

Structures of phi29 DNA polymerase complexed with substrate: the mechanism of translocation in B-family polymerases

Andrea J Berman¹, Satwik Kamtekar^{1,5},
Jessica L Goodman¹, José M Lázaro²,
Miguel de Vega², Luis Blanco², Margarita
Salas² and Thomas A Steitz^{1,3,4,*}

¹Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, USA, ²Centro de Biología Molecular ‘Severo Ochoa’ (CSIC-UAM), Universidad Autónoma, Canto Blanco, Madrid, Spain, ³Department of Chemistry, Yale University, New Haven, CT, USA and ⁴Howard Hughes Medical Institute, Yale University, New Haven, CT, USA

Replicative DNA polymerases (DNAPs) move along template DNA in a processive manner. The structural basis of the mechanism of translocation has been better studied in the A-family of polymerases than in the B-family of replicative polymerases. To address this issue, we have determined the X-ray crystal structures of phi29 DNAP, a member of the protein-primed subgroup of the B-family of polymerases, complexed with primer-template DNA in the presence or absence of the incoming nucleoside triphosphate, the pre- and post-translocated states, respectively. Comparison of these structures reveals a mechanism of translocation that appears to be facilitated by the coordinated movement of two conserved tyrosine residues into the insertion site. This differs from the mechanism employed by the A-family polymerases, in which a conserved tyrosine moves into the templating and insertion sites during the translocation step. Polymerases from the two families also interact with downstream single-stranded template DNA in very different ways.

The EMBO Journal (2007) **26**, 3494–3505. doi:10.1038/sj.emboj.7601780; Published online 5 July 2007

Subject Categories: genome stability & dynamics; structural biology

Keywords: DNA polymerase; phi29; structure; translocation

Introduction

Single-subunit replicative polymerases contain a polymerase domain divided into functional subdomains arranged in a gross common architecture likened to a right hand (Kohlstaedt *et al*, 1992). The thumb and fingers subdomains

form the sides of a ‘U’-shaped cleft, at the bottom of which is the catalytic palm subdomain that utilizes a two-metal ion mechanism for catalyzing phosphodiester bond formation (Steitz *et al*, 1993). The thumb subdomain stabilizes the primer-template duplex product and the fingers subdomain contains basic residues that bind the triphosphate moiety of the incoming nucleotide and the pyrophosphate product of the phosphoryl transfer reaction (Beese *et al*, 1993; Doublie *et al*, 1998).

The coordinated movements of these subdomains have been extensively studied in polymerase families, including family A (bacterial repair polymerases, most bacteriophage replicative polymerases, and T7 RNA polymerase (RNAP)) and family B (viral and eukaryotic genome replicating enzymes) (Rothwell and Waksman, 2005). Structural studies have led to the suggestion that after binding a primer-template DNA substrate, A-family polymerases bind an incoming nucleoside triphosphate at a pre-insertion site located near the fingers subdomain before escorting it into the insertion site (Beese *et al*, 1993; Li *et al*, 1998a; Temiakov *et al*, 2004), whereas it has been proposed from biochemical studies that B-family polymerases bind the incoming nucleoside triphosphate directly in the insertion site at the base of the fingers (Yang *et al*, 2002b). Structural studies of A-family polymerases have also proposed a pre-insertion site for the templating base in the replication cycle of this family (Johnson *et al*, 2003; Temiakov *et al*, 2004; Yin and Steitz, 2004); no evidence for a templating pre-insertion site in the B-family exists. Following the phosphoryl transfer reaction, the newly incorporated nucleotide moves from the insertion site to the priming site, allowing the next incoming nucleotide to bind. This last step, known as translocation, facilitates processive movement of a polymerase along template DNA and is therefore a critical feature of the nucleotide addition cycle of replicative polymerases (Figure 1).

Polymerases are molecular machines that convert chemical energy into mechanical energy, and two models, the power stroke and Brownian-ratchet mechanisms, have been used to explain the energetics of translocation (Hanson and Huxley, 1955; Simon *et al*, 1992). In the context of the polymerization cycle, the power stroke mechanism derives the energy for translocation from the dissociation of the pyrophosphate product of the nucleotidyl transfer reaction, whereas the Brownian-ratchet mechanism utilizes the kinetic energy of primer-template diffusion to facilitate the unidirectional movement of the polymerase along the template strand (Guajardo and Sousa, 1997). Each model has testable predictions, and it is possible that a combination of the two occurs.

Despite a wealth of studies on processive RNAPs and bacterial DNA polymerases (DNAPs), little is known about the translocation step in eukaryotic-like replicative DNAPs. The B-family DNAP of *Bacillus subtilis* bacteriophage phi29 is an appealing system in which to study the structural biology

*Corresponding author. Department of Molecular Biophysics and Biochemistry, Yale University, Room 418, Bass Center, 266 Whitney Avenue, New Haven, CT 06520-8114, USA.

