Annu. Rev. Biochem. 1994. 63:777–822

FUNCTION AND STRUCTURE RELATIONSHIPS IN DNA POLYMERASES

Catherine M. Joyce and Thomas A. Steitz[1]
Department of Molecular Biophysics and Biochemistry, Yale University, Bass Center for Molecular and Structural Biology, 266 Whitney Avenue, New Haven, Connecticut 06520

KEY WORDS: protein structure, X-ray crystallography, DNA-protein interactions, exonuclease, ribonuclease H

CONTENTS

INTRODUCTION

Of fundamental importance to all living organisms is the ability to synthesize DNA efficiently and accurately, thus ensuring the faithful transmission of genetic information from parent to offspring. To achieve this objective, all

[1]and Howard Hughes Medical Institute

0066-4154/94/0701-0777$05.00

free-living organisms encode several DNA polymerases that fulfill the various replicative and repair functions within the cell. Additionally, many viruses bypass the host replicative machinery and encode their own polymerases. At first glance, the DNA polymerases seem a bewilderingly disparate group of enzymes, ranging from the small mammalian repair polymerase β, a single subunit of 39 kDa, to the huge multisubunit replicative polymerases, exemplified by DNA polymerase III holoenzyme of *Escherichia coli*, which has at least 20 subunits and a combined molecular mass close to 900 kDa (1). At an intermediate stage of complexity are polymerases such as *E. coli* DNA polymerase I, having multiple enzymatic activities within a single, fairly large, polypeptide chain.

Despite the rich variety within the DNA polymerase family, it seems likely that, at heart, all DNA polymerases are variations on a single theme. In every case the function of the core polymerase activity is to add deoxynucleotides onto the growing end of a DNA primer strand; the difference between repairing a short patch of DNA in a bacterium and replicating the 46 chromosomes of *Homo sapiens* is one of scale, not of chemistry. The underlying similarity between DNA polymerases is suggested by complementation in vivo between polymerases of quite different types (2–4). Studies of polymerase accuracy, an important attribute of enzymes of this type, also imply that many (though not all) of the physicochemical mechanisms used to discriminate between correct and incorrect basepairs have been preserved throughout this family of enzymes (5, and references cited therein).

Perhaps the most compelling argument for an underlying similarity among polymerases is provided by the analysis of protein sequences; although there is little discernible resemblance overall between the sequences of distantly related polymerases, it appears that a small number of crucial active-site residues are conserved (summarized in Figure 1). The sequence conservation not only encompasses the currently identified families of DNA polymerases (6), but can also be extended to reverse transcriptases, RNA replicases, and DNA-dependent RNA polymerases (7), implying that the basic mechanism of phosphoryl transfer required for polynucleotide synthesis is preserved, but that different enzyme families possess variable features that differentiate between utilization of ribo- or deoxyribo-substrates. That such features are relatively subtle and not central to the overall reaction mechanism is suggested by the observations that polymerases can utilize their "unnatural substrate," albeit at suboptimal efficiency or when minor adjustments are made to the reaction conditions (1, 8), and that the retroviral life-cycle requires that reverse transcriptases use both RNA and DNA templates.

A valuable consequence of the probable similarity between polymerases is that structure-function studies pursued on those polymerases that are

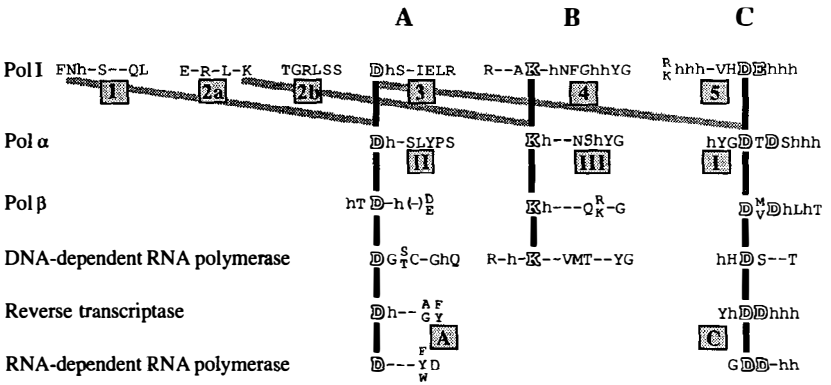


Figure 1 Alignment of the major conserved sequence regions of the polymerase families. The listed motifs are based on published compilations: Ref. 6 for the Pol I and Pol α families, Ref. 104 for the Pol β family and DNA-dependent RNA polymerases, and Ref. 182 for the RNA-dependent polymerases. Positions that are almost invariably occupied by a hydrophobic amino acid are indicated by "h." Hyphens denote nonconserved positions; parentheses are used to indicate length variations within a motif. Following published conventions, conserved sequence blocks in the Pol I family are numbered 1 through 5 according to Delarue et al (7), with the addition of motif 2a according to Blanco et al (183), the Pol α sequences are labeled according to Wong et al (184), and the RNA-dependent polymerase sequences according to Poch et al (182). For clarity, some of the motifs that are conserved within individual families are not shown. The black lines indicate a proposed alignment (7) that gives two motifs, A and C (labeled at the top of the figure), common to the entire polymerase family, containing two invariant aspartates and another highly conserved acidic residue (in motif C). Motif B, containing an invariant lysine, is common to DNA-dependent polymerases (7, 104). As discussed later, mutagenesis experiments provide evidence of the importance of these invariant amino acids (highlighted in the figure). An alternative alignment between the Pol I and Pol α families indicated by shaded lines (183) seems to us to be less satisfactory, since it gives less good conservation of the carboxylates and more variable spacing in regions corresponding to α-helices in the Klenow fragment structure.

simpler and more amenable to such studies are likely to provide information of relevance to all polymerases. Currently, the high-resolution structures of two DNA polymerases have been published (9–11). They are the Klenow fragment of *E. coli* DNA polymerase I and the reverse transcriptase from the human immunodeficiency virus HIV-1. In this review we explore the inferences that can be drawn from a detailed examination of these two structures, together with the available biochemical and genetic data, and the extent to which these deductions can be applied to other polymerases. The focus of our review is confined to structural and mechanistic aspects of the reactions catalyzed by DNA polymerases; to place this information in its biological context, the reader is referred to the book by Kornberg & Baker (1), and to recent reviews in this series on prokaryotic DNA replication

(12), eukaryotic DNA polymerases (13), and DNA replication fidelity (14, 15).

POLYMERASE STRUCTURES

The three-dimensional structures of Klenow fragment and HIV-1 reverse transcriptase will be described in detail. Additionally, though not strictly within the scope of this review, the structure of the DNA-dependent RNA polymerase from bacteriophage T7 (16) will be considered briefly, since its similarity to the other two structures lends additional weight to the idea of a fundamental similarity among all polymerases. Several other polymerase structures are in progress so that, in the not too distant future, it should be possible to evaluate the ideas presented here in the context of a wider range of polymerase structures.

Klenow fragment and HIV-1 reverse transcriptase, representing two distinct classes of DNA polymerase, show substantial differences in their overall molecular properties. Klenow fragment normally uses a DNA template (although it can use RNA, Ref. 8), while reverse transcriptase uses both RNA and DNA. Both enzymes have an associated nuclease domain, but the nucleases serve different functions in the two cases. In Klenow fragment, the associated 3′-5′ exonuclease functions to edit polymerase errors by removing incorrectly incorporated nucleotides from the primer terminus (1). The RNase H activity of reverse transcriptase cleaves the RNA template strand of the RNA-DNA duplex product (17). There are also differences in quaternary structure; Klenow fragment is a monomer, while HIV-1 reverse transcriptase is a dimer of different-sized subunits derived from the same sequence. In contrast to these rather obvious global differences between the two enzymes, there are equally striking similarities between the two when the polymerase active-site regions are examined in detail.

Structure of Klenow Fragment

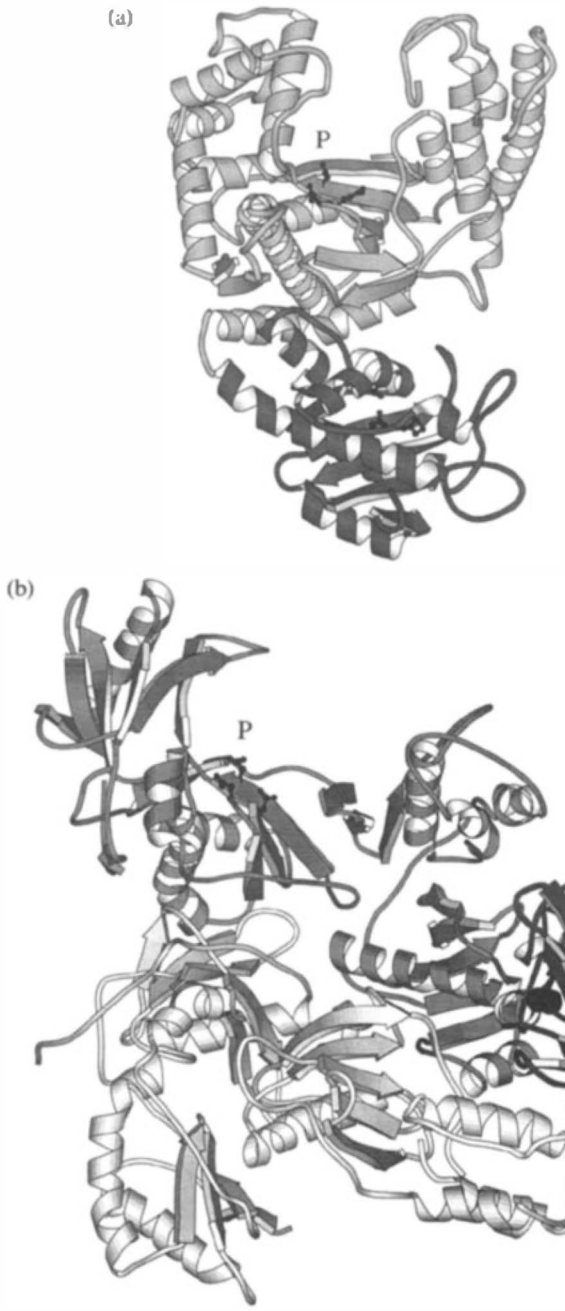
Three different crystalline complexes have been solved from a tetragonal crystal form grown from high salt. The complex of Klenow fragment with deoxynucleoside monophosphate at the exonuclease active site has been refined at 2.5 Å resolution (LS Beese, TA Steitz, unpublished). The complex with oligo(dT)$_4$ at the exonuclease site has been refined at 2.6 Å resolution (18). Finally, an editing complex containing 11 basepairs of duplex DNA has been refined at 3.2 Å resolution (19). A trigonal crystal form grown at low ionic strength from polyethylene glycol diffracts anisotropically to 2.8–3.5 Å resolution and has been partially refined (JM Friedman, TA Steitz, unpublished). It should be noted that the model-built structures derived by Modak and colleagues (20) are incorrect in many details and

should not be considered appropriate substitutes for the experimentally derived structures.

The overall structure of Klenow fragment, initially determined at 3.3 Å resolution, showed very clearly that the 68-kDa molecule consists of two domains (Figure 2*a*): a larger C-terminal domain with a very prominent cleft, and a smaller globular N-terminal domain (9). A large number of experiments have shown conclusively that the 3′-5′ exonuclease active site is located on the small N-terminal domain of Klenow fragment, while the polymerase active site is located in the cleft of the larger C-terminal domain. Thus, in the initial crystallographic studies, deoxynucleoside monophosphate, the product of the exonuclease reaction, was found to bind to the small domain (9); more recently, a binding site for deoxynucleoside triphosphate, the substrate for the polymerase reaction, has been observed on the large domain (21). As described in more detail below, the proposed locations for the exonuclease active site (around the deoxynucleoside monophosphate binding site) and the polymerase active site (within the cleft of the large domain) are strongly supported by site-directed mutagenesis data and protein sequence alignments. Moreover, expression of the DNA encoding the large C-terminal domain gave a protein product that had DNA polymerase activity but no exonuclease activity (22).

The polymerase domain of Klenow fragment can be divided into three subdomains, named "palm," "fingers," and "thumb" for their anatomical analogy to a right hand (Figure 3). The palm subdomain contains a β-sheet that forms the base of the cleft and contains the catalytic residues (see below); the fingers subdomain is virtually all α-helix and forms one wall of the cleft, while the thumb subdomain has two long antiparallel α-helices that interact as coiled coils forming the other side of the cleft. A 50-amino-acid region at the tip of the thumb subdomain was partially disordered in the original description of the structure (9). This region is now seen, both in the low-salt trigonal crystal form and in the complex with duplex DNA, to contain two additional short α-helices and some connecting strands (19). The tips of the fingers and thumb subdomains are in contact so that the cleft is, in fact, more accurately described as a tunnel.

Binding sites for divalent metal ions (magnesium, manganese, and zinc) have been located crystallographically on both the 3′-5′ exonuclease and the polymerase domains. Two metal ions are bound to carboxylate ligands at the 3′-5′ exonuclease active site in the presence of deoxynucleoside monophosphate (18, 23) and have been proposed to play a pivotal role in catalysis, as described in a later section. Magnesium, manganese, or zinc ions bind to the polymerase domain at the bottom of the cleft if the high-salt crystals are transferred from ammonium sulfate into lithium sulfate (LS Beese, TA Steitz, unpublished). No crystallographic information is available


(a)
P
(b)
P

on the binding of metal ions to the polymerase active site in the presence of DNA and deoxynucleoside triphosphate substrates.

Structure of HIV-1 Reverse Transcriptase

HIV-1 reverse transcriptase can employ either RNA or DNA as a template, yielding either RNA-DNA hybrid or duplex DNA products (24). DNA synthesis from the viral RNA as template is initiated in vivo by human $tRNA_3^{Lys}$, whose 3′ end partially unfolds and forms 18 basepairs of duplex with the viral RNA primer-binding site (24, 25). The HIV-1 reverse transcriptase is processed initially from the pol gene products as a 66-kDa polypeptide having an N-terminal polymerase and a C-terminal RNase H domain (26, 27). Subsequent proteolytic cleavage of the homodimer of the 66-kDa subunits removes the RNase H domain from one subunit, leaving a heterodimer containing one 66-kDa subunit (p66) and one 51-kDa subunit (p51) (28, 29). In the p66-p51 heterodimer, only the p66 subunit has polymerase activity, as shown by the observation that mutations of essential active-site residues eliminate polymerase activity only when present on the p66 subunit (30, 31). The p66 subunit also contains the single binding site for non-nucleoside inhibitors such as Nevirapine (32). The heterodimer, but not the isolated subunits, interacts in 1:1 stoichiometry with the tRNA primer (33). A dimeric structure may be typical of most retroviral reverse transcriptases (24); recent results (34) suggest that even those enzymes, such as Moloney murine leukemia virus reverse transcriptase, which are monomeric in solution, may dimerize when bound to their substrate nucleic acid (although in this case the enzymatically active form would be a homodimer).

Since HIV-1 reverse transcriptase is the target of currently approved anti-AIDS drugs, extensive efforts have been directed towards establishing its structure. Initial crystallization experiments led to numerous crystal forms (35, 36), none of which diffracted beyond 6 Å resolution. Two strategies aimed at reducing the inherent flexibility of this polymerase, and consequently improving the resolution to which crystals diffract, have been successful: cocrystallization with an inhibitor, and cocrystallization with a monoclonal antibody F_{ab} fragment and DNA. High-resolution structures have

Figure 2 (*a*) Schematic representation of the Klenow fragment structure, with α-helices shown as spiral ribbons and β-strands as arrows. The extent of the separate 3′-5′ exonuclease domain is indicated by darker shading. The catalytically important carboxylate side chains at the polymerase (P) and exonuclease sites are shown. (*b*) A similar representation of HIV-1 reverse transcriptase. The p51 subunit of the heterodimer is shown with the lightest shading. Within the p66 subunit, the RNase H domain is more darkly shaded to distinguish it from the polymerase and connection regions. The catalytically important carboxylate side chains at the polymerase site (P) are shown; the two metal ions at the RNase H site are shown as black spheres. These figures were made by Joe Jäger.

