNase-I-footprinting experiments that revealed protection of DNA up to base position

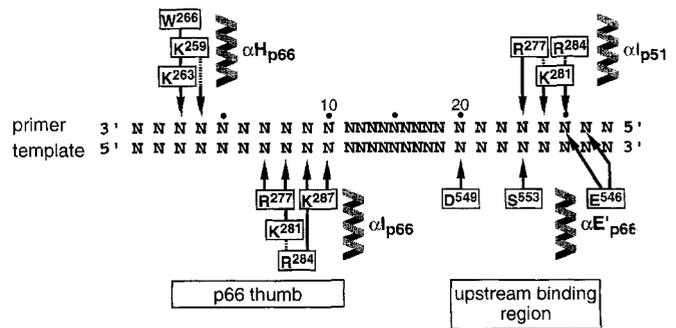


Fig. 2. Possible contacts between the nucleotides of the modeled 27-bp RNA and amino acids in the p66 thumb and the upstream binding region. Residues within helices α H and α I of the thumb and within α E' of the RNase H domain are included [18].

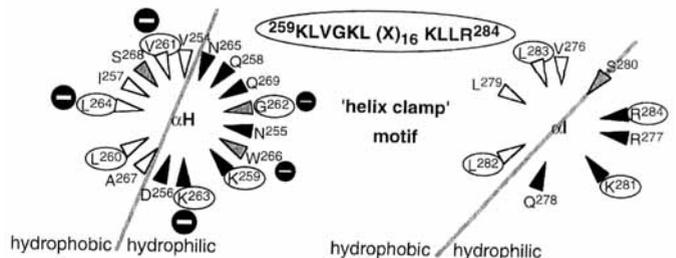


Fig. 3. Helical-wheel representation of helices α H and α I in the RT thumbs. Both helices are amphiphilic. Δ , hydrophobic residues; \blacktriangle , hydrophilic residues. Residues are encircled within the consensus sequence that is conserved in many polymerases [18]. Positions in α H where exchange to alanine lead to significant loss of RT activity [37] are indicated by minus signs.

–23 in the template and up to position –25 in the primer strand [36] are in line with the above hypothesis.

We suggest that two template-binding sites within the p66 and p51 thumb, namely the helix clamps, have important roles in the translocation process. The helix clamp is a bipartite motif located in helices α H and α I (Fig. 3). This motif is conserved in many polymerases, as recently reported [18], and probably participates in template-binding of these polymerases. Evidence for the template-binding capacity of the helix clamps in HIV-1 RT is provided by molecular-modeling [18] and mutational studies [37].

HIV-1 RT interacts differently with templates in the A form and B form. In the previous section, we provided evidence for the presence of two template-binding sites, namely the helix clamp of the p66 thumb and the helix clamp of the p51 thumb in combination with helix α E' of the p66 RNase H domain. The latter is termed the upstream binding region.

Every translocation model for HIV-1 RT has to deal with the observation that RT uses DNA as well as RNA as template. Therefore, it is important to know details about the difference in the interaction of RT with the two template types. There is experimental evidence that shows that the affinity of RT and its processivity are two orders of magnitude higher for RNA than for DNA templates [32, 38], indicating that the template interaction of RT is optimized for A-form substrates. Explanation of this difference on the basis of the known crystal structure of a RT · DNA complex is difficult. The transition of the dsDNA from the A form close to the polymerization site to the B form further upstream might be important for the different template affinities of RT. The finding that the DNA fragment adopts the

A conformation upon binding to RT suggests strong RT-template interaction. Conversely, the transition to the B form in the upstream DNA region might reflect weaker interaction, allowing the dsDNA template to adopt its native form [12, 18]. Therefore, an RNA template with an A conformation in solution [39] should have enhanced affinity and processivity, in agreement with the experimental observation, due to stronger nucleic-acid-protein contacts in the upstream binding region. Consistent with this notion are results from hydroxyl-radical footprinting studies on RT · DNA complexes. The protection pattern shows that the contact region extends to position -18 [27], while DNase I footprinting indicates protection to position -23 on the template [36]. This discrepancy was explained by the assumption that the two probes show different sensitivities to strong and weak interacting sites [36]. From these results, we conclude that, if DNA is used as a template, RT interacts with the upstream binding site weaker than with the binding sites neighboring the polymerase active site. We speculate that the observed difference in the affinity of RT for RNA and DNA templates is due to a lower affinity of DNA to the upstream binding site. Support for this speculation is provided by molecular-modeling studies with RNA and DNA as templates [18].

