

## DNA Polymerases: Structural Diversity and Common Mechanisms\*

Thomas A. Steitz‡

From the Howard Hughes Medical Institute and  
Departments of Molecular Biophysics & Biochemistry  
and Chemistry, Yale University,  
New Haven, Connecticut 06520-8114

Possibly the earliest enzymatic activity to appear in evolution was that of the polynucleotide polymerases, the ability to replicate the genome accurately being a prerequisite for evolution itself. Thus, one might anticipate that the mechanism by which all polymerases work would be both simple and universal. Further, these enzymatic scribes must faithfully copy the sequences of the genome into daughter nucleic acid or the information contained within would be lost; thus some mechanism of assuring fidelity is required. Finally, all classes of polynucleotide polymerases must be able to translocate along the template being copied as synthesis proceeds. The crystal structures of numerous DNA polymerases from different families suggest that they all utilize an identical two-metal-ion-catalyzed polymerase mechanism but differ extensively in many of their structural features.

From amino acid sequence comparisons (1) as well as crystal structure analyses (2), the DNA polymerases can be divided into at least five different families, and representative crystal structures are known for enzymes in four of these families. Perhaps the best studied of these families is the DNA polymerase I (pol I)<sup>1</sup> or A polymerase family, which includes the Klenow fragments of *Escherichia coli* and a *Bacillus* DNA polymerase I, *Thermus aquaticus* DNA polymerase, and the T7 RNA and DNA polymerases, all of whose crystal structures are known (3–11). The second family of DNA-dependent DNA polymerases is DNA polymerase  $\alpha$  (pol  $\alpha$ ) or B family DNA polymerase. All eukaryotic replicating DNA polymerases and the polymerases from phages T4 and RB69 belong to this family, and a crystal structure of the RB69 polymerase shows some similarities to the pol I family enzymes and numerous differences (12). Reverse transcriptases (RT), RNA-dependent RNA polymerases, and telomerase appear to show some common structural similarities, whereas the structure of DNA polymerase  $\beta$  shows no structural relatedness to any of these previous families (13, 14). On the basis of amino acid sequence comparisons but no crystal structures, it appears that the bacterial DNA polymerase III enzymes also form a family that is unrelated to the polymerases of known structure (1).

Independent of their detailed domain structures, all polymerases whose structures are known presently appear to share a common overall architectural feature. They have a shape that can be compared with that of a right hand and have been described as consisting of “thumb,” “palm,” and “fingers” do-

mains (15). The function of the palm domain appears to be catalysis of the phosphoryl transfer reaction whereas that of the fingers domain includes important interactions with the incoming nucleoside triphosphate as well as the template base to which it is paired. The thumb on the other hand may play a role in positioning the duplex DNA and in processivity and translocation. Although the palm domain appears to be homologous among the pol I, pol  $\alpha$ , and RT families, the fingers and thumb domains are different in all four of these families for which structures are known to date (16).

Here the functional and structural similarities and differences among the polymerases of known structure are explored. Of particular interest are the role of editing in the fidelity of copying, the common enzymatic mechanism of polymerases, and the manners in which different domain structures function in the polymerase reaction in analogous ways.

### Structural Differences among Polymerases

Although the palm domains of the pol I, pol  $\alpha$ , and RT families are homologous, the fingers and thumb domains are completely different in the structures from all families (16). In the structure of the DNA polymerase from RB69 five domains are arranged around a central hole (12). In this enzyme the fingers domain consists largely of two very long anti-parallel coiled-coil  $\alpha$ -helices that extend more than 20 Å out the “back.” The thumb domain is seen to be interacting directly with the exonuclease domain and providing some of the binding site for the single-stranded exonuclease substrate. After orienting the palm domains of the RB69 and Klenow fragment enzymes identically, it becomes clear that their exonuclease domains are located in completely different places relative to the polymerase active site. Although in the “standard” orientation the exonuclease domain can be described as being southeast of the polymerase active site in Klenow fragment, the corresponding domain of the RB69 polymerase is located northwest of the polymerase active site. This difference in location of the editing domain may be in part related to the fact that the RB69 enzyme, like that from T4 phage, has an exonuclease activity that is 10<sup>3</sup> times larger than that of the Klenow fragment (17, 18).