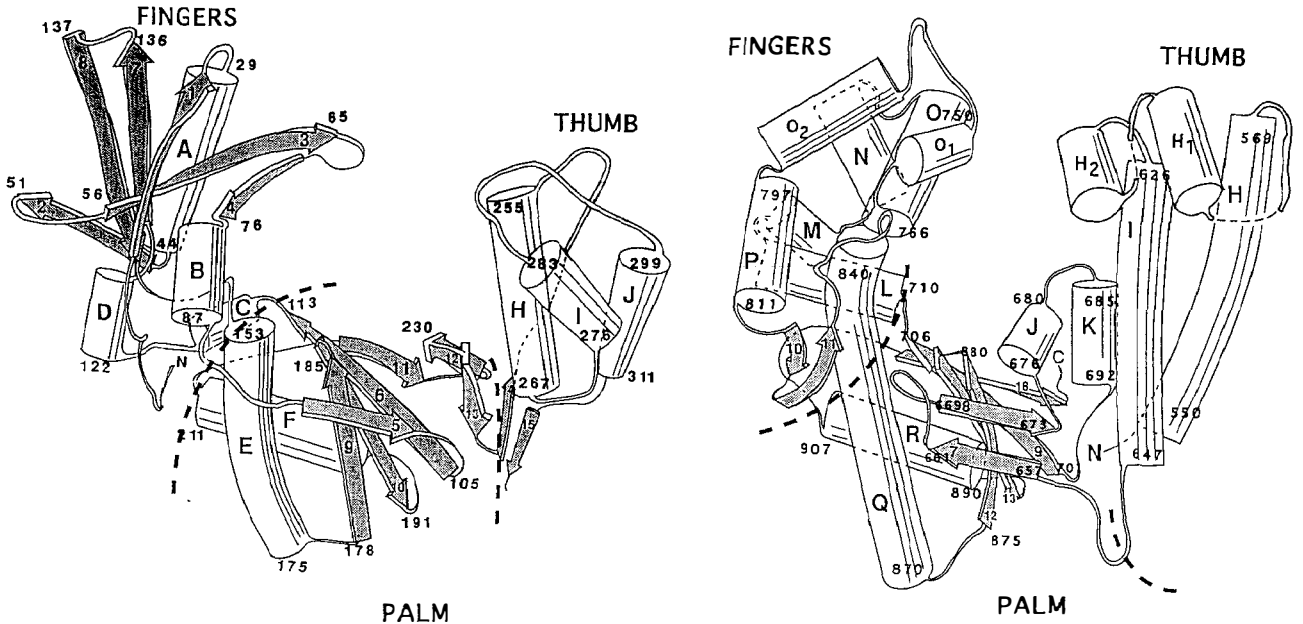


Figure 3 Schematic representation of the polymerase domain of HIV-1 reverse transcriptase (*lefthand panel*) and Klenow fragment (*righthand panel*) with α-helices shown as cylinders and β-strands as arrows (reproduced from Ref. 37). The molecules are positioned with their palm subdomains oriented identically. The amino acid sequence numbers give an approximate indication of the boundaries of secondary structural features. Note that further refinement of the Klenow fragment structure (LS Beese, TA Steitz, *unpublished*) has allowed identification of four additional helices (H_1, H_2, O_1, and O_2) that were not apparent on the original 3.3 Å structure (9). As previously noted (37), the similarities observed in the two polymerase structures do not exactly parallel the sequence alignments proposed by Delarue et al (7). Thus, the two structures diverge immediately following the invariant aspartate of motif A (although the extended sequence alignment may well be structurally justified within individual polymerase families). Conversely, although there is no apparent primary sequence similarity, one can use the structural similarities to extend motif C, which encompasses the sequences of β-strands 12 and 13 in Klenow fragment and 9 and 10 in reverse transcriptase, so as to overlap helices Q (Klenow fragment) with E (reverse transcriptase) and helices R (Klenow fragment) with F (reverse trancriptase).

been reported for the complex of HIV-1 reverse transcriptase with the non-nucleoside inhibitor Nevirapine (initially at 3.5 Å, and now partially refined at 2.9 Å resolution, Refs. 10, 37), and for the complex with an antibody F_{ab} fragment and 18 basepairs of double-stranded DNA (at 3.0 Å resolution, Ref. 11). Additionally, the structure of the isolated RNase H domain of HIV-1 reverse transcriptase has been solved at 2.4 Å (38), and two separate determinations of the structure of the homologous *E. coli* RNase H, at 2.0 and 1.5 Å resolution, have been described (39, 40).

The structure of the polymerase-proficient p66 subunit of HIV-1 reverse transcriptase is analogous to that of Klenow fragment in several ways. It has a very pronounced domain structure (Figure 2*b*), which reflects its enzymatic activities. The RNase H activity resides on a separate C-terminal domain joined via the "connection" subdomain to the three subdomains that form the polymerase region (10). The three polymerase subdomains can be described as fingers, palm, and thumb of a right hand, and together they form a large cleft, reminiscent of the cleft in the polymerase domain of Klenow fragment (Figure 3). The binding sites for two divalent metal ions on the RNase H domain have been located crystallographically (10, 38), but the binding sites for divalent metal ions on the polymerase domain have yet to be determined. Although comparison of the two reverse transcriptase complexes (one with Nevirapine, the other with DNA and an F_{ab} fragment) shows that the overall structure of the protein is similar in the two complexes, there are significant differences in the relative orientations of the subdomains, particularly the RNase H (see Figure 8, in a later section). It is unclear which of the three ligands is responsible for the differences in the structure or whether crystal packing influences the structure.

The structure provides an obvious rationale for the lack of polymerase activity in the p51 subunit of the heterodimer (30, 31). Although the two subunits have the same amino acid sequence throughout the polymerase region, and similar tertiary structures within each subdomain, the overall conformation of the polymerase domain of p66 is astonishingly different from that of p51, due to differences in the relative positioning of subdomains (10). If the palm subdomains of the two subunits are identically oriented, then the p51 subunit has the fingers closer to the palm and the thumb subdomain further from the fingers and palm. The connection subdomain lies within and fills the expanded cleft between thumb and palm in p51, so that the p51 subunit has no cleft and hence no binding site for the primer-template. The polymerase domain of p66 is not related to p51 by a two-fold rotation axis; rather, the two domains interact in a more head-to-tail arrangement that results from a 16 Å translation and an approximately 74° rotation of one palm subdomain relative to the other.

Comparison of Polymerase Domain Structures in Reverse Transcriptase and Klenow Fragment

A comparison of the crystal structures of the two polymerase domains (10, 37, 41) shows that the three subdomains (palm, fingers, and thumb) that form the cleft (Figure 3) appear to be functionally related. The palm subdomains that lie at the bottom of the polymerase cleft in both enzymes show substantial structural similarity (10), such that it is possible to superimpose 45 Cα atoms from two α-helices and three β-strands with an rms deviation of 1.6 Å (Figure 4). Particularly significant is the virtually identical positioning of three carboxylate residues (Asp705, Asp882, and Glu883 of Klenow fragment; Asp110, Asp185, and Asp186 of reverse transcriptase), which have been shown by mutagenesis studies to be crucial for polymerase activity (42–46). In the alignment of all polymerase sequences proposed by Delarue et al (7), two of these residues are the invariant aspartates of sequence motifs A and C; the third (the Asp186/Glu883 residue) is highly conserved (Figure 1). Multiple magnesium or manganese ions have been observed to bind to this trio of carboxyl groups in Klenow fragment (LS Beese, TA Steitz, unpublished; 10, 19). Thus it appears that these carboxyl groups with associated divalent metal ions form the catalytic center of these two polymerases and are a vital feature that is common to all polymerases.

Figure 4 Superposition of a portion of the palm subdomains of Klenow fragment (in darker shading) and HIV-1 reverse transcriptase, showing the α-carbon backbone and the three conserved carboxylate residues. Reproduced from Ref. 37.

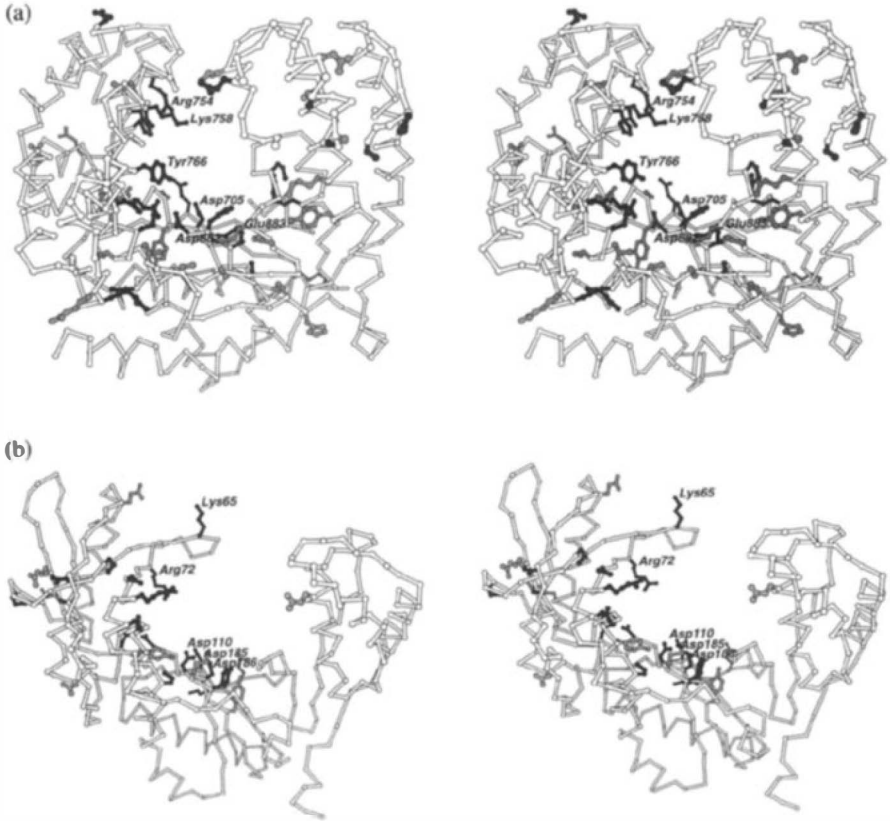


Figure 5 (*a*) Stereo representation of the α-carbon backbone of the polymerase domain of Klenow fragment showing side chains that are conserved in DNA polymerases of the Pol I family (6). The darker-shaded residues are identical in at least eight of the nine sequences; those with lighter shading are conserved in at least six of the sequences. (*b*) A similar representation of the polymerase domain of HIV-1 reverse transcriptase showing side chains that are conserved in the alignment of five reverse transcriptase sequences (26). The darker-shaded side chains are identical in the five sequences; lighter shading indicates side chains that are conserved in four out of the five sequences. In both (*a*) and (*b*), only solvent-accessible side chains are shown. Reproduced from Ref. 37.

In both Klenow fragment and reverse transcriptase, the functional importance of the cleft region could also be inferred from the clustering, particularly at the base of the cleft, of side chains that are invariant or highly conserved in the respective polymerase families (Figure 5).

The thumb subdomains of the two polymerase structures appear to be functionally analogous rather than truly structurally homologous. In reverse transcriptase the thumb consists of three α-helices and one extended strand,

which had previously been tentatively assigned as an α-helix (10, 11, 37). In Klenow fragment the thumb is made up of two longer antiparallel α-helices with a small globular domain at its tip consisting of two short helices and short stretches of random coil (19). Not only do the structures of the thumb regions differ in detail, but their relative location in the gene sequence is also different in the two enzymes. The thumb sequence occurs before the palm and fingers sequences in Klenow fragment, but after these sequences in reverse transcriptase. In both polymerases, however, the thumb subdomain appears to serve the same purpose, interacting with the minor groove of the duplex nucleic acid product of DNA synthesis, and moving in response to the binding of DNA (11, 19, 47) (see below).

The fingers subdomains of Klenow fragment and reverse transcriptase bear no structural resemblance to one another. In Klenow fragment this region is predominantly α-helical, whereas in reverse transcriptase it is mixed α-helix and β-sheet. Consistent with the lack of any structural relatedness, the fingers subdomain of Klenow fragment contains the B sequence motif (Figure 1), which is present only in DNA-dependent polymerases and is not seen in RNA-dependent polymerases such as the reverse transcriptase family (7). This motif is located on the O helix of Klenow fragment (Figure 3), which has no structural counterpart in reverse transcriptase. Nevertheless, as discussed below, it is possible that the fingers region fulfills the same role in both polymerases, that of interacting with the template strand close to the site of synthesis.

Comparison with the Structure of T7 RNA Polymerase

The recently reported structure of T7 RNA polymerase (16) shows a striking similarity to Klenow fragment. Not only do the conserved A, B, and C motifs occur in similar spatial relationships, with conserved residues similarly located, but nearly all the secondary structure elements within the polymerase domain are equivalently located. Mutational studies have demonstrated the importance of Asp537 and Asp812, corresponding to the invariant carboxylates of motifs A and C respectively, and Lys631, an invariant residue of motif B in DNA-directed polymerases (48, 49). Like the two structures described above, T7 RNA polymerase also seems to be built in a modular fashion, having an N-terminal domain that is structurally unrelated to any of the domains of Klenow fragment or reverse transcriptase, and that appears to carry out functions specific to RNA synthesis.

POLYMERASE ACTIVE SITE

A complete structural description of the enzymatic mechanism of the polymerase reaction and the enzyme's role in maintaining accuracy will

require a high-resolution structure of a stable ternary complex containing both substrates, the deoxynucleoside triphosphate and the primer-template. While such information is not yet available, structural data from binary complexes with either deoxynucleoside triphosphate or duplex DNA, combined with extensive site-directed mutagenesis and kinetic studies, provide some useful insights into active-site location and the role of particular side chains in the polymerase reaction.

As indicated above, distinct functions in the polymerase reaction have been proposed for each of the three subdomains of the polymerase region (10). The palm subdomain contains the catalytic center, the binding site for the 3′ terminus of the primer strand, and contributes to the dNTP-binding site. From the available structural and sequence data, it seems likely to be the most conserved part among all polymerases. The fingers subdomain binds and orients the template strand across from the primer terminus and may also form part of the dNTP-binding site. The three published polymerase structures (9–11, 16), together with protein sequence alignments (7), suggest that there are at least two families of structures for the fingers: The mixed α-helix and β-sheet structure found in reverse transcriptase is hypothesized to occur in polymerases that use RNA as a template; the α-helical structure found in Klenow fragment and T7 RNA polymerase is predicted to be typical of DNA-dependent polymerases. The helical thumb, which is attached flexibly to the rest of the polymerase domain, contacts the minor groove of the product duplex (11, 19). The flexibility of this subdomain may be important in allowing access of the primer-template to the binding site, and in translocation of the product after dNTP incorporation.