Evidence for flexible and rigid template-binding sites in HIV-1 RT. As discussed above, there is evidence that the helix-clamp motifs in the thumbs of p51 and p66 provide the main contribution to template-binding in HIV-1 RT. Other parts of RT might also contribute to the affinity, e.g. the primer and template grips. However, for the purpose of the present model, it suffices to show that the contributions can be attributed to either a flexible or a rigid binding site.

The flexibility of the p66 thumb and rigidity of the p51 thumb is the second assumption on which the model is based. Evidence for differences in the flexibilities of the two thumb domains is obtained from X-ray structures. These structures show that the p66 thumb has a smaller contact surface with other domains than has the p51 thumb domain. The latter shares contact surfaces with neighboring subdomains in both subunits, namely an estimated area of 620 \AA^2 with the p51 palm and connection, and an area of 740 \AA^2 with the p66 connection and RNase H [40]. The p66 thumb has a contact surface of only 900 \AA^2 with other subdomains, located exclusively in p66. Direct evidence for flexibility of p66 is provided by comparison of X-ray-structure data of unliganded RT, RT in complex with DNA, and RT in complex with non-nucleosidic inhibitors such as nevirapine, α -anilino-2,6-dibromophenylacetamide, 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine or tetrahydroimidazo[4,5-*l*-*jk*][1,4]benzodiazepin-2(1H)-thione [11, 41–43]. The p66 thumb is rotated into the nucleic-acid-binding cleft by 34° in the absence of nucleic acid [12, 28]. In contrast, the p66 thumb is in an upright orientation when DNA is bound. A similar extended geometry of the thumb was observed in complexes of RT with non-nucleosidic inhibitor [11, 44, 45]. It was suggested that the change of thumb orientation is effected by binding of the non-nucleosidic inhibitor in a hydrophobic pocket close to the hinge of the p66 thumb at its interface with the palm subdomain, which arrests the thumb in a hyperextended state [11, 44]. The binding pocket for non-nucleosidic inhibitors is filled by neighboring amino acid side chains in the unliganded state, and is open in the complex with DNA [12]. These structural comparisons suggest that the RT-inhibiting function of the non-nucleoside drugs is based on arrest of the extended conformation of the p66 thumb. In this way, movement of the thumb relative to the palm subdomain, which contains the polymerase active site, is prevented.