A detailed comparison of the structures from four polymerase families (Fig. 1) shows that the fingers and thumbs are different in all four families for which structures are known (12, 16). (To make a suitable comparison between structural elements having similar functions, the names of the pol  $\beta$  thumb and fingers domains have been switched from the Pelletier *et al.* image (19).) Although the structures of the thumb domains are not homologous, they do exhibit analogous features that consist of largely parallel or anti-parallel  $\alpha$ -helices and in each case at least one  $\alpha$ -helix seems to be making important interactions across the minor groove of the primer-template product. In the case of the pol I family, loops at the top of the thumb also make important and conserved interactions with the DNA backbone (6, 8).

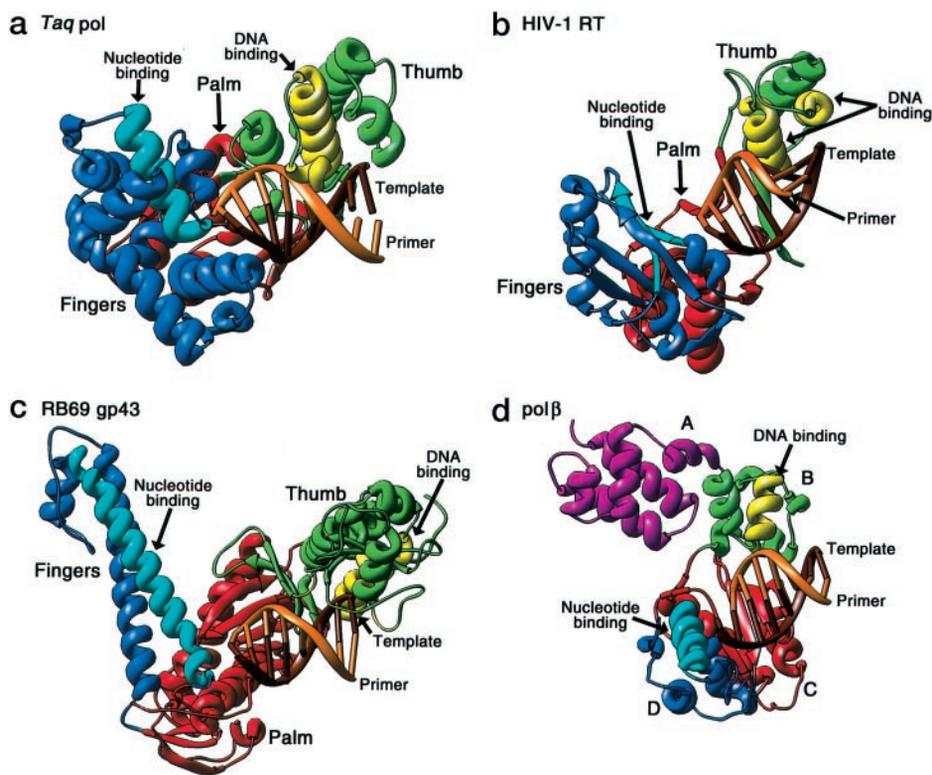
Although the fingers domains of all four families are also not homologous, there are some striking structural analogies among the families as with the thumbs. In three of the four DNA polymerase families with known structures, the pol I, pol  $\alpha$ , and pol  $\beta$  families, an  $\alpha$ -helix in the fingers domain is positioned at the blunt end of the primer-template; it contains side chains that are conserved within the families (the B motif) and provides important orienting interactions with the incoming deoxynucleoside

\* This minireview will be reprinted in the 1999 Minireview Compendium, which will be available in December, 1999.

‡ To whom correspondence should be addressed. Tel.: 203-432-5617; Fax: 203-432-3282.

<sup>1</sup> The abbreviations used are: pol, polymerase; RT, reverse transcriptase(s); HIV, human immunodeficiency virus.