Binding of Primer-Template

Essential to understanding the structural basis of polymerase function is a detailed knowledge of the complex between the primer-template DNA and the enzyme. Although there is as yet no high-resolution structure of a complex between Klenow fragment and duplex DNA in which the 3′ end of the primer strand is bound in the polymerase active site, a complex with an 11-basepair duplex DNA having a 3-nucleotide 3′ overhang bound at the 3′-5′ exonuclease site has been partially refined at 3.2 Å resolution (19). While this structure corresponds to an editing complex, the location of the duplex portion of the DNA may provide useful insights into the polymerase primer-template complex. Recently, the structure of HIV-1 reverse transcriptase complexed with both a monoclonal F_{ab} and an 18-basepair duplex DNA with a single-nucleotide 5′ overhang bound to the polymerase active site has been solved at 3 Å resolution. These structures, taken together with the results of model-building, mutagenesis, and chemical modification,

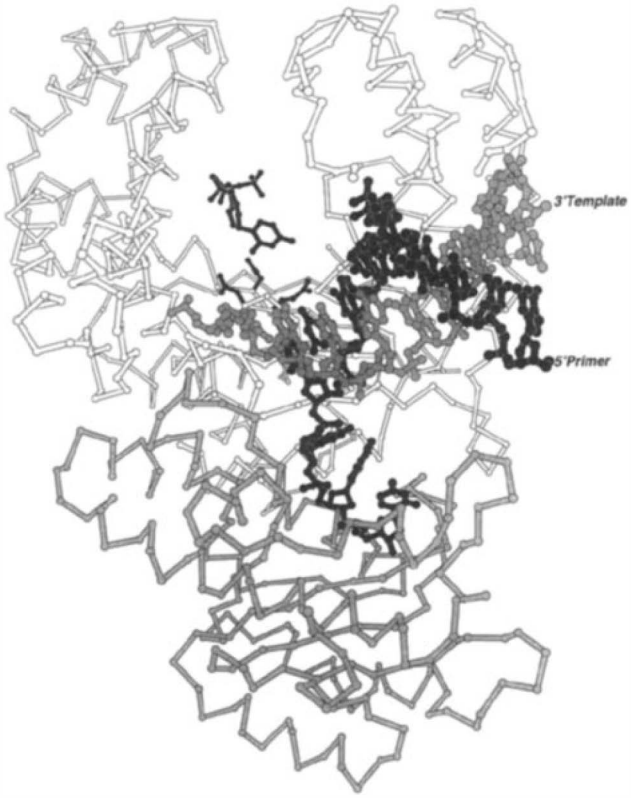


Figure 6 Structure of Klenow fragment complexed with DNA. The α-carbon backbone of Klenow fragment is shown, with the 3′-5′ exonuclease domain lightly shaded. The DNA, shown as a full-atom representation, was derived from the cocrystal structure of an editing complex (19). The position of the dCTP molecule shown was derived from the structure of the enzyme-dNTP binary complex (21). The three conserved carboxylates at the polymerase active site are shown. Reproduced from Ref. 37.

provide a useful, if incomplete, picture of primer-template binding to the polymerase active site.

Although only one cleft was apparent in the initial Klenow fragment structure (9), the cocrystal structure of Klenow fragment with duplex DNA (19) shows the existence of two deep clefts that lie nearly at right angles to each other (Figure 6). The cleft within the polymerase domain was identified as the probable location of the polymerase active site initially because of the cluster of residues that are conserved between T7 DNA polymerase and Klenow fragment (9, 50). Its size and extensive positive

electrostatic potential (51) made it a plausible binding site for the primer-template, and this inference was supported by the location within the cleft of two mutations that cause defects in DNA binding (52). Although the binding of primer-template to the polymerase cleft is now established, the accumulating evidence has recently been recognized to indicate that the initially modeled direction of DNA synthesis (9), in which the primer strand entered the cleft from the end furthest from the 3′-5′ exonuclease domain, is not correct (10, 19). The initial model was inconsistent with the increasingly precise localization of the polymerase active site provided by subsequent mutagenesis studies (45, 46). Moreover, the orientation of DNA in the polymerase cleft of Klenow fragment was opposite to that deduced for HIV-1 reverse transcriptase (10), even though the high degree of structural homology in the polymerase clefts of these two proteins would require a similar mode of binding if the active sites were to function analogously. The cocrystal structure of Klenow fragment with duplex DNA indicated how DNA could be bound to Klenow fragment so as to approach the polymerase active site from the same direction as in reverse transcriptase (19). In the cocrystal, the 11-basepair duplex DNA was bound in a second cleft whose axis is roughly orthogonal to the cleft that contains the polymerase active site (Figures 6 and 7). Four single-stranded nucleotides at the 3′ end of the primer strand bound to the exonuclease active site identically to the previously described complexes with single-stranded DNA (18, 53). This second cleft is in part formed by a significant movement of the thumb towards the DNA with which it is interacting and by stabilization of the subdomain at the tip of the thumb that is highly disordered in the apo-structure (9). Extensive interactions are seen across the minor groove between the backbone phosphates of the duplex DNA and amino acid side chains from the thumb subdomain that are highly conserved in the Pol I family (6). These include Arg631 and Lys635 (region 2a, Figure 1), Asn675 and Asn678 (just beyond region 2b), and a conserved motif (region 1) at the tip of the thumb. The conservation of residues that interact with the duplex DNA implies that the binding site observed in the editing complex is biochemically relevant, and supports the contention that it may also be the duplex DNA-binding site employed when DNA is bound to the polymerase active site. In this case, the primer strand would approach the polymerase catalytic site from the direction of the 3′-5′ exonuclease domain and the large cleft in the polymerase domain would bind the single-stranded template strand beyond the site of DNA synthesis (Figure 7).

The location of the polymerase active site centered on the three carboxylates, Asp705, Asp882, and Glu883 (Figure 6), together with the position of the duplex DNA observed in the cocrystal structure, implies that the duplex DNA upstream of the primer terminus is severely bent when the

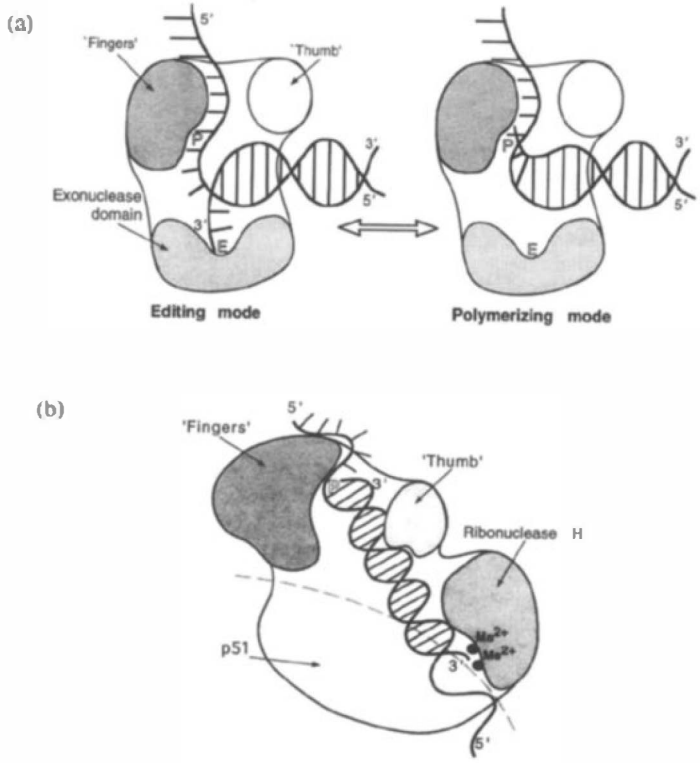


Figure 7 Direction of DNA synthesis and relative location of polymerase ("P") and nuclease active sites in Klenow fragment (*a*) and HIV-1 reverse transcriptase (*b*). For Klenow fragment, "E" indicates the 3′-5′ exonuclease active site; in reverse transcriptase, the two metal ions mark the location of the RNase H active site. Note that the direction of DNA synthesis relative to the conserved polymerase catalytic subdomain appears likely to be the same for both polymerases. Reproduced from Ref. 41.

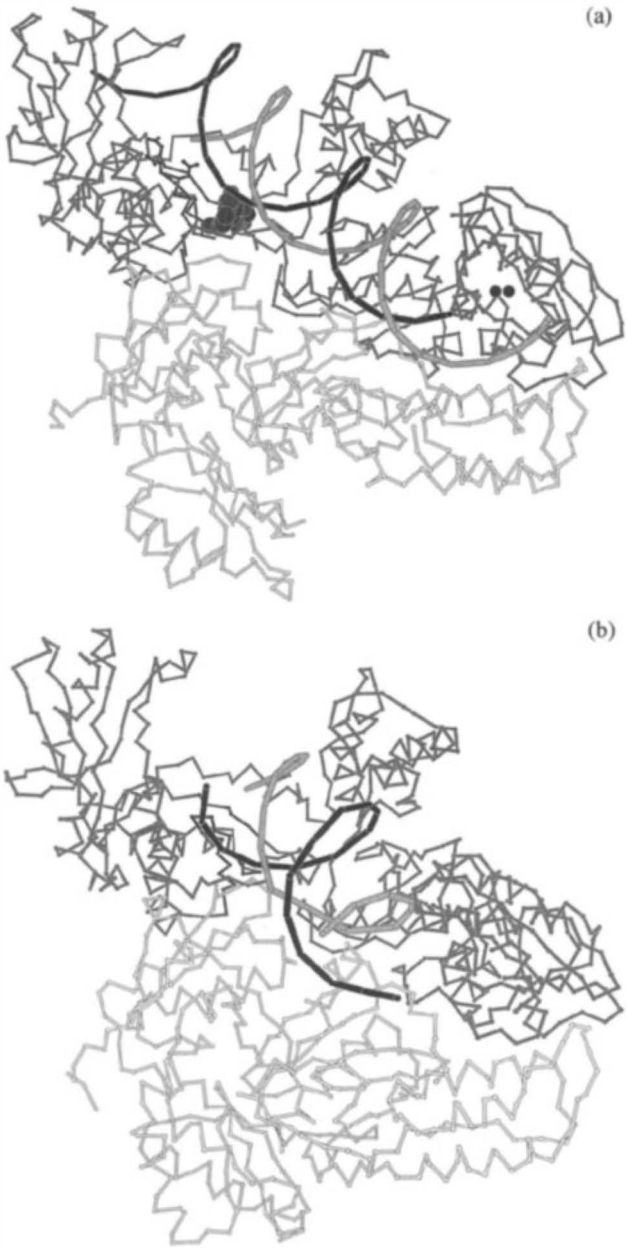
primer terminus is in the polymerase active site. While this predicted DNA bend and the narrowness of the polymerase cleft discouraged detailed model building, placement of the primer terminus near the catalytic carboxylates necessarily positions the template strand in contact with the fingers subdomain (Figure 7). Therefore, in addition to any role the fingers domain might have in binding the deoxynucleoside triphosphate substrate (see below), it must also play a significant role in binding the template strand (19).

The model deduced for binding of DNA to the polymerase site of Klenow fragment is broadly consistent with a number of biochemical studies. It is important to realize, however, that, because a large area of the protein is

proposed to be in contact with the DNA, and because of limitations in the resolution of the techniques used, very few of these experiments provide information at a sufficient level of detail to constrain the model-building process. Thus, the observation that mutation or chemical modification of residues at many different locations within the polymerase cleft has an effect on DNA binding (45, 46, 54–56) is as expected; indeed, it would be surprising if this were not so. The labeling of Tyr766 by a photoaffinity probe at the primer terminus is likewise consistent with the general location of the polymerase catalytic site (57). Chemical footprinting (58), fluorescence (59, 60), and photocrosslinking (57) experiments together indicate that 5–8 basepairs of duplex DNA are covered by Klenow fragment when the primer terminus is at the polymerase site. The range of values probably reflects, on the one hand, the particular requirements for access of the footprinting reagent and, on the other, uncertainties as to the precise orientation of fluorescent or photoactivatable probes attached to DNA; again, the data are consistent with the proposed model but are insufficiently discriminating to rule out alternatives that could be proposed. Using time-resolved fluorescence spectroscopy, it has been concluded that more of the duplex DNA upstream of the primer terminus is drawn into the binding site when the 3′ terminus is at the 3′-5′ exonuclease active site (60). The proposed model (19) suggests a difference of 2–3 basepairs between the polymerase and exonuclease binding modes; it does not appear able to accommodate the 9-nucleotide difference inferred from experiments that compared the effect of a bulky DNA substituent on the two enzymatic reactions (61). This latter experimental result may, however, be reconcilable with the model if, for example, there were a region at the junction of the two clefts where a bulky substituent on the DNA could be tolerated.

When the first high-resolution crystal structure of HIV-1 reverse transcriptase was determined, a model of a binary complex with an A-form DNA-RNA hybrid was constructed by assuming that the 3′ end of the primer strand should be placed near to the catalytic carboxylates, Asp110, Asp185, and Asp186, and that a phosphate of the RNA template strand should lie adjacent to the metal ions at the RNase H active site (10). Employing a rise per residue of 3 Å, 19–20 nucleotides were estimated to lie between these two active sites (Figure 8*a*). Whereas earlier biochemical studies had suggested a distance of 15–16 nucleotides between the polymerase and RNase H active sites (62), more recent experiments, using short incubation times and single-turnover conditions, gave a distance of 18–19 nucleotides, in good agreement with the structural model (63, 64).

More recently, the structure of HIV-1 reverse transcriptase complexed with a duplex DNA oligonucleotide and the F_{ab} fragment of a non-inhibiting antibody has been described at 3.0 Å resolution (11). The positioning of


(a)
(b)

the duplex DNA observed experimentally is nearly identical to the model-built A-form DNA-RNA hybrid. However, the duplex DNA is more nearly B-form except in the vicinity of the polymerase active site (where it is A-form), and is significantly distorted (Figure 8*b*), with a bend at the junction of the A-form and B-form segments. This bend is not as large as the one predicted to occur in the Klenow fragment complex and is in the opposite direction.

The structure of the DNA complex raises the intriguing possibility that reverse transcriptases may constrain their template-primer to adopt an A-form conformation at the polymerase site (11); this would permit the use of either RNA or DNA templates, as required during retroviral replication, since both can be A-form, whereas RNA does not adopt a B-form conformation. Model-building the template strand, either in the RNA-DNA hybrid (10) or by extending from the DNA complex (11) (assuming, in either case, that the template continues as an A-form helix beyond the primer terminus) indicates contacts with the fingers subdomain, with the antiparallel β-ribbon formed by strands 3 and 4 (Figure 3) fitting the minor groove side of a template in A-form (Figure 9). By contrast, DNA polymerases, such as Klenow fragment, might favor binding of a B-form template-primer, substantially reducing the efficiency of RNA-templated reactions (8). Thus, it is tempting to speculate that the fingers subdomain (whose structure, as noted earlier, is very different in Klenow fragment and HIV-1 reverse transcriptase) plays a significant role in determining the template specificity of a polymerase, perhaps by constraining the substrate into the appropriate A-form or B-form conformation at the polymerase active site.

The binding of a DNA primer-template to crystals of reverse transcriptase results in substantial movement of the thumb subdomain towards the DNA (11, 47), as was observed with Klenow fragment. Just as in Klenow fragment, the thumb interacts with the phosphate backbone across the minor groove of the product duplex (10, 11). Consistent with the structural data, UV irradiation of short oligonucleotides (acting as model primers) bound to HIV-1 reverse transcriptase resulted in crosslinking to amino acid residues

Figure 8 (*a*) The α-carbon backbone of the HIV-1 reverse transcriptase heterodimer with a model of an A-form RNA-DNA hybrid duplex. The p66 subunit is shaded to distinguish it from p51. The bound Nevirapine inhibitor is shown in space-filling representation. The 3′ end of the primer (DNA) strand (gray) is positioned adjacent to the carboxylate side chains of Asp110, Asp185, and Asp186 that define the catalytic center of the polymerase active site; these side chains and the three β-strands on which they are located are shown in black. The two spheres indicate the positions of the two divalent metal ions at the RNase H active site and are seen to be adjacent to the backbone of the RNA template strand (black). (*b*) A similar representation to (*a*), showing the experimentally determined position of a B-form DNA duplex (11). The location of the polymerase active site is marked by the three β-strands shown in black. These figures were made by Steve Smerdon.