Tel.: +1 203 432 5617/5619; Fax: +1 203 432 3282;

E-mail: eatherton@csb.yale.edu

⁵Present address: Pfizer Inc., 700 Chesterfield Parkway West, Chesterfield, MO 63017, USA

Received: 26 April 2007; accepted: 30 May 2007; published online: 5 July 2007

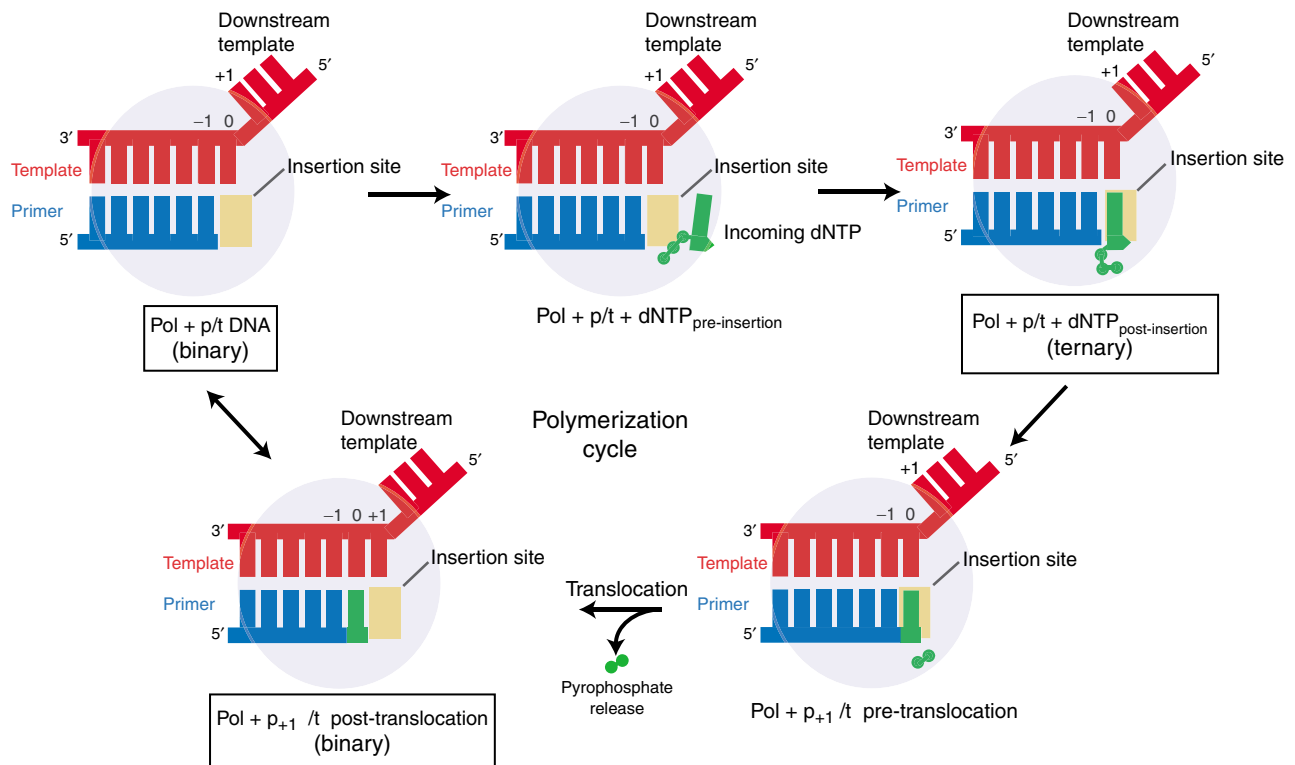


Figure 1 The polymerization cycle. Polymerase (pale blue circle) binds a primer-temple substrate (blue and red) and then the incoming dNTP (green). In some polymerases, the incoming dNTP binds a pre-insertion site before binding the insertion site (yellow) opposite the templating nucleotide. The polymerase then catalyzes the incorporation of the dNTP into the primer strand, resulting in a primer extended by one nucleotide and a molecule of pyrophosphate bound near the active site. Release of the pyrophosphate is associated with translocation of the new primer terminus out of the insertion site and into the priming site (Yin and Steitz, 2004). Boxes indicate the states captured in our crystal structures. For consistency, template strand numbering refers to the base positions in the initial binary complex.

of B-family replication, because it is small and biochemically well characterized (Blanco and Salas, 1996). In addition to the general features of polymerase and exonuclease activities shared by B-family polymerases, phi29 DNAP has a strand displacement capacity and high processivity. It can also initiate replication from a protein primer, terminal protein (TP), a characteristic that is shared by polymerases from several pathogenic viruses, such as adenovirus, poliovirus, and hepatitis virus (Salas, 1991).