**FIG. 1. A comparison of primer-template DNA bound to four DNA polymerases.** The complexes shown in *A*, *B*, and *D* are co-crystal structures, whereas the complex in *C* is a homology model (12). These four structures have been similarly oriented with respect to each other by superposition of the first two base pairs at the primer terminus. The primer strand and template strand are labeled. *A*, *Taq* DNA polymerase bound to DNA (6). As with the other three structures shown in this figure, the DNA stacks against the fingers and is contacted across the minor groove by the thumb domain. *B*, the binary complex of HIV-1 RT and DNA (29). This structure does not have a nucleotide-binding  $\alpha$ -helix in the fingers domains. Instead, a  $\beta$ -hairpin probably performs this function. *C*, the model of DNA bound to RB69 gp43 (12). A likely DNA-binding  $\alpha$ -helix has been highlighted. It appears that the thumb domain would have to move toward the primer terminus to bind DNA analogously to the other polymerases. *D*, the ternary complex of rat pol  $\beta$  with DNA and dideoxy-NTP (19). Domain D plays the role of the fingers and presents an  $\alpha$ -helix at the primer terminus. Domain B is analogous to other polymerase thumb domains and binds the minor groove of the duplex substrate. This figure is reprinted with permission from Brautigam and Steitz (16).



triphosphate. In the case of the reverse transcriptase family, however, it appears that some of these functions have been taken over by an anti-parallel  $\beta$ -ribbon, which lies in a similar position, as seen in a recent ternary complex with the primer-template and deoxynucleoside triphosphate (27).

The perhaps surprising diversity of polymerase structures found in these families leads one to wonder why the structures of DNA polymerases turn out to be so diverse when the structures of most metabolic enzymes, such as glyceraldehyde-3-phosphate dehydrogenase, are almost identical from microbe to man. One possibility is that the RT, pol I, and pol  $\beta$  families are later evolutionary additions to all cellular replicating polymerases, which could turn out to be divergently related to the pol  $\alpha$  family polymerases; however, sequence comparisons do not presently show such a relationship between the eukaryotic pol  $\alpha$  and the prokaryotic pol III DNA polymerases. An alternative speculation might imagine that a ribozyme DNA polymerase originating in the "RNA world" may have persisted beyond the divergence of eukaryotes and prokaryotes and was replaced domain by domain differently.

### Fidelity of Genome Copying

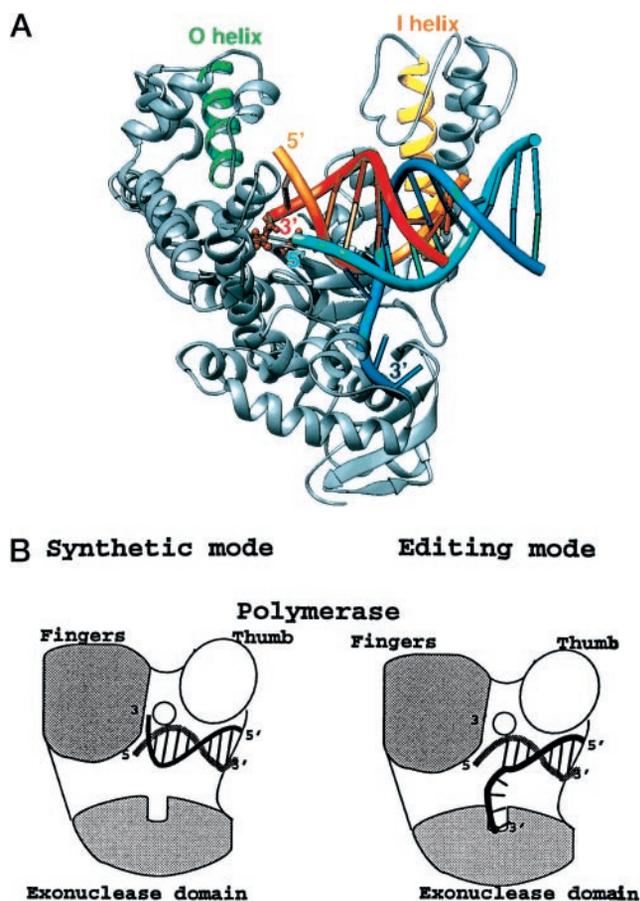
The greatest insights into the mechanisms by which polymerases achieve faithful copying of the template come from structural and biochemical studies of the pol I family of polymerases (2, 20). Fidelity arises both from constraints imposed on base pairing at the polymerase active site as well as the editing of mismatched base pairs at a 3'-exonuclease active site. The crystal structure of the Klenow fragment of DNA polymerase I first showed that this enzyme is divided into two domains, one of which catalyzes the polymerase reaction and the second of which has an active site more than 30 Å away from the polymerase active site and catalyzes the 3'-5'-exonuclease reaction (3). The co-crystal structure of Klenow fragment with duplex DNA containing a 3' overhanging tetranucleotide shows that the single-stranded 3' end binds into the exonuclease active site (4). Extensive structural, mutagenic, and biochemical studies of single-stranded substrates bound to the exonuclease ac-