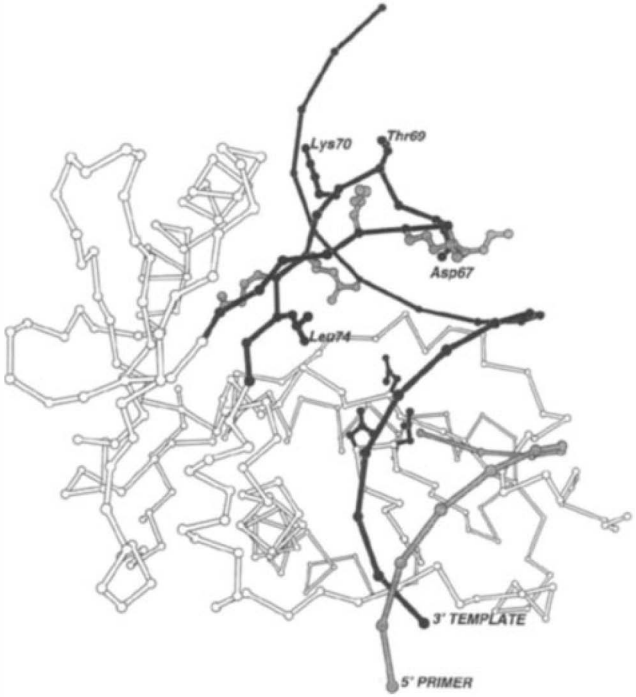


Figure 9 The palm and fingers subdomain of HIV-1 reverse transcriptase, showing a possible interaction between the template strand beyond the site of synthesis and residues in the loop between β-strands 3 and 4 in the fingers subdomain. The model-built primer strand (lightly shaded) has its 3′ terminus close to the three conserved carboxylate side chains, Asp110, Asp185, and Asp186, which are shown in black. The darker template strand of the A-form duplex is suitably placed to interact with the loop, whose backbone atoms are indicated by shading. Several residues whose mutation renders the virus resistant to the nucleoside analogs, AZT, ddI, and ddC, are located in this loop (Asp67, Thr69, Lys70, and Leu74). The approximate positions of the corresponding side chains are shown by dark shading. The remaining positively charged residues (Lys64, Lys65, Lys66, Arg72, and Lys73) are lightly shaded. Reproduced from Ref. 37.

located in the thumb, with Leu289-Thr290 and Leu295-Thr296 as the probable sites of labeling (65, 66).

Unlike other retroviral reverse transcriptases, HIV-1 reverse transcriptase is inhibited noncompetitively by several distinct classes of non-nucleoside compounds whose chemical structures appear largely unrelated to each other, although the isolation of cross-resistant mutants suggests that these compounds share a common binding site (67). When HIV-1 reverse transcriptase was cocrystallized with the non-nucleoside inhibitor Nevirapine (68), the molecule was observed to bind to a hydrophobic pocket that lies at the base

of the thumb at the junction between the thumb and palm subdomains (10) (Figure 8*a*). Consistent with its behavior as a noncompetitive inhibitor (68), the Nevirapine-binding site does not overlap the observed or expected binding sites for either substrate of the polymerase reaction. The structure suggests that inhibition by this molecule may be a consequence of a reduction in the mobility of the thumb subdomain (10), since the binding of Nevirapine may immobilize the thumb in one orientation relative to the palm. It is possible that this fixed orientation may not be precisely correct for positioning the DNA substrate in a manner appropriate for catalysis. Alternatively (or additionally) the reduced ability of the thumb to undergo necessary conformational changes (a sort of molecular arthritis) may inhibit translocation of the duplex product.

The template-directed synthesis of DNA by HIV-1 reverse transcriptase is initiated by the 3′ end of human $tRNA_3^{Lys}$. Eighteen nucleotides at the 3′ end of the tRNA pair with a specific complementary sequence in the viral template (the primer-binding site). It has been hypothesized that the D-stem and anticodon stem of the tRNA, which remain folded, interact with the p51 subunit (10); this is consistent with the observed binding of the tRNA primer to the heterodimer of reverse transcriptase, but not to the isolated subunits (33). Chemical crosslinking suggests that the anticodon region is in close proximity to the protein (33). Moreover, comparisons of the binding of unmodified tRNAs made in vitro and $tRNA_3^{Lys}$ modified in vivo (69, 70) indicate that the modifications on the anticodon loop contribute to the specificity of the interaction of HIV-1 reverse transcriptase with its primer. Other interactions between the tRNA and the protein are independent of RNA sequence (69, 70). Additional selectivity for the primer tRNA is provided by complementary basepairing interactions with the viral RNA template, involving the unraveled 18 nucleotides at the 3′ end of the tRNA, and 6 nucleotides of the anticodon loop (69; LA Kohlstaedt, TA Steitz, unpublished). In the case of HIV-2 RNA, an additional interaction of about 6 nucleotides of the T-stem with the RNA template is predicted (71). Since reverse transcriptase alone can initiate synthesis from the tRNA primer (69), it must facilitate the partial unfolding of the tRNA and formation of a complex structure with the template.

Nucleotide Binding

Our understanding of the polymerase reaction will not be complete without the identification of the side chains responsible for positioning the dNTP substrate, particularly as it is likely that a subset of these interactions will play important roles as the reaction proceeds. Information on polymerase-dNTP contacts can, in principle, be derived from a combination of crystallographic studies of complexes containing the dNTP, affinity labeling and

chemical modification experiments, and the study of mutations that affect the binding of nucleotides or their analogs. With all of these approaches, however, there are caveats that preclude an unambiguous interpretation of the results, so that our understanding of dNTP-binding remains at a rather primitive level. In the first place, although a polymerase-dNTP binary complex can be formed, such a complex is not catalytically competent (72). This makes sense since the requirement for complementarity between the incoming dNTP and the DNA template dictates a substantial degree of contact between the nucleotide base and the template-primer. However, it means that any structural or labeling data obtained with the binary complex must be interpreted with caution. Secondly, a clear-cut interpretation of mutational data is difficult, since a mutation could influence dNTP binding either via its effect on dNTP contact residues, or via an effect on template contacts, resulting in a subtle change in position of the opposing template base at the dNTP-binding site. With these caveats in mind, we will examine the available experimental data on dNTP binding.

High-resolution crystallographic studies have been carried out on a binary complex formed between Klenow fragment and dNTP after the tetragonal crystals grown in high salt were transferred to low ionic strength (21). Although a complex between 5-Hg UTP, primer-template DNA, and HIV-1 reverse transcriptase has also been reported (47), the only information presented was the position of the Hg substituent at 7 Å resolution. In the Klenow fragment-dNTP complex, the multiple transfers of the crystals resulted in some loss of resolution; however, a plausible model for the binary complex could be constructed (Figure 10). Consistent with the ideas discussed in the preceding paragraph, the conformation of the nucleotide base was not uniquely defined, but varied depending on the identity of the dNTP. In each case the base made van der Waals and hydrophobic interactions with residues at the bottom of the polymerase cleft, interactions that cannot occur when the base is hydrogen-bonded to the opposing template base. The location of the nucleotide base is consistent with chemical crosslinking experiments that indicated proximity to Tyr766 and His881 in the binary polymerase-dNTP complex (73, 74), implying that the crystalline and solution binary complexes are the same.

Although no useful conclusions could be drawn from the positions of the nucleotide base or sugar, it is possible that the crystalline binary complex may be informative in identifying contacts between Klenow fragment and the dNTP phosphate groups. The β and γ phosphates appear to make interactions with three positively charged side chains within the cleft (Figure 10): Arg754 and Lys758 on helix O of the fingers subdomain, and Arg682 on the thumb. Both Arg682 and Lys758 had previously been suggested as dNTP contacts on the basis of affinity-labeling experiments (75, 76); the

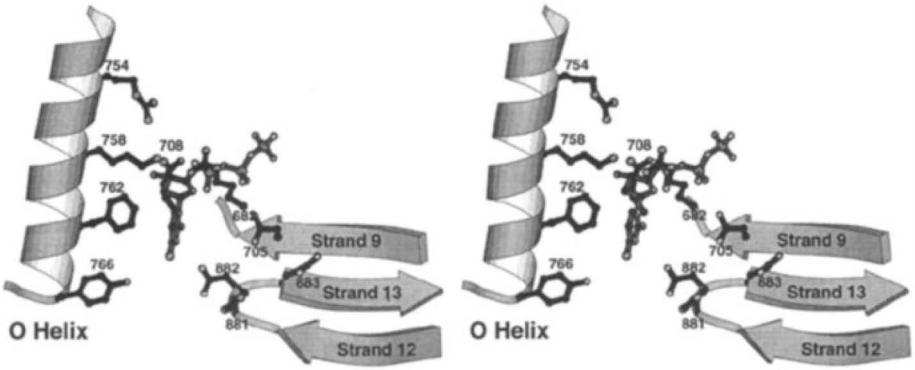


Figure 10 Stereo drawing of the dCTP-binding site of Klenow fragment (21). The protein is viewed looking down into the polymerase cleft, approximately perpendicular to the view shown in Figure 3. The side chains of the trio of carboxylates, and of other residues that contact the dCTP molecule, are shown in black and numbered. The dCTP molecule is shown in gray. Relevant secondary structural features are indicated. This figure was made by Joe Jäger.

data for Arg682 seem particularly persuasive since labeling was carried out in the presence of primer-template DNA, so as to probe dNTP binding in the ternary complex (75). However, it is disturbing that mutation of either of these residues had little or no effect on $K_{m(dNTP)}$ (77; M Astatke, CM Joyce, unpublished work). The positioning of the β and γ phosphates inferred from the Klenow fragment-dNTP complex is consistent with data from a binary complex of Klenow fragment with pyrophosphate (21), and would place the deoxyribose close to Phe762 and the α-phosphate about 6 Å from the catalytically important side chains of Asp705 and Asp882.

While biochemical experiments with HIV-1 reverse transcriptase have identified the general region of the polymerase active site, a detailed interpretation of the role of particular residues cannot be made with the methods used. Photoaffinity labeling of a binary complex with dTTP resulted in crosslinking to Lys73 (78), and an antibody inhibition study suggested that dNTP binds in the vicinity of residues 65–73 (79). However, the residues implicated by these studies are located on the β-hairpin that has been proposed, from crystallographic work, to bind to the template strand (Figure 9), and Lys73 is more than 15 Å from the primer terminus (10, 11, 37; J Jäger, TA Steitz, unpublished). It is possible that in the binary complex the thymine base is not in a fixed orientation (as is the case in the dNTP complex with Klenow fragment), allowing crosslinking with Lys73 to occur. Labeling experiments using pyridoxal phosphate were interpreted

as indicating a role in dNTP binding for Lys263 (80), which is, however, located on the thumb subdomain, also at a considerable distance from the position of the primer terminus (37). Furthermore, a subsequent mutational study (81) likewise argues against this interpretation, and kinetic data suggest that pyridoxal phosphate is not a good probe for the dNTP site, at least for this enzyme (82).

A different approach aimed at pinpointing active-site residues that serve as important enzyme-substrate contacts is provided by kinetic experiments that probe the binding of dNTPs to mutant DNA polymerases. In Klenow fragment, mutations that have been found to affect the binding of dNTP in the ternary complex (as reflected in $K_{m(dNTP)}$) are located on one side of the polymerase cleft within or close to the fingers subdomain. Positions identified thus far encompass the N terminus of helix Q (Arg841 and Asn845), the exposed face of helix O (Tyr766, Phe762, and Arg754), and neighboring residues closer to the catalytic center (Asp705 and Glu710) (45, 46; M Astatke, CM Joyce, unpublished). Since these regions include the conserved sequence motifs A (around Asp705) and B (helix O), it is gratifying to note that mutational analyses of the corresponding regions of human DNA-polymerase α and ϕ29 DNA polymerase have also identified side chains that may play a role in dNTP binding (83–85). An advantage of the kinetic approach is that it probes the ternary complex; however, as discussed above, it is impossible, in the absence of other structural evidence, to distinguish direct effects from those mediated via template interactions. Moreover, the side chains listed above encompass an area much larger than the dNTP molecule and therefore cannot all be in direct contact with it. Since the region of Klenow fragment implicated by these studies is thought to make extensive contacts with the template strand, a reasonable interpretation is that a subset of the residues mentioned above are in direct contact with the dNTP, while the remainder bind the template DNA.

In a similar way, studies on mutations that influence the binding of DNA polymerase inhibitors have the potential to provide information on polymerase-dNTP contacts. However, the location of mutations in HIV-1 reverse transcriptase that result in resistance of the virus to one or more of the chain-terminating nucleoside analogs AZT, ddI, and ddC (86–89) is entirely consistent with the conclusion that the mutated side chains are involved in positioning the template strand (10). Among the residues whose mutation confers drug resistance, Asp67, Thr69, and Lys70 are on the antiparallel β-hairpin that is proposed to bind and orient the template strand (Figure 9). Residues Thr215 and Lys217, at the junction between fingers and palm, also appear to interact with the model-buiit template strand (10), while Leu74 and Met41 could stabilize and position the antiparallel β-hairpin. An intriguing parallel may also exist in Klenow fragment: Mutation of Tyr766,

which is located on the fingers domain in the vicinity of the model-built template strand, affects the discrimination between deoxy and dideoxy nucleotide substrates (CM Joyce, unpublished) and also influences polymerase fidelity at the level of dNTP insertion (90).

In conclusion, it is clear that in the catalytically competent ternary complex the dNTP α-phosphate must be close to the carboxylates at the polymerase catalytic center. The precise position of the rest of the molecule remains to be established, although the available data are suggestive of contacts with the neighboring part of the palm subdomain and perhaps with the fingers subdomain. Contacts with the dNTP base must be provided by the primer terminus and the template strand, whose position also seems likely to be determined by contacts with the fingers region.

Catalysis of the Polymerase Reaction

The polymerase active site catalyzes a nucleophilic attack by the 3′ hydroxyl of the primer terminus on the dNTP α-phosphate, with release of pyrophosphate. Allowing for relatively subtle variations to accommodate ribo- or deoxyribo-derivatives, it seems reasonable to predict an essentially similar chemical pathway for all polymerases, regardless of whether synthesis is templated by DNA or RNA or whether the final product is DNA or RNA. The stereochemical outcome of the reaction, determined for several DNA polymerases, is consistent with an "in-line" nucleophilic displacement proceeding via a pentacovalent transition state or intermediate (91–96). The reaction pathway has been extensively characterized for several polymerases, including Klenow fragment (97), the DNA polymerases of bacteriophages T7 (98) and T4 (99), and HIV-1 reverse transcriptase (63, 100, 101). In every case there is an obligatory order of substrate binding, with initial formation of the enzyme–nucleic acid binary complex being required for productive binding of the dNTP to form a catalytically competent ternary complex. For at least some of the enzymes studied, the reaction pathway includes, in addition to the chemical step, one or more nonchemical transformations (63, 97, 98). Thus, for Klenow fragment, the chemical step is both preceded and followed by slow nonchemical processes that may play a part in maintaining polymerase accuracy (102, 103). These nonchemical steps are frequently described as conformational changes, although the precise nature of such changes is unknown at present. It is important to bear in mind that such a change could involve either enzyme or substrate molecules (or both), and need not necessarily be large in structural terms in order to be kinetically significant.

An active-site residue that plays a role in catalysis may do so by accelerating any one of the steps of the reaction. Depending on the nature of the side chain, it could intervene directly in the chemical step, e.g. as

a general acid or base, it could accelerate either the chemical or nonchemical steps by preferential binding to the relevant transition state, or it could be binding a catalytically essential cofactor such as a metal ion. It is probably unrealistic to categorize active-site residues as involved exclusively in either substrate binding or catalysis, since some overlap between these functions is inevitable; moreover, an unambiguous assignment of mechanistic roles is virtually impossible in the absence of detailed structural information on the relevant substrate complexes. Thus far, most of the experimental evidence for the participation of particular residues in catalysis has been provided by mutagenesis studies. Removal of a putative catalytic residue by mutation should cause a substantial decrease in the reaction rate; further studies on the mutant protein can then dissect further the involvement (if any) of the side chain in substrate binding, and identify the particular step of the reaction that is affected. In some cases, the study of a series of substitutions at a particular position may lead to inferences about the mechanistic role of the mutated side chain.