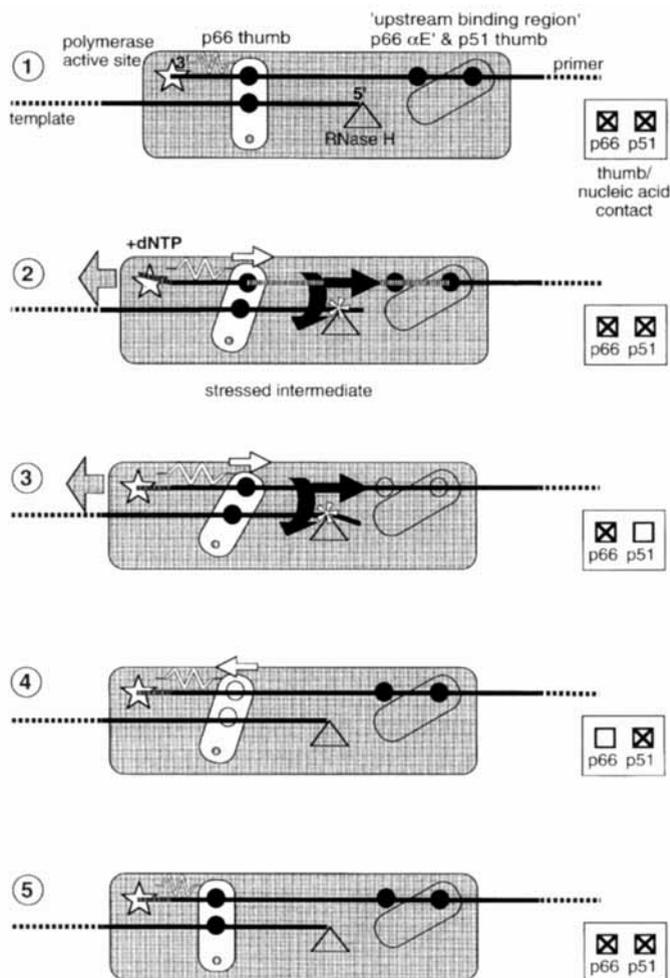


Fig. 4. Mechanism of HIV-1 RT translocation on an RNA template according to the model described in the text. The nucleic-acid-binding structures are shown as open circles. Binding of nucleic acid is marked with filled circles. The primer/template hybrid is displayed as black lines. Newly incorporated nucleoside triphosphate is in grey. Movement of the p66 thumb subdomain is indicated by open arrows, that of the nucleic acid relative to the polymerase by black arrows. Binding or loss of contact is indicated on the right side.

A model for processive translocation of HIV-1 RT: two thumbs work in concert. The translocation model that we suggest assumes that strain is induced in the template by incorporation of a new nucleotide at the 3' end of the primer. This strain enables communication between the two template-binding sites. Thus, coordinated association and dissociation of the two sites is assured in such a way that contact with at least one template-binding site is maintained throughout the translocation. While the flexible template-binding site in the p66 thumb follows the movement of the template substrate, the upstream binding region acts on the template, like a lock on a rack.

Fig. 4 shows the different phases of translocation. In the absence of nucleotide triphosphates, the thumbs of p66 and p51 bind to primer/template so that the 3'-OH terminus of the primer is positioned close to the polymerase active site (Fig. 4). Due to incorporation of a nucleotide at the 3' end of the primer (Fig. 4) strain is induced in the nucleic acid substrate, which is clamped between the two binding sites of RT. We propose that during formation of the phosphodiester bond, strain is generated by substitution of the nucleotide at position n by the incoming nucleotide $n+1$. It has been suggested recently by Arnold's group [25] that incorporation of the incoming base is facilitated by a stack-

ing interaction of bases n and $n+1$. Nucleotide substitution, in concert with phosphodiester-bond formation, leads to movement of the template and the p66 thumb (Fig. 4), which is flexible enough to follow the movement of the nucleic acid. In this way, contact of the p66 thumb is maintained with the template. Due to the rigidity of the upstream binding region, strain in the template facilitates its dissociation from the upstream binding site, which leads to relief of strain within the template and rebinding of the upstream binding region to the template one position further downstream (Fig. 4). We suggest that this re-association facilitates dissociation of the p66 thumb. While the p66 thumb is reverting to its initial upright position, the upstream binding region stabilizes the primer/template on RT.

It is reasonable to assume that the strain in the primer/template has two components, namely a longitudinal component and a torsional component. This correlates with the movement of the primer/template helix, which is translocated along its axis by one base and has to be rotated to ensure that the initial conditions for nucleotide incorporation are re-established.

The different processivity rates of RT on DNA and RNA templates can be explained in the context of the proposed translocation model. If the affinity between RT and DNA is lower due to weaker contacts at the upstream binding site, as discussed above, the DNA can dissociate much more easily.

Conclusion, implications and suggested experimental proofs.

During translocation, RT goes through a labile state, since RT has to reorganize the synthesizing complex after incorporation of the previous nucleoside. One group of RT inhibitors, the non-nucleosidic inhibitors, e.g. tetrahydroimidazo[4,5,*l*-jk][1,4]benzodiazepin-2(1H)-thione, are proposed to act by interfering with the translocation process. To acquire an understanding of the inhibitory function of this group of inhibitors, more information about the translocation process is required. We have presented a model for processive translocation of HIV-1 RT during reverse transcription, which could serve as basis for further discussion. The model can be tested by analysis of RT that contains mutations in the helix-clamp motif, which is assumed to constitute the template-binding site of RT. Use of mixed reconstituted RT that contains one wild-type subunit and one mutated subunit, is one possible approach to determine whether the proposed two-template-site model is correct. The model predicts that reconstituted RT containing mutated p51 has reduced processivity on an RNA template and a processivity similar to that of wild-type RT on a DNA template.