Initial structural studies provided insights into the intrinsic strand displacement, processivity, and protein priming activities of phi29 DNAP. The structure of the *apo* phi29 DNAP exhibited two globular domains, an N-terminal exonuclease domain and a C-terminal polymerase domain. This structure contains three tunnels, one of which is formed by the exonuclease domain and the palm and TP-interacting region 2 (TPR2) subdomains of the polymerase domain. Homology modeling of a substrate complex using the primer-temple DNA and incoming nucleotide substrates from the ternary complex of the B-family DNAP from bacteriophage RB69 (Franklin *et al*, 2001) identified this tunnel at the location where phi29 DNAP would bind the downstream 5' region of single-stranded template DNA and suggested mechanisms for processivity and strand displacement (Kamtekar *et al*, 2004); truncation of the TPR2 subdomain reduced processivity and strand displacement, consistent with the proposed mechanism (Rodríguez *et al*, 2005). These structures also showed that the two subdomains, TPR1 and TPR2, which are only present in protein-primed polymerases, interact with the

intermediate and priming domains of TP, respectively (Kamtekar *et al*, 2006).

Here, we present four crystal structures of complexes of phi29 DNAP with substrates. These include the structure of polymerase bound to a primer-temple substrate (binary complex) in the post-translocated state, before the next incoming nucleotide binds the polymerase, as well as the structures of complexes of polymerase bound to two different primer-templates and their complementary incoming nucleotides (ternary complexes). Finally, we describe the structure of polymerase bound to single-stranded DNA (ssDNA) (ssDNA complex). Comparison of the structures of these complexes allows us to understand ssDNA and double-stranded DNA binding in B-family DNAPs and to propose a mechanism of translocation in this family.

Results

Four views of substrate binding: binary, ternary, exonuclease, and ssDNA template complexes

The structures described here represent different stages of the replication cycle. One of them contains a molecule of polymerase bound to ssDNA in a tunnel that lies downstream of the active site, a complex that is relevant to understanding protein-primed initiation. This structure also contains ssDNA bound at the exonuclease active site. The binary complex structure contains polymerase bound to a primer-temple DNA substrate that is in the post-translocated position. The two ternary complexes contain primer-temple DNA and an

Table I Data collection and refinement statistics for phi29 DNAP-substrate complexes

	Ternary1	Ternary2	Binary	ssDNA complex
<i>Substrates</i>				
Primer ^a	5'-GACTGCTT ACAT	5'-GACTGCTT ACG	5'-GACTGCTT AC	
Template ^b	3'-CTGACGAATG TACA	3'-CTGACGAATG CACA	3'-CTGACGAATG CACAATC	
ssDNA				5'GGACTTT
<i>Data collection</i>				
Resolution limit (Å)	50.0–2.2 (2.28–2.20)	50.0–2.03 (2.10–2.03)	50.0–2.6 (2.69–2.6)	50.0–1.6 (1.66–1.6)
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁	C2	P2 ₁
<i>Copies in the AU</i>				
Protein	1	2	4	2
DNA	1	3	2/2 ^c	5
<i>Cell parameters</i>				
a, b, c (Å)	58.9, 78.2, 157.8	72.8, 114.7, 104.8	216.6, 146.3, 115.1	54.2, 200.2, 67.0
α, β, γ (deg)	90, 90, 90	90, 94.1, 90	90, 117.9, 90	90, 109.4, 90
Unique reflections	36737	107450	95762	172311
I/σ(I)	8.9 (1.2)	16.3 (2.4)	5.8 (1.0)	49.9 (3.7)
Redundancy	4.8 (2.5)	3.5 (2.3)	2.1 (1.9)	6.9 (5.5)
Completeness (%)	97 (83.7)	98 (83.9)	98.4 (95.1)	97.8 (92.2)
R _{merge}	0.166 (0.85)	0.076 (0.353)	0.152 (0.773)	0.045 (0.44)
<i>Refinement statistics</i>				
Data range	40–2.20 (2.26–2.20)	40–2.03 (2.09–2.03)	50–2.6 (2.67–2.60)	50–1.6 (1.64–1.60)
R _{work} /R _{free}	19.5/25.4 (26.8/34.0)	18.7/23.4 (23.5/32.3)	21.9/27.3 (36.6/39.7)	16.1/19.4 (22.2/27.5)
<i>RMSD</i>				
Bond lengths (Å)	0.010	0.009	0.009	0.011
Bond angles (deg)	1.3	1.3	1.2	1.4
<i>Number of atoms</i>				
Protein	4680	9367	18488	9267
Nucleic acid	534	1493	1207	553
Ions	2	4	0	0
Water	357	969	89	1494
<i>Average B-factors (Å²)</i>				
Protein	32.9	22.1	27.3	20.7
Nucleic acid	39.7	23.6	29.3	23.9
PDBID	2PYL	2PYJ	2PZS	2PY5

^aThe incoming dNTP of each ternary complex is indicated in bold.