In the complex of HIV-1 reverse transcriptase with an oligonucleotide (11), the DNA primer terminus interacts with the palm region of the polymerase domain, placing the 3′-hydroxyl, the attacking nucleophile in the polymerase reaction, close to the three proposed catalytic aspartate side chains (positions 110, 185, and 186) (Figure 11*a*). The structurally and functionally equivalent residues in Klenow fragment are Asp705, Asp882, and Glu883. Substitutions of Asp705 or Asp882 are among the most deleterious mutations isolated to date in Klenow fragment, while the effect of mutations at Glu883 is more modest (45, 46). The role of the three carboxylate residues in Klenow fragment has been examined further by determining the elemental effect for incorporation of an α-thio-substituted dNTP. The results indicated that Asp882 and Glu883 may play some role in the chemical step of the polymerase reaction, with the side chain of Asp882 being located extremely close to the dNTP α-phosphate at which nucleophilic substitution takes place (46). By contrast, the kinetically important contribution of Asp705 is felt at a different step of the reaction, perhaps the preceding conformational change (46; CM Joyce, unpublished work).

Other residues in Klenow fragment that are important in accelerating the reaction include Gln849 (implicated by elemental effect measurements in the chemical step), Arg668, and Lys758 (46; M Astatke, CM Joyce, unpublished). Although widely separated on the primary sequence, the side chains of these residues and the three important carboxylates are sufficiently close in the tertiary structure to define an active site. Gln849 is located on helix Q of Klenow fragment, which is structurally analogous to helix E of the p66 polymerase domain of reverse transcriptase (10, 37) (Figure 3).

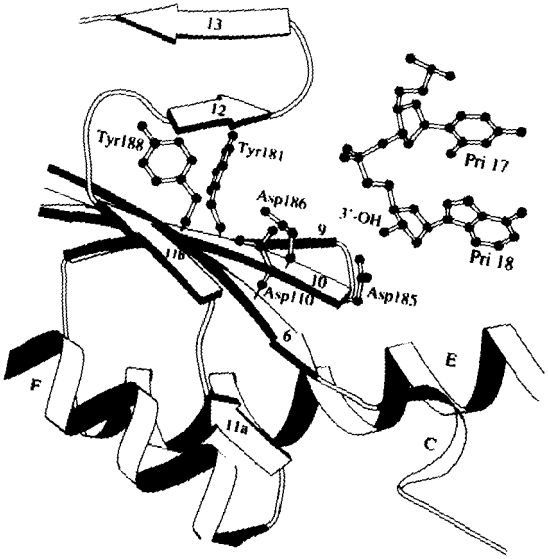
Arg668, a residue that is invariant in the Pol I family of polymerases, is on a neighboring β-hairpin that has no obvious counterpart in the reverse transcriptase structure. Lys758, an invariant residue in the sequence motif B of DNA-dependent polymerases (Figure 1; Refs. 7, 104), is located on helix O in the fingers subdomain of Klenow fragment. Although this region shows no structural resemblance to reverse transcriptase (10), mutagenesis data raise the possibility that the fingers subdomain of reverse transcriptase may also contribute a catalytically important lysine residue, Lys65 (44, 105).

The presence of a few highly conserved residues within the entire polymerase family could reflect a common active-site architecture, and mutagenesis experiments on polymerases for which no structural data are available lend additional support to this idea. Within the two-carboxylate motif C (Figure 1), even conservative substitutions at the carboxylate positions cause a dramatic loss of activity in a variety of polymerases: several members of the DNA polymerase α family (e.g. Refs. 106–109), mammalian DNA polymerase β (110), and the RNA replicase of encephalomyocarditis virus (111). The effects of substitutions at the noncarboxylate positions in motif C are more variable (106, 107, 110) and may, in at least some cases, reflect the degree of structural disruption caused by the mutation (e.g. Ref. 112).

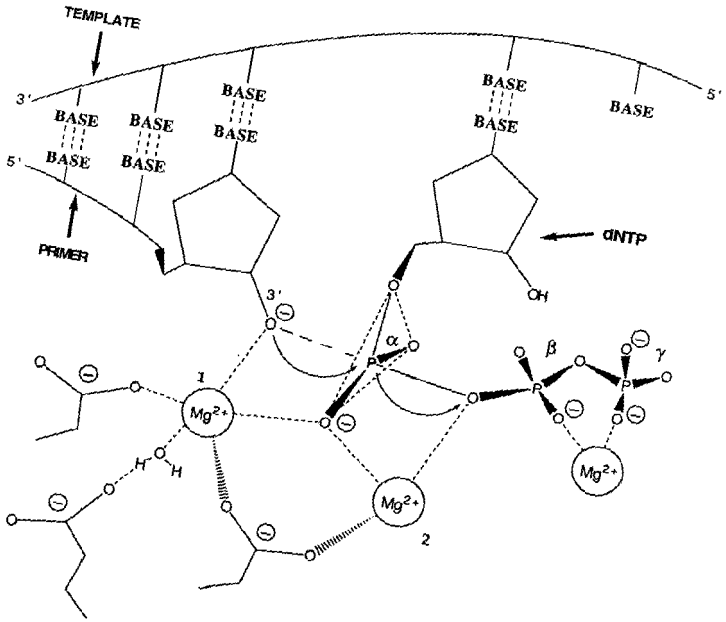
In a similar way, mutagenesis studies on ϕ29 and PRD1 DNA polymerases (108, 113) and encephalomyocarditis virus RNA replicase (111) support the importance of the invariant aspartate of motif A. Surprisingly, however, mutation of the corresponding residue had little effect on the polymerase activity of human DNA polymerase α, the prototype sequence for the ϕ29 and PRD1 polymerases (83). Mutagenesis of the invariant lysine of motif B in PRD1 DNA polymerase and T7 RNA polymerase suggests an important role for this side chain (49, 108). Other mutations within motifs A and B generally had less dramatic overall effects on polymerase activity, but frequently affected the binding of dNTPs or nucleotide analogs (83–85, 113), consistent with the idea, discussed above, that some of the residues in motifs A and B may be involved—either directly or indirectly—in positioning the dNTP.

In summary, the available structural data together with mutagenesis studies in a variety of polymerases suggest a common polymerase active-site structure containing three crucial carboxylate side chains and probably also a lysine. Additional polar residues, e.g. Arg668 and Gln849 in Klenow fragment, may well be important, but their degree of conservation within the polymerase superfamily is at present unclear. There are many ways in which a constellation of polar side chains of this type could catalyze the polymerase reaction. One possibility, given the large number of active-site

(a)


13
12
Tyr188
Tyr181
Pri 17
Asp186
3'-OH
9
Pri 18
11b
10
Asp110
Asp185
6
E
F
11a
C

(b)


TEMPLATE
3'
5'
BASE
PRIMER
dNTP
OH
α
β
γ
P
O
Mg2+
1
2
H

carboxylates, is that these side chains serve to anchor a pair of divalent metal ions that promote catalysis by a mechanism (Figure 11*b*) analogous to that described below for the 3′-5′ exonuclease reaction (18, 41, 53, 114). In this mechanism, one Mg^{2+} (number 1) would promote the deprotonation of the 3′ hydroxyl of the primer strand, analogous to the role of metal A in the exonuclease active site. The other Mg^{2+} (number 2) would facilitate the formation of the pentacovalent transition state at the α-phosphate of the dNTP and the loss of pyrophosphate. Consistent with the idea of catalysis mediated by metal ions, crystallographic experiments have shown binding of Mg^{2+} and Mn^{2+} to Asp705 and Asp882 of Klenow fragment (LS Beese, TA Steitz, unpublished; Refs. 10, 19). Despite the attractiveness of the two-metal-ion mechanism, the available evidence does not exclude related mechanisms in which the necessary deprotonation of the attacking 3′ hydroxyl is achieved by one of the active-site carboxylates acting as a general base. Clearly, a full understanding of the mechanism of catalysis is unlikely in the absence of structural data for a complex containing both substrates at the polymerase active site.

In addition to understanding the mechanistic pathway of the polymerase reaction, it has long been a goal to describe polymerase fidelity in molecular terms. The effect on fidelity of mutations that alter the relative activities of the polymerase and editing functions is well understood (14), but less is known about the two fidelity-determining processes that take place at the polymerase site, namely the discrimination against incorrect dNTPs in the insertion step and against further extension of a mismatched primer terminus. Biochemical studies on polymerase fidelity have suggested a model based on geometric criteria; the polymerase tolerates an incorrect substrate to a degree related to its resemblance to normal DNA geometry (14). From this model one would predict that at least some of the residues that determine accuracy would be those responsible for positioning the dNTP or the primer terminus in the active site and that mutations in these residues would have a mutator or antimutator phenotype. Mutations that affect the fidelity of the polymerase reaction have been described for several DNA polymerases (e.g. 83, 90, 115–119). However, the absence of a large mutational database, particularly for polymerases whose structure is known, precludes a coherent molecular description of polymerase fidelity at present. Comparison of HIV-1

Figure 11 (*a*) The polymerase active-site region in HIV-1 reverse transcriptase, showing the position of the 3′ terminal dinucleotide of the primer strand, the trio of catalytic carboxylates (Asp110, Asp185, and Asp186), and Tyr181 and Tyr188, which are involved in binding non-nucleoside inhibitors, such as Nevirapine. Reproduced from Ref. 11. (*b*) A possible mechanism for the polymerase reaction, involving catalysis mediated by two divalent metal ions. For details see the text. Based on the conclusions of Burgers & Eckstein (91), a third Mg^{2+} is shown chelated by the dNTP β and γ phosphates. Reproduced from Ref. 41.

reverse transcriptase with Klenow fragment suggests one structural feature that may play a role in polymerase fidelity. HIV-1 reverse transcriptase has an exceptionally high rate of synthesis errors that appear to result from template-primer misalignments (120). It is tempting to speculate that the reverse transcriptase molecule with its more open DNA-binding cleft (Figure 2) may be better able to accommodate the postulated misaligned intermediates (37).

OTHER ACTIVITIES PRESENT IN DNA POLYMERASES

Many DNA polymerases also have nuclease activities, which are frequently part of the same polypeptide chain as the polymerase region. The available evidence suggests that the arrangement seen in the Klenow fragment and HIV-1 reverse transcriptase structures, in which the nuclease active site is located on a separate structural domain, is widespread. Polymerases can therefore be thought of as made up of separate enzymatically active modules that appear in various combinations in the different polymerase families.

3′-5′ Exonuclease

Many DNA polymerases (exemplified by Klenow fragment) have an associated 3′-5′ exonuclease that acts in opposition to the direction of DNA synthesis and serves to remove, or proofread, polymerase errors (1). In most cases the 3′-5′ exonuclease is part of the same polypeptide chain as the DNA polymerase, although DNA polymerase III of *E. coli* has its editing exonuclease as a separate subunit (ε) within the core polymerase (121). The preferred substrate for the exonuclease is single-stranded DNA (122), and crystallographic studies have demonstrated the molecular basis for this preference. Thus, when an eight-basepair duplex DNA was cocrystallized with Klenow fragment, four single-stranded nucleotides resulting from partial melting of the duplex were seen bound to the exonuclease active site in a binding pocket that is capable of binding single-stranded (but not duplex) DNA (53). Biochemical experiments using covalently crosslinked DNA duplexes have also shown a requirement for local melting at the 3′ terminus (61). It has therefore been concluded that the DNA primer terminus is bound as a "frayed" or single-stranded end at the 3′-5′ exonuclease active site, even when the enzyme is working on a duplex DNA substrate (53, 58).

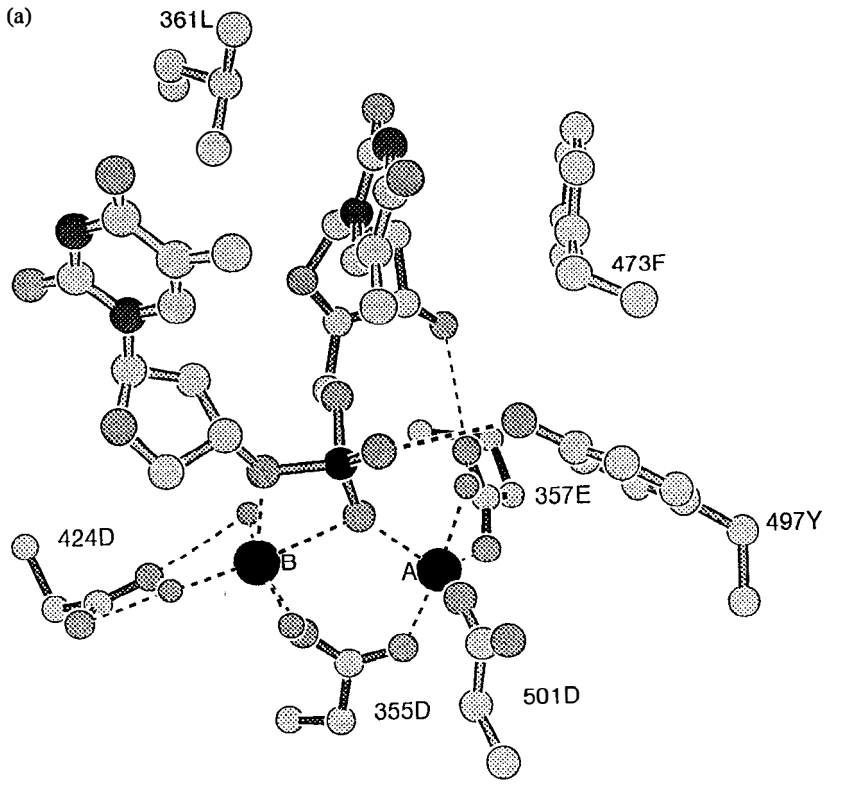
The mechanism of the 3′-5′ exonuclease reaction is probably the best understood of the reactions catalyzed by DNA polymerases, thanks to the availability of high-resolution structural data for crystalline complexes of Klenow fragment with the product (dNMP) or the substrate (a single-stranded DNA oligomer) bound at the exonuclease active site (9, 18, 53). The amino

acid side chains that bind the metal ions and the deoxynucleotides, and thus form the active site, were identified from these structures (Figure 12*a*), thereby providing the basis for a detailed mutagenesis study of their roles in the reaction (23, 123).

The structural and mutagenesis studies demonstrated that the major catalytic role in the 3′-5′ exonuclease reaction is played by a pair of divalent metal ions, A and B (Figure 12*a*), 4 Å apart, that are coordinated to the oxygens of the phosphodiester bond that is to be cleaved (18). In addition to the phosphate oxygen, metal A is bound to the protein in distorted tetrahedral geometry by the carboxylate groups of Asp355, Glu357, and Asp501. Metal B has octahedral coordination; it shares the Asp355 ligand with the metal A site and is coordinated via bridging water molecules to Asp424. Crystallographic studies have indicated that both sites can be filled by Mg^{2+}, Zn^{2+}, or Mn^{2+} in the presence of dNMP (18); likewise, biochemical studies have shown that any of these three metal ions alone can support the 3′-5′ exonuclease reaction (124). When Klenow fragment crystals were placed in a solution containing Mg^{2+} and Zn^{2+}, site A bound Zn^{2+} and site B bound Mg^{2+}, as expected from the tetrahedral geometry of site A and the octahedral geometry of site B. This may be indicative of the preferred combination of metal ions in vivo. Although spectroscopic data suggest the binding of a third metal ion at or close to the active site (125), no additional binding site has been observed crystallographically in either the substrate or the product complex using Mg^{2+} at concentrations up to 10 mM (18).

Mutations of the aspartic acid side chains that serve as ligands to the metal ions (primarily Asp355, Asp424, and Asp501) were shown crystallographically to disrupt metal binding, and also caused dramatic decreases in 3′-5′ exonuclease activity (23, 123). Since the mutations included one (D424A) that abolished binding of metal B (but not A) (23), and one (D355A) that caused loss of metal A (but not B) (19), it is clear that both metal ions are required for the reaction. Moreover, since mutations in the three aspartate ligands caused the largest decreases ($>10^4$ fold) in exonuclease activity of any active-site mutation that has been tested (123), these two metal ions alone appear likely to play the central role in the chemistry of catalysis.