The authors thank the *Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie* and the *Deutsche Forschungsgemeinschaft*, for support, and Thomas Meier, Tatjana Singer, Anne Whelan and especially Georgi Muskhelishvili, for valuable discussion during preparation of the manuscript. T. H. thanks the *Studienstiftung des deutschen Volkes* for support.

REFERENCES

- Poch, O., Sauvaget, I., Delarue, M. & Tordo, N. (1989) *EMBO J.* **8**, 3867–3874.
- Delarue, M., Poch, O., Tordo, N., Moras, D. & Argos, P. (1990) *Protein Eng.* **3**, 461–467.
- Kamer, G. & Argos, P. (1984) *Nucleic Acids Res.* **12**, 7269–7282.
- Argos, P. (1988) *Nucleic Acids Res.* **16**, 9909–9916.
- Polesky, A. H., Steitz, T. A., Grindley, N. D. & Joyce, C. M. (1990) *J. Biol. Chem.* **265**, 14579–14591.
- Beese, L. S., Friedman, J. M. & Steitz, T. A. (1993) *Biochemistry* **32**, 14095–14101.
- Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G. & Steitz, T. A. (1985) *Nature* **313**, 762–766.

- Kim, Y., Eom, S. H., Wang, J., Lee, D.-S., Suh, S. W. & Steitz, T. A. (1995) *Nature* **376**, 612–616.
- Korolev, S., Nayal, M., Barnes, W. M., DiCera, E. & Waksman, G. (1995) *Proc. Natl Acad. Sci. USA* **92**, 9264–9268.
- Arnold, E., Jacobo-Molina, A., Nanni, R. G., Williams, R. L., Lu, X., Ding, J., Clark, A. D. Jr, Zhang, A., Ferris, A. L., Clark, P., Hizi, A. & Hughes, S. H. (1992) *Nature* **357**, 85–89.
- Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. & Steitz, T. A. (1992) *Science* **256**, 1783–1790.
- Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D. Jr, Lu, X., Tantillo, C., Williams, R. L., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H. & Arnold, E. (1993) *Proc. Natl Acad. Sci. USA* **90**, 6320–6324.
- Sousa, R., Chung, Y. J., Rose, J. P. & Wang, B.-C. (1993) *Nature* **364**, 593–599.
- Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H. & Kraut, J. (1994) *Science* **264**, 1891–1903.
- Darst, S. A., Kubalek, E. W. & Kornberg, R. D. (1989) *Nature* **340**, 730–732.
- Schultz, P., Celia, H., Riva, M., Sentenac, A. & Oudet, P. (1993) *EMBO J.* **12**, 2601–2607.
- Joyce, C. M. & Steitz, T. A. (1994) *Annu. Rev. Biochem.* **63**, 777–822.
- Hermann, T., Meier, T., Götte, M. & Heumann, H. (1994) *Nucleic Acids Res.* **22**, 4625–4633.
- Kong, X.-P., Onrust, R., O'Donnell, M. & Kuriyan, J. (1992) *Cell* **69**, 425–437.
- Krishna, T. S. R., Kong, X.-P., Gary, S., Burgers, P. M. & Kuryian, J. (1994) *Cell* **79**, 1233–1243.
- Wyman, C. & Botchan, M. (1995) *Curr. Biol.* **5**, 334–337.
- Tabor, S., Huber, H. E. & Richardson, C. R. (1987) *J. Biol. Chem.* **262**, 16212–16223.
- Metzger, W., Schickor, P. & Heumann, H. (1989) *EMBO J.* **8**, 2745–2754.
- Chamberlin, M. J. (1994) *Harvey Lect.* **88**, 1–21.
- Patel, P. H., Jacobo-Molina, A., Ding, J., Tantillo, C., Clark, A. D., Raag, R., Nanni, R. G., Hughes, S. H. & Arnold, E. (1995) *Biochemistry* **34**, 5351–5363.
- Zaychikov, E., Denissova, L. & Heumann, H. (1995) *Proc. Natl Acad. Sci. USA* **92**, 1739–1743.
- Metzger, W., Hermann, T., Schatz, O., LeGrice, S. F. J. & Heumann, H. (1993) *Proc. Natl Acad. Sci. USA* **90**, 5909–5913.
- Rodgers, D. W., Gamblin, S. J., Harris, B. A., Ray, S., Culp, J. S., Hellmig, B., Woolf, D. J., Debouck, C. & Harrison, S. C. (1995) *Proc. Natl Acad. Sci. USA* **92**, 1222–1226.
- Skalka, A. M. & Goff, S. P. (1993) *Reverse transcriptase*. Cold Spring Harbor Laboratory Press, New York.
- Goff, S. P. (1990) *AIDS (Phila.)* **3**, 817–831.
- Gilboa, E., Mitra, S. W., Goff, S. P. & Baltimore, D. (1980) *Cell* **18**, 93–100.
- Reardon, J. E. (1993) *J. Biol. Chem.* **268**, 8743–8751.
- Cheng, N., Merrill, B. M., Painter, G. R., Frick, L. W. & Furman, P. A. (1993) *Biochemistry* **32**, 7630–7634.
- Kumar, A., Kim, H.-R., Sobol, R. W., Becerra, S. P., Lee, B. J., Hatfield, D. L., Suhadolnik, R. J. & Wilson, S. H. (1993) *Biochemistry* **32**, 7466–7474.
- Nanni, R. G., Ding, J., Jacobo-Molina, A., Hughes, S. H. & Arnold, E. (1993) *Perspect. Drug Discov. Des.* **1**, 129–150.
- Wöhrl, B. M., Tantillo, C., Arnold, E. & LeGrice, S. F. J. (1995) *Biochemistry* **34**, 5343–5350.
- Beard, W. A., Stahl, S. J., Kim, H. R., Bebenek, K., Kumar, A., Strub, M. P., Becerra, S. P., Kunkel, T. A. & Wilson, S. H. (1994) *J. Biol. Chem.* **269**, 28091–28097.
- DeStefano, J. J., Bambara, R. A. & Fay, P. J. (1993) *Biochemistry* **32**, 6908–6915.
- Saenger, W. (1983) *Principles of nucleic acid structure*, Springer, New York.
- Ding, J., Jacobo-Molina, A., Tantillo, C., Lu, X., Nanni, R. G. & Arnold, E. (1994) *J. Mol. Recognit.* **7**, 157–161.
- Ding, J., Das, K., Tantillo, C., Zhang, W., Clark, A. D., Jessen, S., Lu, X., Hsiou, Y., Jacobo-Molina, A., Andries, K., Pauwels, R., Moerrees, H., Koymans, L., Janssen, P. A. J., Smith, R. H., Koepke, M. K., Michejda, C. J., Hughes, S. H. & Arnold, E. (1995) *Structure (Lond.)* **3**, 365–379.