^bThe 0 position on the template strand is underlined.

^cIn the binary crystal form, there are two copies of primer-template substrate and two copies of single-stranded template.

incoming nucleotide (dNTP) that is complementary to the templating base (0 position) (Figure 1). The sequences of the substrates differ at several positions in these ternary complexes, facilitating sequence-specific comparisons (Table I).

Although the exonuclease and polymerase domains move slightly with respect to each other, the global structures of the polymerases complexed with substrate remain largely unchanged when compared to the *apo* structures of polymerase (PDBID 1X1I and 1XHX). The root mean squared deviations (RMSD) calculated over all pairs of polymerases in the *apo* and complex structures range from 0.7 to 2.8 Å over 572 Cα. The catalytic palm subdomains (residues 190–260 and 427–530) of all of these copies of polymerase are very similar, with an RMSD range of 0.3–1.0 Å over 173 Cα.

Substrate binding

Single-stranded DNA complex. Phi29 DNAP bound to ssDNA crystallizes in space group P2₁ and diffracts to better than 1.6 Å resolution (Table I). The two non-crystallographically related copies in the asymmetric unit are very well ordered, except for amino acid residues 305–311 in copy A and

residues 305–313 in copy B. These residues are part of a mobile loop in TPR1 that is only well ordered in the presence of TP (Kamtekar *et al*, 2006) or duplex DNA product. The two copies of polymerase are very similar, with an RMSD of 0.8 Å over the 561 Cα atoms, but exhibit significant differences in exonuclease residues 140–145 and at residue Y165 that may have functional implications for exonuclease activity (Supplementary Figure S1). Each copy of polymerase in this crystal form binds the 3' end of one ssDNA emerging from the downstream template tunnel at the polymerase active site, and the 3' end of another ssDNA at the exonuclease active site. One copy of polymerase binds a third ssDNA in a biologically non-relevant location (Supplementary Figure S3c).

Binary complex. The crystals of the binary complex diffract to 2.6 Å resolution and contain four copies of polymerase per asymmetric unit related by pseudo-222 non-crystallographic symmetry (Table I). Two of the copies were modeled with a primer-template substrate (Supplementary Figure S3a). The other two copies of polymerase in this crystal form have density for a single-stranded 5' template overhang in the

downstream template tunnel, but are missing amino acids 306–314 and the duplex region of the primer-template due to disorder. Averaged electron density maps indicated the presence of one of the two missing primer-template duplex regions, but the quality of this density was poor and it was therefore not included in the final model.

The binary complex is representative of a post-translocation state (Figure 1), as the primer terminus occupies the priming site. The phosphate moiety of the priming nucleotide interacts with the invariant Y500 in motif KxY as predicted (Figure 2D) (Blasco *et al*, 1995). Two residues, Y254 and Y390, occupy the insertion site. The former (Y254) is called the steric gate residue, because its location in the active site

would lead to a steric clash with the 2'-hydroxyl of a ribonucleotide (Gao *et al*, 1997; Franklin *et al*, 2001), thereby preventing the incorporation of ribonucleotides into the primer strand (Bonnin *et al*, 1999). The latter amino acid, Y390, is from a B-family conserved sequence motif at the base of the fingers subdomain. Neither of the catalytic metal ions is observed in this complex, and, similar to the binary complexes from the A-family (Li *et al*, 1998b; Johnson *et al*, 2003), one of the catalytic aspartates (D249) is not properly oriented for catalysis (Figure 2A and C).

Ternary complexes. The ternary1 and ternary2 complexes crystallize in space groups $P2_12_12_1$ and $P2_1$, and diffract to