Refinement of the structures of complexes having single-stranded DNA at the 3′-5′ exonuclease active site gave a detailed view of substrate interactions and suggested a mechanism for catalysis of the exonuclease reaction (Figure 12*b*) (18). Since the stereochemistry of the reaction implies an associative in-line displacement (126), the attacking nucleophile must approach the phosphodiester bond from the direction of Tyr497. At 2.6 Å resolution, there is electron density consistent with an appropriately posi-



tioned water molecule (or hydroxide ion) as the fifth ligand of the distorted tetrahedral coordination to metal A. Evidence from pH-dependence studies suggests that metal A, and not one of the protein side chains, serves to deprotonate the water molecule and facilitate nucleophilic attack (123). In addition to promoting formation of the nucleophile, the two metal ions probably accelerate the reaction by stabilizing the developing negative charge on the pentacovalent transition state or intermediate and by facilitating departure of the leaving group. Two-metal-ion catalysis of the type proposed for the 3′-5′ exonuclease reaction may be a recurrent theme in phosphoryl transfer reactions. In addition to the likely parallels with the phosphoryl transfer reactions catalyzed by alkaline phosphatase (127), ribonuclease H (38, 39), and by ribozymes (18, 53, 128, 129), one cannot ignore the possibility (discussed above) that a similar catalytic mechanism may be used at the polymerase active site (18).

The mutagenesis results suggest that active-site residues other than the

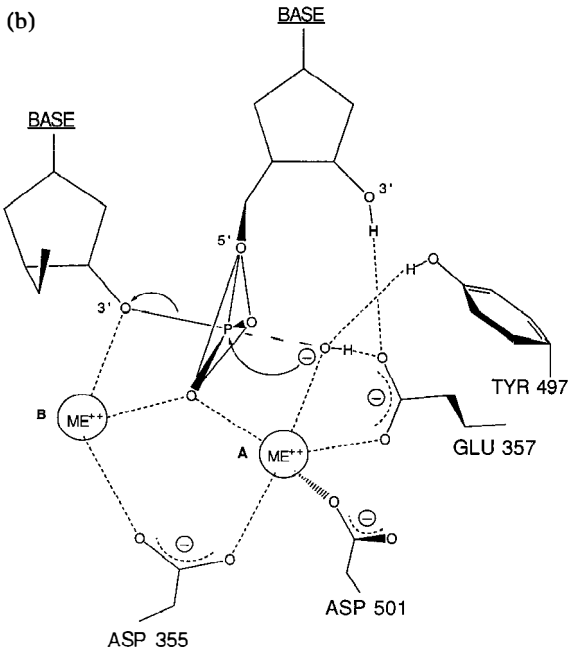


Figure 12 (*a*) Structure of the 3′-5′ exonuclease active site containing a bound dinucleotide. The catalytically essential metal ions, A and B, are shown as large black balls. Other atoms are represented as smaller balls, with phosphorus black, and carbon, oxygen, and nitrogen represented by increasingly darker shades of gray. Water molecules are shown as smaller gray spheres. Reproduced from Ref. 18. (*b*) The proposed transition state for the 3′-5′ exonuclease reaction. The mechanism is thought to involve catalysis mediated by the two bound divalent metal ions. Metal ion A facilitates the formation of the attacking hydroxide ion, whose lone-pair electrons are oriented towards the phosphorus by interactions with metal A, Tyr497, and Glu357. Metal ion B stabilizes the geometry and charge of the pentacovalent transition state and facilitates the departure of the 3′ hydroxyl group. Reproduced from Ref. 18.

aspartates at 355, 424, and 501 play an important but secondary role in the exonuclease reaction (123). From the structural data, a probable role for these side chains is that of holding the DNA substrate and the attacking nucleophile in the correct orientation for efficient catalysis. In the complex with single-stranded DNA, Leu361 and Phe473 anchor the 3′-terminal two residues by stacking with the bases; Glu357 is hydrogen-bonded to the 3′-hydroxyl as well as binding metal ion A, while Tyr497 and Glu357 may serve to orient the attacking nucleophile (18). A combination of polar interactions with the sugar-phosphate backbone and hydrophobic interactions with the nucleotide bases defines a binding pocket capable of binding 3–4

nucleotides of single-stranded DNA in a sequence-independent manner (18, 53). This size is in good agreement with biochemical data indicating that the 3′-5′ exonuclease of Klenow fragment requires the melting of 4 or 5 terminal basepairs (61), but raises questions about the energetics of such extensive melting. It has been suggested (19) that the bent configuration modeled for duplex DNA bound to the polymerase site (Figure 7) may serve to destabilize the duplex terminus and facilitate the melting necessary for the exonuclease reaction. Direct measurement by time-resolved fluorescence spectroscopy indicates that the energetics of melting is indeed in an appropriate range, since a perfectly basepaired duplex DNA molecule partitions about 7:1 in favor of the polymerase site (60).

Protein sequence alignments suggest strongly that a very similar 3′-5′ exonuclease active site will be found in all DNA polymerases that have an editing function, and preliminary structural data for the N-terminal 45-kDa domain of T4 DNA polymerase confirm this prediction (J Wang, P Yu, WH Konigsberg, TA Steitz, unpublished). The important active-site residues identified in the Klenow fragment studies are conserved in all proofreading polymerases, while the rest of the protein scaffold within this region is very divergent. Although some homologies had previously been noted, Bernad et al (130) were the first to recognize that three sequence motifs (Exo I, II, and III) were present in the majority of DNA polymerase sequences. More recently, several groups have arrived independently at essentially identical adjustments to the alignment proposed by Bernad et al (6, 104, 131, 132). The three Exo sequence motifs parallel almost exactly the active-site residues described above. The Exo I motif contains the core sequence DXE, where the two acidic residues correspond to Asp355 and Glu357 of Klenow fragment. Exo II has the sequence $NX_{2-3}(F/Y)D$; in Klenow fragment the last residue is Asp424, and the motif is immediately preceded by Gln419, whose side chain interacts with the penultimate phosphodiester bond of the DNA substrate (18, 53). The Exo III motif has the sequence YX_3D, containing the active-site residues Tyr497 and Asp501 in Klenow fragment. There is no obvious conservation of the active-site residues Leu361 (potentially part of the Exo I region) and Phe473, although in the latter case several candidate aromatic residues could be proposed in the sequences between Exo II and Exo III. Perhaps the problem of anchoring the terminal base has been solved in different ways by different polymerases.

In the absence of structural data, the validity of the above sequence alignments in predicting active-site residues has been borne out by site-directed mutagenesis studies; the majority of these (exemplified by references 98, 130, 131, 133–135) involved mutation of one or both of the invariant carboxylate residues in the Exo I motif, although a few studies have addressed the importance of residues in Exo II and/or Exo III (130, 133–137).

Despite the overall similarity implied by the conservation of active-site residues, individual polymerases have very different rates for the 3′-5′ exonuclease reaction (e.g. Refs. 99, 123, 138). This is true even when comparing the rates of degradation of single-stranded DNA and thus avoiding complications due to the need for melting at a duplex terminus. Since the available kinetic data do not establish with certainty that the observed rate in every case indicates the rate of chemical catalysis, it remains to be seen whether the differences in rate reflect structural differences at the exonuclease active site.

In Klenow fragment, the polymerase and exonuclease active sites are located about 30 Å apart on separate structural domains. Moreover, in Klenow fragment and a number of other polymerases, point mutations at one active site usually have a negligible effect on the other (23, 98, 107, 108, 123, 130, 137, 139), implying that this arrangement of separate and independent active sites is widespread. Exceptions, such as the DNA polymerase from herpes simplex virus, where the two activities cannot be cleanly separated by mutation (140), are rare, and the underlying structural causes are unclear at present. Although the two catalytic sites usually behave independently of one another, the available evidence suggests that DNA-binding functions are shared between polymerase and exonuclease domains. Even in *E. coli* DNA polymerase III, where the ϵ subunit is independently active as an exonuclease (121), the activity of ϵ is enhanced, particularly on duplex DNA, by assembly with the polymerase subunit, α (141). In Klenow fragment, the structural data place duplex DNA in a cleft between polymerase and exonuclease domains (19); moreover, the polymerase domain contributes the contact residue His660, which interacts with single-stranded DNA bound at the exonuclease site (18, 53). Experiments with the separate domains of Klenow fragment are consistent with idea of a mutual reliance of one domain on the other for DNA-binding functions (although other explanations have not been ruled out). Thus, removal of the exonuclease domain weakens duplex DNA binding to the polymerase (22, 142), while removal of the polymerase domain leaves an exonuclease domain with no detectable enzymatic or single-stranded DNA-binding activity (124).

Despite the physical separation of polymerase and 3′-5′ exonuclease active sites, these two activities must cooperate functionally to proofread polymerase errors. The interplay between the single-stranded DNA-binding properties of the exonuclease region and the duplex DNA binding associated with the polymerase site provides a straightforward structural rationale for the editing process (53, 143). Thus, editing by the 3′-5′ exonuclease is determined by the single-stranded character of the primer terminus, which favors its binding to the associated single-stranded DNA-binding site. Detailed studies of the exonuclease activity of T4 and T7 polymerases demonstrate how an increas-

ing number of terminal mismatches favor the exonuclease reaction both by increasing the proportion of termini in a "frayed" configuration and by accelerating the melting process that converts a duplex molecule into a substrate for the exonuclease (15, 99, 138). At the same time, the polymerase active site plays a role in enhancing the effectiveness of the editing process. Not only is a mispaired terminus preferred over a correctly paired terminus as a substrate for the exonuclease, it is also a very poor substrate for further rounds of dNTP addition at the polymerase site (15, 144, 145). This delay increases the time window for editing of a mismatch and allows even the slow 3′-5′ exonuclease of Klenow fragment to serve as an effective proofreader. Because the efficiency of editing is determined by a physical property (notably, the tendency to melt or fray) of the DNA, effective editing of a polymerase error can be achieved even if the DNA product travels from polymerase to exonuclease site via dissociation. Consistent with this prediction, both intermolecular and intramolecular transfers between active sites have been observed (e.g. 99, 138, 146, 147), as a natural consequence of the balance between the rates of DNA dissociation and of the reactions under consideration (which, in turn, may be related to the in vivo function of a particular enzyme).

Ribonuclease H

The ribonuclease H activity associated with retroviral reverse transcriptases acts as an endonuclease (148, 149) to degrade the viral RNA template during (−) strand DNA synthesis. Additionally, though beyond the scope of this review, RNase H participates in a series of strand-transfer reactions that are necessary for the complete cycle of retroviral replication (150–152). The structure of the RNase H domain of HIV-1 reverse transcriptase has been solved at 2.4 Å resolution from the isolated domain (38) and at 2.9 Å resolution in the context of the heterodimer (10, 37). The isolated RNase H domain, from Tyr427 to the C-terminus, has exactly the same structure as in the heterodimer except for a small loop at the active site containing His539, which is disordered only in the isolated domain (10). A comparison (38) of the structure of the HIV-1 RNase H domain with that of *E. coli* RNase H (39, 40) shows that they are very similar in terms of the overall fold and location of secondary structure elements, as well as the position of active-site residues (see below). An important difference is that the *E. coli* enzyme has an insertion of some 15 residues, 6 of which are basic, which form an additional α helix at the surface of the protein. It has been suggested (40, 153) that this basic protrusion on the *E. coli* enzyme forms part of the nucleic acid–binding site, contacting the RNA template strand about 10 residues 5′ to the catalytic site. This location is similar to the contacts proposed between the HIV-1 reverse transcriptase heterodimer and

a DNA-RNA duplex (10), although in reverse transcriptase the contacts are provided by residues from the polymerase domain. The structural comparisons therefore suggest that the polymerase domain of HIV-1 reverse transcriptase provides an important part of the nucleic acid–binding site for the RNase H reaction and also stabilizes the active-site loop containing His539. Either or both of these functions could be the reason why the isolated p15 RNase H domain is only weakly active (27) unless reconstituted with p51 (154). Interestingly, the RNase H region of Moloney murine leukemia virus contains sequences similar to the additional basic region in *E. coli* RNase H and, in contrast to the situation in HIV-1 reverse transcriptase, the murine retroviral RNase H is active as an isolated domain (155).

The RNase H domain consists of a five-stranded β-sheet flanked on each side by α-helices. On this β-sheet are four carboxylate residues (three Asp, one Glu) that are conserved among all RNase H sequences (156). Mutagenesis studies of HIV-1 reverse transcriptase and *E. coli* RNase H indicate that at least three of these carboxylates are crucial for RNase H activity (157–159). In the structure of the HIV-1 RNase H domain, the conserved carboxylates are seen to bind two Mn^{2+} ions spaced 4 Å apart, inviting comparison with the 3′-5′ exonuclease active site of Klenow fragment. In fact, there is substantial structural similarity in the active-site regions of these two nucleases, although the overall topology of folding implies that they do not share a common evolutionary ancestor. As shown in Figure 13, it is possible to superimpose three of the conserved carboxylates, the two metal ions, and a substantial region of the surrounding β-strand scaffold in the two structures (37). The superposition implies a correspondence between Asp443, Glu478, and Asp549 of HIV-1 reverse transcriptase and, respectively, Asp355, Asp424, and Asp501 of Klenow fragment. The suggested equivalence between Asp549 (reverse transcriptase) and Asp501 (Klenow fragment) may explain why the homologous residue in *E. coli* RNase H, Asp134, seems to be dispensable for enzymatic activity even though this residue is conserved in all RNase H sequences (156). The analysis of the role of Asp134 in *E. coli* RNase H made use of an asparagine substitution that had no effect on the enzymatic reaction (159); the same substitution at position 501 of Klenow fragment was likewise without effect, although alanine or glutamate substitutions caused dramatic decreases in exonuclease activity (123). Thus, for RNase H, just as for the Klenow fragment 3′-5′ exonuclease, the requirement at this position may be merely for a metal ligand of the correct stereochemistry.

As in the Klenow fragment 3′-5′ exonuclease, mutational analysis of the RNase H active site suggests that the most important side chains are the carboxylates, implying a central role in catalysis for the metal ions to which

they serve as ligands. In addition to the carboxylates, there are three residues (Ser, His, and Asn) that are conserved in all retroviral and bacterial RNase H sequences (156). Although these side chains are located at the RNase H active site, mutations at these positions have much less effect than mutations in the carboxylates, and it has been suggested that these residues may play some role in substrate binding (159–161). Indeed, model building of an RNA-DNA substrate suggests that His539 may interact with the phosphate backbone (J Jäger, TA Steitz, unpublished). The apparent structural similarity between the 3′-5′ exonuclease and RNase H active sites, together with the necessity for a similar reaction pathway (both cleave so as to leave a 5′ phosphate product), suggests that RNase H also may use a two-metal-ion catalytic mechanism (38, 39), as described for the 3′-5′ exonuclease (Figure 12*b*). While the observation of two bound metal ions, with the appropriate spacing, in the HIV-1 RNase H domain structure (38) provides good evidence for such a mechanism, the presence of only a single magnesium ion in the *E. coli* RNase H structure (40) has been taken as evidence for the related reaction mechanism in which the necessary deprotonation of the attacking water molecule is facilitated by one of the carboxylates acting as a general base (162). However, it is also possible that the putative second metal site in the *E. coli* enzyme could be obstructed by basic residues from a neighboring molecule (39, 40). Alternatively, as is the case with the 3′-5′ exonuclease, binding of substrate or product may be required for occupancy of the second metal site by Mg^{2+} (37). Only Mn^{2+} binds to both metal sites of the 3′-5′ exonuclease in the absence of a bound phosphate at the cleavage site (18).