tive site gave rise to the proposal of a two-metal ion mechanism of phosphoryl transfer (20, 21, 23). When duplex DNA is bound to the homologous *T. aquaticus* DNA polymerase, which does not contain a functioning exonuclease active site, the 3' end of the primer strand is found to lie in the polymerase active site adjacent to highly conserved carboxylate residues known to be important for the polymerase reaction (6). Likewise in the ternary complex between T7 DNA polymerase, primer-template DNA, and dNTP, the 3' end is in the polymerase active site. Although the duplex portion of substrates lies in the same approximate position adjacent to the thumb whether the primer-template is bound to the enzyme in polymerase mode or in exonuclease mode, the 3' ends are located in active sites that are separated by more than 30 Å.

The mechanism whereby the exonuclease domain exerts its editing function is proposed (Fig. 2) to involve a competition between these two active sites for the 3' end of the primer strand and a rapid shuttling of the primer terminus between them (2, 4, 21). The 3'-exonuclease active site binds single-stranded DNA whereas the polymerase active site binds duplex DNA with the ratio of about 1 to 10 for correctly Watson-Crick base-paired duplex DNA (18). Mismatched base pairs destabilize the duplex DNA and thereby enhance the binding of the 3' single-stranded DNA to the exonuclease active site. Furthermore, polymerization is stalled after incorporation of mismatched base pairs presumably because of misorientation of the 3'-hydroxyl group of the primer terminus onto which the next nucleotide is to be added. Once again, this stalling of the polymerization reaction serves to enhance the probability of excision by the exonuclease activity. The additional role that the polymerase domain plays in fidelity is considered below.