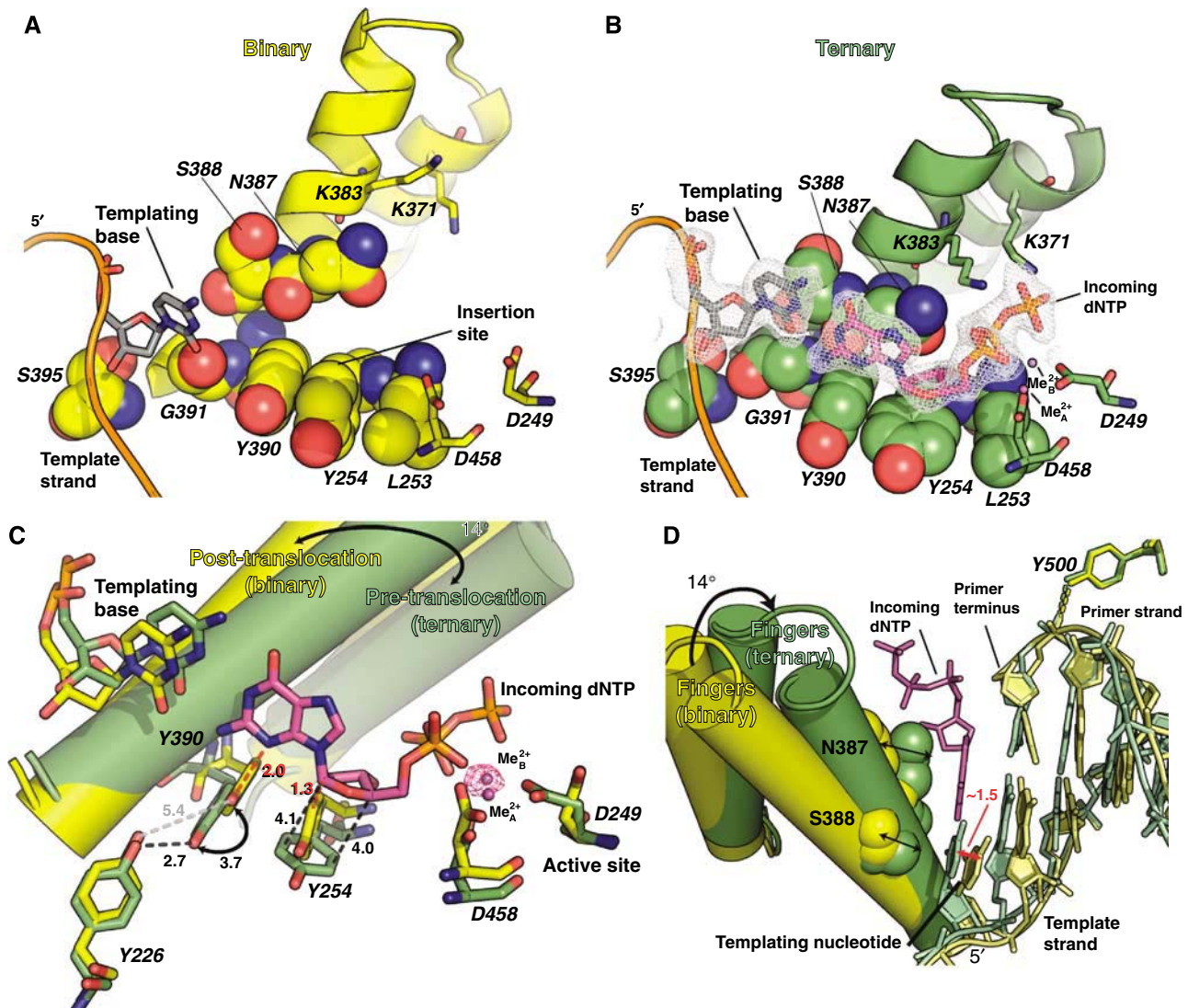


Figure 2 Comparison of the binary and ternary complexes. The binary complex is shown in yellow and the ternary complex in green. Metals are indicated as magenta spheres. The incoming dNTP from the ternary complex is shown as magenta sticks. The fingers subdomain rotates 14° in going from the opened binary complex to the closed ternary complex. (A) Binary complex. The residues that form the nascent base pair-binding site in the ternary complex are shown as spheres and the active site carboxylates are shown as sticks. The fingers subdomain is shown in cartoon representation. Two conserved tyrosine residues occupy the insertion site. (B) Ternary complex. The conserved lysine residues that interact with the phosphates are also shown. The density from a simulating annealing omit map using phases calculated from a model with the nascent-base pair omitted and amplitudes from the ternary2 data contoured at 2.5σ is shown as gray mesh for the nascent base pair. (C) Comparison of the binary and ternary complex structures. All of the mechanically significant amino-acid movements are indicated. Black dashed lines represent interactions. Red dashed lines indicate steric clashes. The distances indicated are in Å. The density shown for metal ion B (manganese) is from an anomalous difference Fourier map calculated using data from 50.0–2.03 Å resolution and contoured at 6σ . (D) The propagated shift of the DNA base pair planes between the binary and ternary complexes. When the fingers close, residues S388 and N387 (shown as spheres) stack against the templating nucleotide and incoming dNTP, respectively, completing the nascent base pair-binding pocket. Y500 interacts with the phosphate moiety of the priming nucleotide.

2.2 and 2.0 Å resolution, respectively (Table I). The orthorhombic crystal form contains one copy of polymerase bound to primer-template DNA and incoming nucleotide per asymmetric unit. The monoclinic crystal form contains two ternary complexes and a third copy of primer-template DNA bound in a biologically non-relevant manner (Supplementary Figure S3b). The RMSDs over the 173 C α in the catalytic palm subdomain among the three copies of ternary complex range from 0.6–0.8 Å. We have chosen copy A of the ternary2 complex, except where noted, as a representative ternary complex in all of the following discussion, because it is well ordered in electron density maps (Figure 2B).