The recently determined NMR structure of an RNA-DNA hybrid (163) has provided a possible mechanism whereby RNase H may discriminate among RNA-RNA, RNA-DNA, and DNA-DNA duplexes in its cleavage specificity. The minor groove of the A-form RNA-DNA duplex is narrower than that of RNA-RNA, though wider than the B-form DNA-DNA minor groove. Fedoroff et al (163) suggest that only this narrower A-form phosphate backbone spacing will fit into the RNase H active site.

The structure of HIV-1 reverse transcriptase shows polymerase and RNase H activities located on separate structural domains, with about 60 Å distance between the two active sites. Although inhibition studies (164) and the properties of many point mutants (e.g. 44, 157) suggest that the two active sites are functionally distinct, there is clearly a substantial degree of communication between the domains. Thus, attempts to separate the two activities genetically using deletion and insertion mutations were largely unsuccessful (165–167). By contrast, the polymerase and RNase H activities of the reverse transcriptase of Moloney murine leukemia virus seem much more independent since they can be cleanly separated both by subcloning

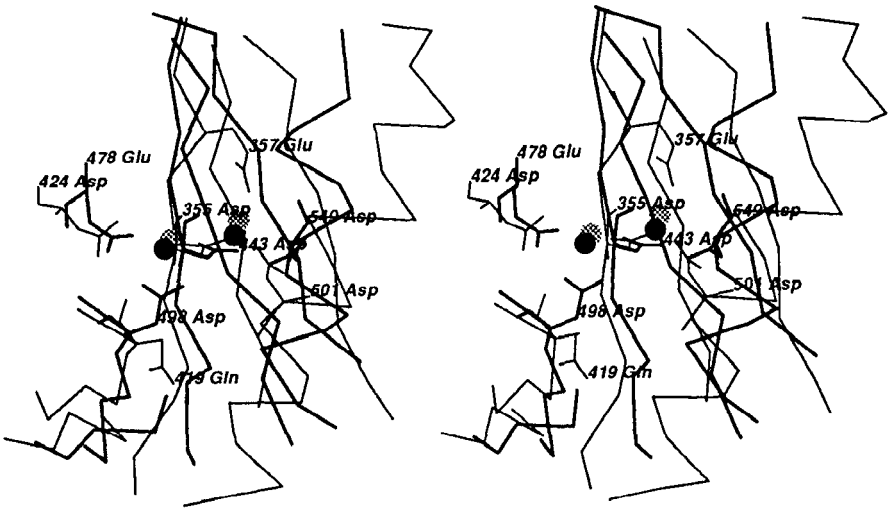


Figure 13 Stereo representation of a least-squares overlap of the 3′-5′ exonuclease domain of Klenow fragment (thin lines) with the RNase H domain of HIV-1 reverse transcriptase (thick lines). The superposition was carried out by first overlapping the C_γ atoms of the three pairs of matching carboxylic acid residues and the two pairs of metal ions, and was then extended to include 37 C_α atoms from each structure in the three-stranded β-sheet and a short section of the α-helix. The RNase H metal ions are shown as black circles and those of the 3′-5′ exonuclease as lightly shaded circles. Reproduced from Ref. 37.

and by linker-insertion mutagenesis (155). As already discussed, some of the differences between these two reverse transcriptases may be due to a greater requirement in the HIV-1 RNase H reaction for substrate contacts provided by the polymerase domain. Additionally, a large area of the reverse transcriptase molecule is involved in the dimerization interface, and mutations that affect these interactions would be likely to influence reactions at both active sites. Consistent with this idea, recent experiments with the reverse transcriptase of Moloney murine leukemia virus have shown that, in spite of the apparent independence of the two activities in this system, some RNase H–defective mutations exert subtle effects on the polymerase reaction by interfering with dimerization (34).

The mechanism of retroviral reverse transcription requires that the two separate active sites cooperate so as to achieve degradation of the original viral RNA template as the (−) strand DNA is synthesized, and raises questions about the temporal and spatial relationship of the two activities. As noted in an earlier section, the 19-nucleotide distance between the RNase H cleavage position and the primer terminus is in good agreement with model-building of an A-form DNA-RNA hybrid onto the reverse transcriptase structure (Figure 8*a*). Moreover, the demonstration that both reac-

tions can occur during a single processive cycle of DNA synthesis (64) is consistent with this model, which allows simultaneous placement of the DNA primer terminus at the polymerase active site and the RNA template phosphodiester backbone at the RNase H active site. Although it is clear that both activities can take place during a single polymerase–nucleic acid encounter (64), the available evidence argues against a tight or obligatory coupling between them (63, 168). More likely, the sequence of reactions that takes place will be determined by the relative rates of the polymerase and RNase H reactions in relation to dissociation of the complex (169). An elegant experiment, showing complementation between a polymerase-defective and an RNase H–defective reverse transcriptase, argues against any strict requirement that the two activities be physically coupled, even in vivo (170).

5′-3′ Exonuclease

Most of the Pol I family of enzymes have an associated 5′-3′ exonuclease activity specific for double-stranded DNA. In the bacterial DNA polymerase I enzymes, this activity is part of the same polypeptide chain as the polymerase. By contrast, bacteriophages T4, T5, and T7 encode essentially the same enzymatic activity in a separate polypeptide (171–173). In several cases, these enzymes have been shown to have RNase H activity with the same 5′-3′ polarity (171, 174–176), an observation that fits well with the in vivo requirement for removal of RNA primers during lagging-strand DNA replication. It seems likely that some eukaryotic and viral polymerases may have similar 5′-3′ exonuclease activities, either covalently attached to the polymerase (177) or as an associated subunit (178).

The conventionally used term "5′-3′ exonuclease" is an inaccurate description of the activity of these enzymes. As was originally shown for *E. coli* DNA polymerase I (179), and more recently for several thermophilic DNA polymerases (175), the preferred substrate is a displaced 5′ end of the type that could be generated by strand-displacement synthesis. Cleavage takes place at or close to the junction between the single-stranded and duplex DNA, though it seems likely that the enzyme requires a free 5′ end for access to the substrate (175). The term "structure-specific endonuclease" is therefore a better description of this enzymatic activity. Specificity for a particular substrate structure seems to be intrinsic to the nuclease and not to require presentation of the substrate by an associated polymerase, since both the isolated exonuclease domain of *E. coli* DNA polymerase I and the T7 exonuclease showed similar rates and positions of cleavage on model substrates as did a polymerase-associated 5′-3′ exonuclease (175, 179). A strong inference from both these model studies is that the well-known stimulation of the 5′-3′ exonuclease by polymerase action may be due to

production of the preferred substrate and need not imply any physical or temporal coupling between the two activities.

The lack of structural data for any 5′-3′ exonuclease precludes any definitive conclusions on structure-function relationships, but protein sequence alignments and mutational studies provide some tantalizing clues. Characterization of 5′-3′ exonuclease-proficient amber fragments of *E. coli* DNA polymerase I indicates that the exonuclease active site must be within the first 297 residues of the protein (180). Alignment of the protein sequences of six 5′-3′ exonuclease regions from bacterial polymerases and four bacteriophage 5′-3′ exonucleases indicates six highly conserved sequence motifs containing 14 invariant amino acids (181). Nine of these invariant residues are acidic, inviting the speculation that the 5′-3′ exonuclease, like the other enzymatic activities so far described, may require the coordination of divalent metal ions at the active site. A conservative mutation, D13N, at one of the invariant residues in *E. coli* DNA polymerase I results in a dramatic decrease in 5′-3′ exonuclease activity (CM Joyce, unpublished data). Intriguingly, the sequence surrounding Asp13 (ILVDGSSY) is quite similar to the sequence in *E. coli* RNase H that contains the active-site residue Asp10 (IFTDGSCL). Inspection of the 5′-3′ exonuclease sequence alignment does not, however, suggest any obvious counterparts to the other carboxylates, Glu48 and Asp70, that have been shown to be important for RNase H activity (159).

An entirely different structural question concerns the way in which the 5′-3′ exonuclease domain fits together with the Klenow fragment structure; this question should be resolved by the ongoing structural studies on *Taq* polymerase (Y-S Kim, TA Steitz, unpublished data). The polymerase active site, at the base of a cleft, and the 5′-3′ exonuclease active site, on an independent structural domain, must necessarily be separated in space, and yet the two must work together in vivo so as to leave a ligatable nick. The way in which this cooperation is achieved poses some intriguing mechanistic questions.

CONCLUDING REMARKS

The available evidence leads us to predict that the common polymerase active-site architecture observed in the three published polymerase structures will be widespread, perhaps universal, within the polymerase superfamily. The major global differences between individual polymerases can be attributed to additional functions that are assembled onto the polymerase core, frequently in a modular fashion. Obvious examples are the various nuclease modules (described above) that are present in particular polymerase families. Additionally, some polymerases contain subdomains responsible for the

protein-protein interactions that allow assembly of large multisubunit replicative complexes, whose overall molecular properties appear very different from those of the simple polymerase modules that are their functional core. Finally, there is suggestive, though by no means conclusive, evidence that the phosphoryl transfer reactions catalyzed by polymerases and their associated nucleases may share a common mechanism of catalysis mediated by divalent metal ions.

ACKNOWLEDGMENTS

We are grateful to the many coworkers who have shared our investigations of polymerase structure and function over the years. We are particularly grateful to Joe Jäger and Steve Smerdon for invaluable help with the figures, to Nigel Grindley for comments on the manuscript, and to Kate Tatham for her patience in preparing the manuscript. We also thank the large number of colleagues who sent us preprints of their work. Polymerase research in the authors' laboratories is supported by NIH grants GM-28550 and GM-39546, ACS grant BE-52, and HHMI funding.

This paper is dedicated to the memory of Hatch Echols.

Literature Cited

1. Kornberg A, Baker TA. 1992. *DNA Replication*. San Francisco: Freeman. 931 pp. 2nd ed.
2. Bryan S, Chen H, Sun Y, Moses RE. 1988. *Biochim. Biophys. Acta* 951:249–54
3. Witkin EM, Roegner-Maniscalco V. 1992. *J. Bacteriol.* 174:4166–68
4. Sweasy JB, Loeb LA. 1992. *J. Biol. Chem.* 267:1407–10
5. Joyce CM, Sun XC, Grindley NDF. 1992 *J. Biol. Chem.* 267: 24485–500
6. Braithwaite DK, Ito J. 1993. *Nucleic Acids Res.* 21:787–802
7. Delarue M, Poch O, Tordo N, Moras D, Argos P. 1990. *Protein Eng.* 3:461–67
8. Richetti M, Buc H. 1993. *EMBO J.* 12:387–96
9. Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA. 1985. *Nature* 313:762–66
10. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. 1992. *Science* 256:1783–90
11. Jacobo-Molina A, Ding J, Nanni RG, Clark AD Jr, Lu X, et al. 1993. *Proc. Natl. Acad. Sci. USA* 90:6320–24
12. Marians KJ. 1992. *Annu. Rev. Biochem.* 61:673–719
13. Wang TS-F. 1991. *Annu. Rev. Biochem.* 60:513–52
14. Echols H, Goodman MF. 1991. *Annu. Rev. Biochem.* 60:477–511
15. Johnson KA. 1993. *Annu. Rev. Biochem.* 62:685–713
16. Sousa R, Chung YJ, Rose JP, Wang B-C. 1993. *Nature* 364:593–99
17. Crouch RJ. 1990. *New Biol.* 2:771–77
18. Beese LS, Steitz TA. 1991. *EMBO J.* 10:25–33
19. Beese LS, Derbyshire V, Steitz TA. 1993. *Science* 260:352–55
20. Yadav PNS, Yadav JS, Modak MJ. 1992. *Biochemistry* 31:2879–86
21. Beese LS, Friedman JM, Steitz TA. 1993. *Biochemistry* 32:14095–101

22. Freemont PS, Ollis DL, Steitz TA, Joyce CM. 1986. *Proteins* 1:66–73
23. Derbyshire V, Freemont PS, Sanderson MR, Beese L, Friedman JM, et al. 1988. *Science* 240:199–201
24. Goff SP. 1990. *J. Acquired Immune Defic. Syndr.* 3:817–31
25. Weiss R, Teich N, Varmus HE, Coffin J, eds. 1985. *Molecular Biology of Tumor Viruses: RNA Tumor Viruses*, Parts 1 & 2. Cold Spring Harbor, NY: Cold Spring Harbor Lab. 2nd ed.
26. Johnson MS, McClure MA, Feng D-F, Gray J, Doolittle RF. 1986. *Proc. Natl. Acad. Sci. USA* 83:7648–52
27. Hansen J, Schulze T, Mellert W, Moelling K. 1988. *EMBO J.* 7:239–43
28. Di Marzo Veronese F, Copeland TD, DeVico AL, Rahman R, Oroszlan S, et al. 1986. *Science* 231:1289–91
29. Lightfoote MM, Coligan JE, Folks TM, Fauci AS, Martin MA, Venkatesan S. 1986. *J. Virol.* 60:771–75
30. LeGrice SFJ, Naas T, Wohlgensinger B, Schatz O. 1991. *EMBO J.* 10:3905–11
31. Hostomsky Z, Hostomska Z, Fu T-B, Taylor J. 1992. *J. Virol.* 66:3179–82
32. Wu JC, Warren TC, Adams J, Proudfoot J, Skiles J, et al. 1991. *Biochemistry* 30:2022–26
33. Barat C, Lullien V, Schatz O, Keith G, Nugeyre MT, et al. 1989. *EMBO J.* 8:3279–88
34. Telesnitsky A, Goff SP. 1993. *Proc. Natl. Acad. Sci. USA* 90:1276–80
35. Lowe DM, Aitken A, Bradley C, Darby GK, Larder BA, et al. 1988. *Biochemistry* 27:8884–89
36. Lloyd LF, Brick P, Mei-Zhen L, Chayen NE, Blow DM. 1991. *J. Mol. Biol.* 217:19–22
37. Steitz TA, Smerdon S, Jäger J, Wang J, Kohlstaedt LA, et al. 1993. *Cold Spring Harbor Symp. Quant. Biol.* 58:In press
38. Davies JF, Hostomska Z, Homstomsky Z, Jordan SR, Matthews DA. 1991. *Science* 252:88–95
39. Yang W, Hendrickson WA, Crouch RJ, Satow Y. 1990. *Science* 249:1398–405
40. Katayanagi K, Miyagawa M, Matsushima M, Ishikawa M, Kanaya S, et al. 1992. *J. Mol. Biol.* 223: 1029–52
41. Steitz TA. 1993. *Curr. Opin. Struct. Biol.* 3:31–38
42. Larder BA, Purifoy DJM, Powell KL, Darby G. 1987. *Nature* 327:716–17
43. Larder BA, Kemp SD, Purifoy DJM. 1989. *Proc. Natl. Acad. Sci. USA* 86:4803–7
44. Boyer PL, Ferris AL, Hughes SH. 1992. *J. Virol.* 66:1031–39
45. Polesky AH, Steitz TA, Grindley NDF, Joyce CM. 1990. *J. Biol. Chem.* 265: 14579–91
46. Polesky AH, Dahlberg ME, Benkovic SJ, Grindley NDF, Joyce CM. 1992. *J. Biol. Chem.* 267:8417–28
47. Arnold E, Jacobo-Molina A, Nanni RG, Williams RL, Lu X, et al. 1992. *Nature* 357:85–89
48. Bonner G, Patra D, Lafer EM, Sousa R. 1992. *EMBO J.* 11:3767–75
49. Osumi-Davis PA, de Aguilera MC, Woody RW, Woody AY. 1992. *J. Mol. Biol.* 226:37–45
50. Ollis DL, Kline C, Steitz TA. 1985. *Nature* 313:818–19
51. Warwicker J, Ollis D, Richards FM, Steitz TA. 1985. *J. Mol. Biol.* 186: 645–49
52. Joyce CM, Fujii DM, Laks HS, Hughes CM, Grindley NDF. 1985. *J. Mol. Biol.* 186:283–93
53. Freemont PS, Friedman JM, Beese LS, Sanderson MR, Steitz TA. 1988. *Proc. Natl. Acad. Sci. USA* 85:8924–28
54. Basu A, Williams KR, Modak MJ. 1987. *J. Biol. Chem.* 262:9601–07
55. Basu S, Basu A, Modak MJ. 1988. *Biochemistry* 27:6710–16
56. Mohan PM, Basu A, Basu S, Abraham KI, Modak MJ. 1988. *Biochemistry* 27:226–33
57. Catalano CE, Allen DJ, Benkovic SJ. 1990. *Biochemistry* 29:3612–21
58. Joyce CM, Steitz TA. 1987. *Trends Biochem. Sci.* 12:288–92
59. Allen DJ, Darke PL, Benkovic SJ. 1989. *Biochemistry* 28:4601–7
60. Guest CR, Hochstrasser RA, Dupuy CG, Allen DJ, Benkovic SJ, Millar DP. 1991. *Biochemistry* 30:8759–70
61. Cowart M, Gibson KJ, Allen DJ, Benkovic SJ. 1989. *Biochemistry* 28: 1975–83
62. Furfine ES, Reardon JE. 1991. *J. Biol. Chem.* 266:406–12
63. Kati WA, Johnson KA, Jerva LF, Anderson KS. 1992. *J. Biol. Chem.* 267:25988–97
64. Gopalakrishnan V, Peliska JA, Benkovic SJ. 1992. *Proc. Natl. Acad. Sci. USA* 89:10763–67
65. Basu A, Ahluwalia KK, Basu S, Modak MJ. 1992. *Biochemistry* 31: 616–23
66. Sobol RW, Suhadolnik RJ, Kumar A, Lee BJ, Hatfield DL, Wilson SH. 1991. *Biochemistry* 30:10623–31
67. Sardana VV, Emini EA, Gotlib L, Graham DJ, Lineberger DW, et al. 1992. *J. Biol. Chem.* 267:17526–30