42. Ren, J., Esnouf, R., Garman, E., Somers, D., Ross, C., Kirby, I., Keeling, J., Darby, G., Jones, Y., Stuart, D. & Stammers, D. (1995) *Nat. Struct. Biol.* 2, 293–302.
43. Ding, J., Das, K., Moereels, H., Koymans, L., Andries, K., Janssen, P. A. J., Hughes, S. H. & Arnold, E. (1995) *Nat. Struct. Biol.* 2, 407–415.
44. Smerdon, S. J., Jäger, J., Wang, J., Kohlstaedt, L. A., Chirino, A. J., Friedman, J. M., Rice, P. A. & Steitz, T. A. (1994) *Proc. Natl Acad. Sci. USA* 91, 3911–3915.
45. Tantillo, C., Ding, J., Jacobo-Molina, A., Nanni, R. G., Boyer, P. L., Hughes, S. H., Pauwels, R., Andries, K., Janssen, P. A. J. & Arnold, E. (1994) *J. Mol. Biol.* 243, 369–387.
46. Kornberg, A. & Baker, T. A. (1991) *DNA replication*, Freeman, New York.
47. Bonner, C. A., Stukenberg, P. T., Rajagopalan, M., Eritja, R., O'Donnell, M., McEntee, K., Echols, H. & Goodman, M. F. (1992) *J. Biol. Chem.* 267, 11431–11438.
48. Ricchetti, M. & Buc, H. (1993) *EMBO J.* 12, 387–396.
49. Dong, Q., Copeland, W. C. & Wang, T. S. (1993) *J. Biol. Chem.* 268, 24175–24182.
50. McDonald, W. F. & Traktman, P. (1994) *J. Biol. Chem.* 269, 31190–31197.
51. Tsurumi, T. (1991) *Biochem. J.* 280, 703–708.