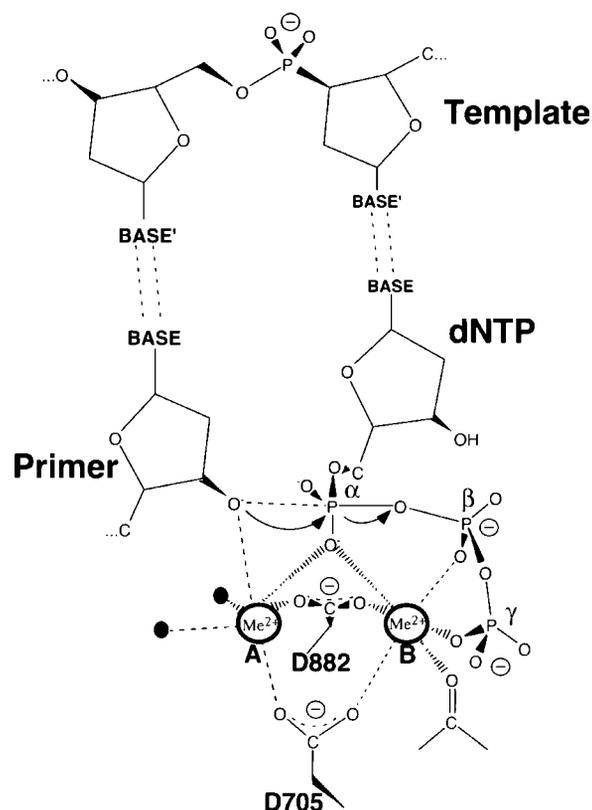
### Polymerase Mechanism

Structural studies as well as sequence comparisons among polymerases strongly suggest the hypothesis that the phosphoryl transfer reaction of all polymerases is catalyzed by a two-metal ion mechanism (Fig. 3) originally proposed (22, 24) by analogy to the well studied two-metal ion mechanism in the



**FIG. 2. The shuttle mechanism of editing in DNA polymerases (21).** A, superposition of DNAs bound in editing and in polymerizing modes. To orient the two DNAs, the polymerase domains of the Klenow fragment editing complex (4) and the *Taq* polymerase synthetic complex (6) were superimposed, and the DNA from the editing complex was added to the *Taq* polymerase-DNA complex. The 3' end of the primer strand in polymerizing mode is duplex and lies near three catalytically important carboxylates in the polymerase active site. The 3' end of the primer strand in editing mode is single-stranded and lies in the 3'-5'-exonuclease domain active site (6). B, the shuttling model for polymerase editing proposes that the equilibrium between the 3' end of the primer strand being bound as a single strand in the exonuclease active site (right) and bound as duplex at the polymerase active site (left) is shifted toward the editing mode by mismatched base pairs, which destabilize duplex DNA and retard addition of the next nucleotide. The shuttling of the 3' end between the two active sites is fast compared with the rate of next nucleotide addition. This figure is from Silvan *et al.* (L. Silvan, J. Wang, and T. A. Steitz, unpublished results).

3'-exonuclease reaction (20–23). The first observation of a polymerase complex with both primer-template DNA and dNTP-Mg<sup>2+</sup> bound to the polymerase active site that directly showed the structural basis of a two-metal ion mechanism was a complex with rat pol  $\beta$  (19). These two-metal ions are bound by three carboxylates contained in a domain that is not homologous to other polymerases (14). A higher resolution structure (2.1 Å) of human pol  $\beta$  complexed with a gapped DNA substrate and dideoxy-CTP shows precise detail of the interaction of two hexacoordinated, partially hydrated Mg<sup>2+</sup> ions interacting with the three phosphates (25). In the homologous “palm” domains of the pol I and RT families these two-metal ions (normally magnesium ions) are observed to bind to the enzyme through two completely conserved carboxylate residues but only in the primer-template complex with the dNTP (8). These metal ions are separated by just under 4 Å in the ternary complex of T7 DNA polymerase complexed with primer-template DNA and dNTP (8). Metal ion A interacts with the 3'-hydroxyl of the primer strand and is proposed (24) to lower the



**FIG. 3. The two-metal ion mechanism of DNA polymerase (24).** The two conserved aspartates have the *E. coli* DNA polymerase I numbers. The active site features two metal ions that stabilize the resulting pentacoordinated transition state. Metal ion A activates the primer's 3'-OH for attack on the  $\alpha$ -phosphate of the dNTP. Metal ion B plays the dual role of stabilizing the negative charge that builds up on the leaving oxygen and chelating the  $\beta$ - and  $\gamma$ -phosphates. This figure is reprinted with permission from Brautigam and Steitz (16).

$pK_a$  of the hydroxyl, facilitating its attack on the  $\alpha$ -phosphate of the incoming dNTP. Metal ions A and B are also proposed to stabilize both the structure and charge of the pentavalent transition state that occurs during the course of this reaction. Finally, metal ion B binds to and is proposed to facilitate the leaving of the  $\beta$ - and  $\gamma$ -phosphates. Chemically similar mechanisms of two-metal ion-catalyzed phosphoryl transfer reactions are used by many enzymes including ribozymes (26).