68. Merluzzi VJ, Hargrave KD, Labadia M, Grozinger K, Skoog M, et al. 1990. *Science* 250:1411–13
69. Kohlstaedt LA, Steitz TA. 1992. *Proc. Natl. Acad. Sci. USA* 89:9652–56
70. Barat C, LeGrice SFJ, Darlix J-L. 1991. *Nucleic Acids Res.* 19:751–57
71. Leis J, Aiyar A, Cobrinik D. 1993. In *Reverse Transcriptase,* ed. AM Skalka, SP Goff, pp. 33–47. Cold Spring Harbor, NY: Cold Spring Harbor Lab.
72. Bryant FR, Johnson KA, Benkovic SJ. 1983. *Biochemistry* 22:3537–46
73. Rush J, Konigsberg WH. 1990. *J. Biol. Chem.* 265:4821–27
74. Pandey VN, Williams KR, Stone KL, Modak MJ. 1987. *Biochemistry* 26: 7744–48
75. Pandey VN, Kaushik NA, Pradhan DS, Modak MJ. 1990. *J. Biol. Chem.* 265:3679–84
76. Basu A, Modak MJ. 1987. *Biochemistry* 26:1704–9
77. Pandey VN, Kaushik N, Sanzgiri RP, Patil MS, Modak MJ, Barik S. 1993. *Eur. J. Biochem.* 214:59–65
78. Cheng N, Merrill BM, Painter GR, Frick LW, Furman PA. 1993. *Biochemistry* 32:7630–34
79. Wu J, Amandoron E, Li X, Wainberg MA, Parniak MA. 1993. *J. Biol. Chem.* 268:9980–85
80. Basu A, Tirumalai RS, Modak MJ. 1989. *J. Biol. Chem.* 264:8746–52
81. Martin JL, Wilson JE, Furfine ES, Hopkins SE, Furman PA. 1993. *J. Biol. Chem.* 268:2565–70
82. Mitchell LLW, Cooperman BS. 1992. *Biochemistry* 31:7707–13
83. Dong Q, Copeland WC, Wang TS-F. 1993. *J. Biol. Chem.* 268:24163–74
84. Blasco MA, Lázaro JM, Bernad A, Blanco L, Salas M. 1992. *J. Biol. Chem.* 267:19427–34
85. Blasco MA, Lázaro JM, Blanco L, Salas M. 1993. *J. Biol. Chem.* 268: 16763–70
86. Larder BA, Kemp SD. 1989. *Science* 246:1155–58
87. Fitzgibbon JE, Howell RM, Haberzettl CA, Sperber SJ, Gocke DJ, Dubin DT. 1992. *Antimicrob. Agents Chemother.* 36:153–57
88. St. Clair MH, Martin JL, Tudor-Williams G, Bach MC, Vavro CL, et al. 1991. *Science* 253:1557–59
89. Kellam P, Boucher CAB, Larder BA. 1992. *Proc. Natl. Acad. Sci. USA* 89:1934–38
90. Carroll SS, Cowart M, Benkovic SJ. 1991. *Biochemistry* 30:804–13
91. Burgers PMJ, Eckstein F. 1979. *J. Biol. Chem.* 254:6889–93
92. Brody RS, Frey PA. 1981. *Biochemistry* 20:1245–52
93. Romaniuk PJ, Eckstein F. 1982. *J. Biol. Chem.* 257:7684–88
94. Bartlett PA, Eckstein F. 1982. *J. Biol. Chem.* 257:8879–84
95. Brody RS, Adler S, Modrich P, Stec WJ, Leznikowski ZL, Frey PA. 1982. *Biochemistry* 21:2570–72
96. Hopkins S, Furman PA, Painter GR. 1989. *Biochem. Biophys. Res. Commun.* 163:106–10
97. Kuchta RD, Mizrahi V, Benkovic PA, Johnson KA, Benkovic SJ. 1987. *Biochemistry* 26:8410–17
98. Patel SS, Wong I, Johnson KA. 1991. *Biochemistry* 30:511–25
99. Capson TL, Peliska JA, Kaboord BF, Frey MW, Lively C, et al. 1992. *Biochemistry* 31:10984–94
100. Reardon JE. 1992. *Biochemistry* 31: 4473–79
101. Reardon JE. 1993. *J. Biol. Chem.* 268:8743–51
102. Dahlberg ME, Benkovic SJ. 1991. *Biochemistry* 30:4835–43
103. Eger BT, Benkovic SJ. 1992. *Biochemistry* 31:9227–36
104. Heringa J, Argos P. 1994. *Evolutionary Biology of Viruses,* ed. SS Morse. New York: Raven. In press
105. Boyer PL, Ferris AL, Hughes SH. 1992. *J. Virol.* 66:7533–37
106. Copeland WC, Wang TS-F. 1993. *J. Biol. Chem.* 268:11028–40
107. Bernad A, Lázaro JM, Salas M, Blanco L. 1990. *Proc. Natl. Acad. Sci. USA* 87:4610–14
108. Jung G, Leavitt MC, Schultz M, Ito J. 1990. *Biochem. Biophys. Res. Commun.* 170:1294–300
109. Joung I, Horwitz MS, Engler JA. 1991. *Virology* 184:235–41
110. Date T, Yamamoto S, Tanihara K, Nishimoto Y, Matsukage A. 1991. *Biochemistry* 30:5286–92
111. Sankar S, Porter AG. 1992. *J. Biol. Chem.* 267:10168–76
112. Wakefield JK, Jablonski SA, Morrow CD. 1992. *J. Virol.* 66:6806–12
113. Blasco MA, Lázaro JM, Blanco L, Salas M. 1993. *J. Biol. Chem.* 268: 24106–13
114. Joyce CM. 1991. *Curr. Opin. Struct. Biol.* 1:123–29
115. Dong Q, Copeland WC, Wang TS-F. 1993. *J. Biol. Chem.* 268:24175–82
116. Reha-Krantz LJ, Nonay RL, Stocki S. 1993. *J. Virol.* 67:60–66
117. Copeland WC, Lam NK, Wang TS-F. 1993. *J. Biol. Chem.* 268:11041–49

118. Reha-Krantz LJ. 1988. *J. Mol. Biol.* 202:711–24
119. Hall JD. 1988. *Trends Genet.* 4:42–46
120. Bebenek K, Abbotts J, Roberts JD, Wilson SH, Kunkel TA. 1989. *J. Biol. Chem.* 264:16948–56
121. Scheuermann RH, Echols H. 1984. *Proc. Natl. Acad. Sci. USA* 81:7747–51
122. Brutlag D, Kornberg A. 1972. *J. Biol. Chem.* 247:241–48
123. Derbyshire V, Grindley NDF, Joyce CM. 1991. *EMBO J.* 10:17–24
124. Derbyshire V. 1990. *Studies of the 3′ to 5′ exonuclease of DNA polymerase I of Escherichia coli*. PhD thesis. Yale Univ., New Haven. 111 pp.
125. Han H, Rifkind JM, Mildvan AS. 1991. *Biochemistry* 30:11104–8
126. Gupta AP, Benkovic SJ. 1984. *Biochemistry* 23:5874–81
127. Kim EE, Wykoff HW. 1991. *J. Mol. Biol.* 218:449–64
128. Piccirilli JA, Vyle JS, Caruthers MH, Cech TR. 1993. *Nature* 361:85–88
129. Steitz TA, Steitz JA. 1993. *Proc. Natl. Acad. Sci. USA* 90:6498–502
130. Bernad A, Blanco L, Lázaro JM, Martin G, Salas M. 1989. *Cell* 59:219–28
131. Morrison A, Bell JB, Kunkel TA, Sugino A. 1991. *Proc. Natl. Acad. Sci. USA* 88:9473–77
132. Blanco L, Bernad A, Salas M. 1992. *Gene* 112:139–44
133. Foury F, Vanderstraeten S. 1992. *EMBO J.* 11:2717–26
134. Reha-Krantz LJ, Nonay RL. 1993. *J. Biol. Chem.* 268:27100–8
135. Simon M, Giot L, Faye G. 1991. *EMBO J.* 10:2165–70
136. Soengas MS, Esteban JA, Lázaro JM, Bernad A, Blasco MA, et al. 1992. *EMBO J.* 11:4227–37
137. Frey MW, Nossal NG, Capson TL, Benkovic SJ. 1993. *Proc. Natl. Acad. Sci. USA* 90:2579–83
138. Donlin MJ, Patel SS, Johnson KA. 1991. *Biochemistry* 30:538–46
139. Eger BT, Kuchta RD, Carroll SS, Benkovic PA, Dahlberg ME, et al. 1991. *Biochemistry* 30:1441–48
140. Gibbs JS, Weisshart K, Digard P, de Bruynkops A, Knipe DM, Coen DM. 1991. *Mol. Cell. Biol.* 11:4786–95
141. Maki H, Kornberg A. 1987. *Proc. Natl. Acad. Sci. USA* 84:4389–92
142. Derbyshire V, Astatke M, Joyce CM. 1994. *Nucleic Acids Res.* 21: 5439–48
143. Joyce CM, Friedman JM, Beese L, Freemont PS, Steitz TA. 1988. *DNA Replication and Mutagenesis*, ed. RE Moses, WC Summers, pp. 220–26. Washington, DC: Am. Soc. Microbiol.
144. Kuchta RD, Benkovic P, Benkovic SJ. 1988. *Biochemistry* 27:6716–25
145. Wong I, Patel SS, Johnson KA. 1991. *Biochemistry* 30:526–37
146. Joyce CM. 1989. *J. Biol. Chem.* 264: 10858–66
147. Reddy MK, Weitzel SE, von Hippel PH. 1992. *J. Biol. Chem.* 267:14157–66
148. Oyama F, Kikuchi R, Crouch RJ, Uchida T. 1989. *J. Biol. Chem.* 264: 18808–17
149. Krug MS, Berger SL. 1989. *Proc. Natl. Acad. Sci. USA* 86:3539–43
150. Luo GX, Taylor J. 1990. *J. Virol.* 64:4321–28
151. Champoux JJ. 1993. In *Reverse Transcriptase*, ed. AM Skalka, SP Goff, pp. 103–17. Cold Spring Harbor, NY: Cold Spring Harbor Lab.
152. Peliska JA, Benkovic SJ. 1992. *Science* 258:1112–18
153. Kanaya S, Katsuda-Nakai C, Ikehara M. 1991. *J. Biol. Chem.* 266:11621–27
154. Hostomsky Z, Hostomska Z, Hudson GO, Moomaw EW, Nodes BR. 1991. *Proc. Natl. Acad. Sci. USA* 88:1148–52
155. Tanese N, Goff SP. 1988. *Proc. Natl. Acad. Sci. USA* 85:1777–81
156. Doolittle RF, Feng D-F, Johnson MS, McClure MA. 1989. *Q. Rev. Biol.* 64:1–29
157. Mizrahi V, Usdin MT, Harington A, Dudding LR. 1990. *Nucleic Acids Res.* 18:5359–63
158. Schatz O, Cromme FV, Grüninger-Leitch F, LeGrice SFJ. 1989. *FEBS Lett.* 257:311–14
159. Kanaya S, Kohara A, Miura Y, Sekiguchi A, Iwai S, et al. 1990. *J. Biol. Chem.* 265:4615–21
160. Tisdale M, Schulze T, Larder BA, Moelling K. 1991. *J. Gen. Virol.* 72:59–66
161. Wöhrl BM, Volkmann S, Moelling K. 1991. *J. Mol. Biol.* 220:801–18
162. Nakamura H, Oda Y, Iwai S, Inoue H, Ohtsuka E, et al. 1991. *Proc. Natl. Acad. Sci. USA* 88:11535–39
163. Fedoroff OY, Salazar M, Reid BR. 1993. *J. Mol. Biol.* 233:509–23
164. Tan C-K, Zhang J, Li Z-Y, Tarpley WG, Downey KM, So AG. 1991. *Biochemistry* 30:2651–55
165. Prasad VR, Goff SP. 1989. *Proc. Natl. Acad. Sci. USA* 86:3104–08
166. Hizi A, Barber A, Hughes SH. 1989. *Virology* 170:326–29
167. Hizi A, Hughes SH, Shaharabany M. 1990. *Virology* 175:575–80
168. DeStefano JJ, Buiser RG, Mallaber

LM, Myers TW, Bambara RA, Fay PJ. 1991. *J. Biol. Chem.* 266:7423–31
169. Dudding LR, Mizrahi V. 1993. *Biochemistry* 32:6116–20
170. Telesnitsky A, Goff SP. 1993. *EMBO J.* 12:4433–38
171. Hollingsworth HC, Nossal NG. 1991. *J. Biol. Chem.* 266:1888–97
172. Sayers JR, Eckstein F. 1990. *J. Biol. Chem.* 265:18311–17
173. Kerr C, Sadowski PD. 1972. *J. Biol. Chem.* 247:311–18
174. Berkower I, Leis J, Hurwitz J. 1973. *J. Biol. Chem.* 248:5914–21
175. Lyamichev V, Brow MAD, Dahlberg JE. 1993. *Science* 260:778–83
176. Shinozaki K, Okazaki T. 1978. *Nucleic Acids Res.* 5:4245–61
177. Crute JJ, Lehman IR. 1989. *J. Biol. Chem.* 264:19266–70
178. Siegal G, Turchi JJ, Myers TW, Bambara RA. 1992. *Proc. Natl. Acad. Sci. USA* 89:9377–81
179. Lundquist RC, Olivera BM. 1982. *Cell* 31:53–60
180. Kelley WS, Joyce CM. 1983. *J. Mol. Biol.* 164:529–60
181. Gutman PD, Minton KW. 1993. *Nucleic Acids Res.* 21:4406–7
182. Poch O, Sauvaget I, Delarue M, Tordo N. 1989. *EMBO J.* 8:3867–74
183. Blanco L, Bernad A, Blasco MA, Salas M. 1991. *Gene* 100:27–38
184. Wong SW, Wahl AF, Yuan P-M, Arai N, Pearson BE, et al. 1988. *EMBO J.* 7:37–47