In the ternary complexes, the dNTP is bound at the insertion site, poised for catalysis (Figure 1) and the primer terminus occupies the priming site. The priming nucleotide in each of the ternary complexes interacts with Y500 of motif KxY (Figure 2D) (Blasco *et al*, 1995). In each complex, the base moiety of the dNTP forms a Watson–Crick base pair with the templating nucleotide and its deoxyribose ring stacks on the phenolic side chain of the steric gate residue, Y254. This steric gate side chain occupies a less favorable rotameric state, which is also observed in other ternary complexes (Huang *et al*, 1998; Franklin *et al*, 2001), suggesting that the stacking interaction with a sugar moiety of a nucleotide stabilizes its unusual conformation. Along with Y254, the side chain of the tyrosine from conserved sequence motif 2a (Y390) forms part of the nascent base pair-binding pocket. Y390 also interacts with the hydroxyl of Y226 through a hydrogen bond. Both aspartates that bind the catalytic magnesium ions participate in the active site (Figure 2B and C).

The phosphates of the incoming dNTP interact with the basic side chains of residues K371 and K383 from the fingers subdomain. Consistent with mutational data, the conserved sequence motif B residue K383 (Saturno *et al*, 1997) interacts with the α - and γ -phosphates of the incoming dNTP, and the pre-B motif residue K371 (Truniger *et al*, 2002) interacts with the γ -phosphate (Figure 2B). It is possible that these residues comprise part of a pre-insertion-binding site for the nucleotide, as, in the *apo* structure, they were observed to interact with sulfate ions which are sterically and electrostatically similar to the phosphate groups of a nucleotide (PDBID: 1XHX) (Kamtekar *et al*, 2004), although kinetics experiments with the B-family DNAP from bacteriophage RB69 have been interpreted to indicate the absence of a pre-insertion site in that system (Yang *et al*, 2002b). Biochemical experiments (Truniger *et al*, 2004) and sequence alignments with a conserved arginine from the A-family (Doublé *et al*, 1998) have implicated a third lysine residue, K379, in dNTP binding. The structure shows that K379 interacts with the γ -phosphate indirectly through a network of water molecules.

The ternary complex structure contains both metal ions, A and B, which have respectively been assigned as a magnesium ion and a manganese ion in the ternary2 complex based on an anomalous difference Fourier map (Figure 2C). The α - and γ -phosphates of the incoming dNTP, the catalytic aspartate residues (D249 and D458), and the carbonyl of V250 of the palm subdomain coordinate the catalytic metals. A 2', 3'-dideoxynucleotide was incorporated at the primer terminus to facilitate the formation of a ternary complex, and the absence of the 3' hydroxyl results in a slightly skewed coordination geometry of metal ion A.

While the structure of this ternary complex confirms the general substrate positioning in phi29 DNAP that was predicted from our previous homology modeling using the ternary complex of RB69 DNAP, some features are clearly different. The presence of the TPR2 subdomain in phi29 DNAP that is absent in RB69 DNAP results in a slight shift in the position of the DNA from its homology model placement (Franklin *et al*, 2001; Kamtekar *et al*, 2004). As predicted, the upstream duplex is topologically encircled by the thumb and TPR2 subdomains, but the base pairs distal to the active site interact with subdomain TPR2, resulting in their displacement of ~ 5 Å off the active site relative to the DNA in the homology-modeled complex. The structure also resolves the minor clashes between the thumb and the upstream duplex observed in the modeling. Likewise, the single-stranded downstream template enters the active site through the downstream template tunnel formed from the exonuclease domain and the TPR2, palm, and fingers subdomains, as predicted, but the interactions within the tunnel were unpredicted, and have implications for sequence-independent recognition of template DNA as well as for the mechanism of translocation.

Comparison of the pre-translocation ternary and post-translocation binary complexes. The binding of the incoming dNTP triggers a 14° rotation of the fingers subdomain toward the polymerase active site (Figure 2), corresponding to a ~ 7 Å movement of the tip of the fingers. As in other polymerases, the triphosphate moiety of the incoming nucleotide acts as an electrostatic crosslink between conserved residues of the fingers and the catalytic metal ions chelated to the conserved carboxylates, thereby keeping the fingers closed (Doublé *et al*, 1998; Huang *et al*, 1998; Li *et al*, 1998b; Franklin *et al*, 2001; Yin and Steitz, 2004). Once closed, the fingers complete the nascent base pair-binding pocket (Figure 2B and C).