Although amino acid sequence comparisons had suggested that there were three highly conserved carboxylate-containing residues in the active sites of all classes of polymerases, comparison of the crystal structures of Klenow fragment, HIV-1 RT, RB69 pol  $\alpha$  polymerase, and the T7 RNA polymerase (12) shows that only two aspartic acid residues are structurally conserved among these four enzymes. Furthermore, the crystal structure of T7 DNA polymerase complexed with a primer-template and a deoxynucleoside triphosphate (8) shows that the two divalent metal ions are bound by only one aspartic acid residue emanating from conserved sequence motif A and one aspartate residue emanating from conserved sequence motif C (residues Asp-705 and Asp-882 of Klenow fragment). Further, comparison of the pol  $\alpha$  and pol I structures shows that previous sequence alignments of the motif C, which positioned the first Asp of a DXD sequence on the first Asp of a DE sequence (1), needs to be modified so that, rather, the second Asp of the pol  $\alpha$  family DXD sequence superimposes on the first Asp of the RT family DE sequence or the pol I family DD sequence (1).

In spite of the fingers domains of the pol  $\beta$ , RT, pol I, and pol  $\alpha$  DNA polymerases all having different evolutionary origins, they share some similar functional features. The binding of

dNTP to the pol  $\beta$ , RT, and T7 DNA polymerases complexed with primer-template DNA results in a significant rotation of the fingers domain when compared with the corresponding binary polymerase complexes with either DNA or dNTP (8, 19, 27). The orientation of the dNTP in the binary complex differs considerably from its orientation in the ternary complex. Only in the presence of the next correct dNTP is a ternary complex formed in which the fingers rotate and the incoming nucleotide makes a base pair with the template. In those ternary complexes in which the fingers do not rotate, the dNTP binds as in the dNTP binary complex. All of the structures together (8, 19, 25, 27) are consistent with the early proposal that dNTP first binds to a primer-template complex with the polymerase in a non-template-dependent fashion, but after a rate-limited conformational change (finger rotation?) tight dNTP binding is template sequence-specific (28). The proposal has been further made that only with the correct Watson-Crick base pairing between template and dNTP does the fingers rotation essential for catalysis occur (25). That is, fidelity for incorporation of the correct nucleotide at the polymerization step is enhanced by this catalytically essential, induced fit conformational change, which detects the presence of a correct base pair.

Furthermore, the four known non-homologous fingers domains present similar residues to the incoming dNTP for the same functional reasons. Arg-72 in HIV-1 RT and Lys-522 in T7 DNA pol interact with a non-bridging oxygen of the  $\alpha$ -phosphate in a way that would help stabilize the additional negative charge of the pentacovalent transition state. Furthermore, analogous basic residues from the fingers domains interact with the  $\gamma$ -phosphate of dNTP.

### Conclusions

From the extensive crystallographic, biochemical, and genetic studies of polynucleotide polymerases there are several general conclusions that can be drawn about this class of enzymes. First, all polynucleotide polymerases may use the same two-metal ion mechanism to catalyze the polymerase phosphoryl transfer reactions (24). It is perhaps of interest to note that such a mechanism, which involves only the properties of two correctly positioned divalent metal ions, could easily be used by an enzyme made entirely of RNA and thus could function in an all RNA world. Second, the fidelity of DNA synthesis results from a combination of "enforced" Watson-Crick interactions at the polymerase active site (8, 19, 25) and competitive editing at the 3'-exonuclease active site (4, 6, 21). Misincorporated nucleotides retard further synthesis, destabilize duplex DNA, and enhance binding to the exonuclease active site. Third, although the catalytically important palm domains are seen to be homologous in the pol I, RT, and pol  $\alpha$  families, the pol  $\beta$  family palm domain is decidedly unrelated (14); likewise, the catalytic domains from the DNA polymerase III and the multisubunit RNA polymerase families are likely to be different as judged from amino acid sequence comparisons. Fourth, the thumb and finger domains are structurally different in all of the polymerase families for which representative crystal structures are now known, although analogously positioned secondary structures function in similar ways.