The structure of the duplex DNA in the binary complex is distorted compared to its structure in the ternary complex. The nucleotide bases in the binary structure are substantially displaced, with the entire nucleotide at the -1 position of the template strand lifted almost 2 Å off the active site, whereas the positions of the phosphate backbones shift with an RMSD of less than 1 Å. The distortion of the duplex DNA appears to be a consequence of the position of the templating nucleotide. When the fingers are closed, the nascent base pair-binding pocket holds the templating nucleotide in position and the upstream bases of the template strand stack accordingly. However, in the binary complex, where the fingers are opened, the residues completing the nascent base pair binding pocket are too far away to stabilize the nucleotide in the templating position. This results in the displacement of the templating nucleotide by ~ 1.5 Å upstream from its position in the ternary complex; the stacking of the upstream nucleotides follows, slightly distorting the duplex (Figure 2D). Similar systematic shifts are observed in comparing the binary and ternary complexes of the A-family polymerases from *B. stearothermophilus* (Johnson *et al*, 2003) and *Thermus aquaticus* (Li *et al*, 1998b), and the X-family polymerase β from rat (Pelletier *et al*, 1994, 1996), suggesting that this could be the more stable DNA conformation in the absence of an incoming dNTP.

Despite these shifts in the binary and ternary complexes, an extensive water network mediating most of the protein–

nucleic acid interactions is conserved among different complexes. In both the binary and ternary complexes, the polymerase makes contacts with the sugar-phosphate backbone of duplex DNA through a few direct interactions and through multiple water-mediated hydrogen bonds (Figure 3 and Supplementary Figure S2). The only direct side chain contact with the minor groove of the duplex product is made by a highly conserved lysine (K498) that interacts with the N3 of a purine or the O2 of a pyrimidine at the primer strand -2 position (Figure 3). More than thirty ordered water molecules facilitate hydrogen bonds between conserved and nonconserved amino acids and the DNA duplex in each of the ternary complexes and in the binary complex to maintain flexibility in duplex binding (Figure 3 and Supplementary Figure S2). Several of these water molecules are also present in the *apo* polymerase structures. These water molecules thus act as surrogate side chains, as the entropic penalty for their immobilization is independent of the binding of DNA duplex.

The opening of the fingers that occurs in the transition from the ternary complex to the binary complex is accompanied by several mechanistically significant changes. When the fingers open, the side chain of Y390 from conserved sequence motif 2a moves into the insertion site, such that the newly incorporated nucleotide can no longer reside there.

This observation is consistent with biochemical data suggesting that Y390 interacts either directly or indirectly with the incoming dNTP (Blasco *et al*, 1992). If no nucleotide occupies the insertion site, the steric gate residue (Y254) can flip to its most favorable rotamer. This rotamer places the phenolic ring of the steric gate residue directly in the insertion site, stacking on the conserved tyrosine at the base of the fingers (Y390), one of the most energetically stable tyrosine-tyrosine interactions (Chelli *et al*, 2002) (Figure 2A and C). The positions of both of these tyrosine residues in the insertion site preclude the primer terminus from binding at the insertion site while the fingers are opened. Therefore, the primer terminus must move to the priming site, resulting in translocation of the DNA by one nucleotide.

The rotation of Y390 breaks its hydrogen bond with Y226 (Figure 2C), a residue in the conserved B-family I/YxGG/A sequence motif that has been proposed to be involved in template binding at the active site and in protein priming (Truniger *et al*, 1996, 1999; Brenkman *et al*, 2001). In the structures of these complexes, this motif stabilizes the nucleotides in the -1 and -2 positions of the template strand by van der Waals and hydrogen-bonding interactions (Figure 4A), as predicted by mutagenesis studies in B-family polymerases (Truniger *et al*, 1999; Brenkman *et al*, 2001).

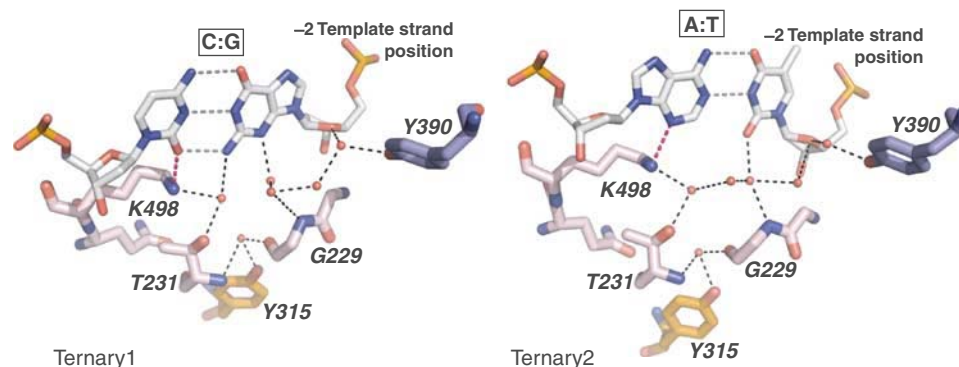


Figure 3 Water-mediated interactions maintain sequence nonspecific binding. The C:G base pair is from the ternary1 complex, and the A:T base pair is from the ternary2 complex. Red spheres are water molecules and black dashes are hydrogen bonds. Amino acids are colored by subdomain as in Kamtekar *et al* (2004).