### REFERENCES

- Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990) An attempt to unify the structure of polymerases. *Protein Eng.* **3**, 461–467
- Joyce, C. M., and Steitz, T. A. (1994) Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem.* **63**, 777–822
- Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. (1985) Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* **313**, 762–766
- Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993) Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* **260**, 352–355
- Kim, Y., Eom, S. H., Wang, J., Lee, D. S., Suh, S. W., and Steitz, T. A. (1995) Crystal structure of *Thermus aquaticus* DNA polymerase. *Nature* **376**, 612–616
- Eom, S. H., Wang, J., and Steitz, T. A. (1996) Structure of *Taq* polymerase with DNA at the polymerase active site. *Nature* **382**, 278–281
- Kiefer, J. R., Mao, C., Braman, J. C., and Beese, L. S. (1998) Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* **391**, 304–307
- Doublé, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**, 251–258
- Sousa, R., Chung, Y. J., Rose, J. P., and Wang, B.-C. (1993) Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Å resolution. *Nature* **364**, 593–599
- Jeruzalmi, D., and Steitz, T. A. (1998) Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme. *EMBO J.* **17**, 4101–4113
- Cheetham, G. M. T., Jeruzalmi, D., and Steitz, T. A. (1999) *Nature* **399**, 80–83
- Wang, J., Sattar, A. K. M. A., Wang, C. C., Karam, J. D., Konigsberg, W. H., and Steitz, T. A. (1997) Crystal structure of a pol  $\alpha$  replication DNA polymerase from bacteriophage RB69. *Cell* **89**, 1087–1089
- Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H., and Kraut, T. (1994) Crystal structure of rat DNA polymerase  $\beta$ : Evidence for a common polymerase mechanism. *Science* **264**, 1930–1935
- Steitz, T. A., Smerdon, S. J., Jäger, J., and Joyce, C. M. (1994) A unified mechanism for nonhomologous DNA and RNA polymerases. *Science* **266**, 2022–2025
- Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) Crystal structure at 3.5 Å of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**, 1781–1790
- Brautigam, C. A., and Steitz, T. A. (1998) Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* **8**, 54–63
- Huang, W. H., and Lehman, I. R. (1972) On the direction of translation of the T4 deoxyribonucleic acid polymerase gene *in vivo*. *J. Biol. Chem.* **247**, 7663–7667
- Capson, T. L., Peliska, J. A., Kaboord, B. F., Frey, M. W., Lively, C., Dahlberg, M., and Benkovic, S. J. (1992) Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. *Biochemistry* **31**, 10984–10994
- Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994) Structure of ternary complexes of rat DNA polymerase  $\beta$ , a DNA template-primer, and dCTP. *Science* **264**, 1981–1993
- Brautigam, C. A., and Steitz, T. A. (1998) Structural principles for the inhibition of the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I by phosphorothioates. *J. Mol. Biol.* **277**, 363–377
- Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R., and Steitz, T. A. (1988) Cocrystal structure of an editing complex of Klenow fragment with DNA. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 8924–8928
- Beese, L. S., and Steitz, T. A. (1991) Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.* **10**, 25–33
- Derbyshire, V., Grindley, N. D. F., and Joyce, C. M. (1991) The 3'-5' exonuclease of DNA polymerase I of *Escherichia coli*: contribution of each amino acid at the active site to the reaction. *EMBO J.* **10**, 17–24
- Steitz, T. A. (1993) DNA- and RNA-dependent DNA polymerases. *Curr. Opin. Struct. Biol.* **3**, 31–38
- Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J., and Pelletier, H. (1997) Crystal structures of human DNA polymerase  $\beta$  complexed with gapped and nicked DNA: Evidence for an induced fit mechanism. *Biochemistry* **36**, 11205–11212
- Steitz, T. A., and Steitz, J. A. (1993) A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 6468–6502
- Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: Implications for drug resistance. *Science* **282**, 1669–1675
- Bryant, F. R., Johnson, K. A., and Benkovic, S. J. (1983) Elementary steps in the DNA polymerase I reaction pathway. *Biochemistry* **22**, 3537–3546
- Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., and Clark, P., Hizi, A., Hughes, S. H., and Arnold, E. (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 6320–6324