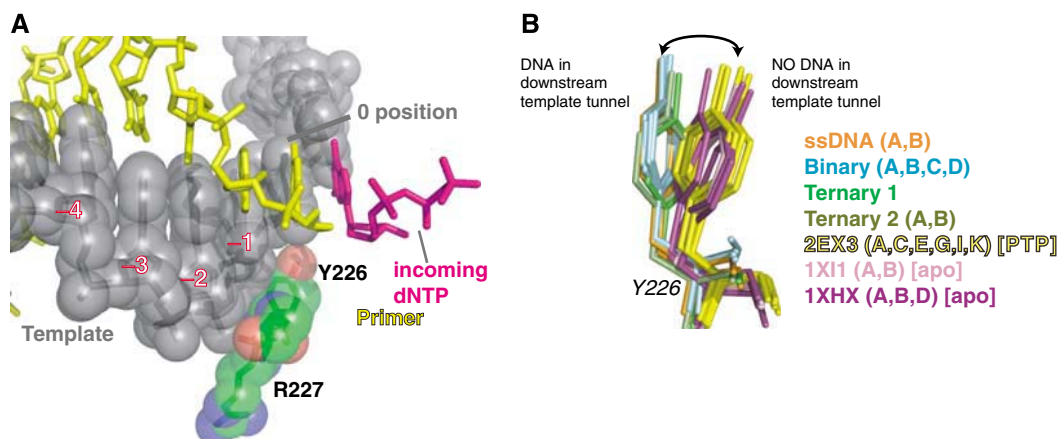


Figure 4 The I/YxGG/A motif. (A) The primer and template strands from the ternary complex are shown as yellow and gray sticks, respectively. The template strand and the residues of the I/YxGG/A motif are shown as spheres. (B) The two distinct populations of Y226 are shown in sticks based on a superposition of the palm subdomain. The residues are colored by crystal structure.

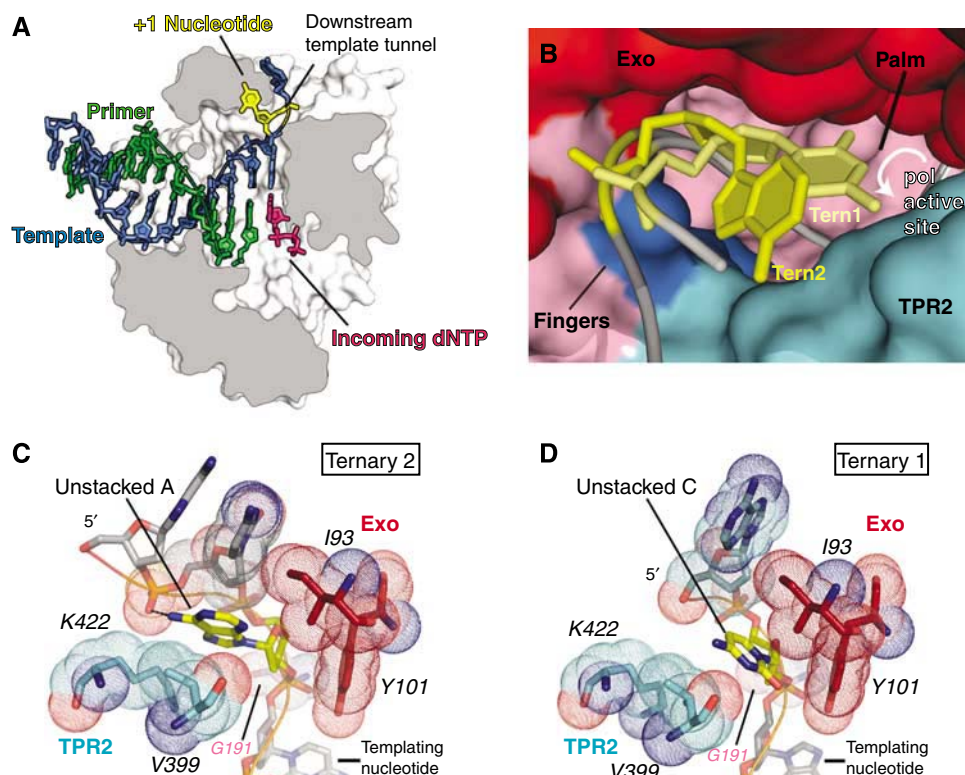


Figure 5 ssDNA in the downstream template tunnel. (A) A space filling representation of polymerase sliced through a plane showing the primer-template substrate and the single-stranded 5' template overhang in the downstream template tunnel. The +1 nucleotide is shown in yellow, the incoming nucleotide in magenta, the primer strand in green and the template strand in blue. (B–D) The ssDNA substrate is gray, the +1 nucleotide is yellow. Polymerase residues are colored by domain and subdomain as in Kamtekar *et al* (2004). (B) An overlay of the substrates from the ternary1 (light yellow and light gray) and ternary2 (dark yellow and dark gray) crystal forms. The polymerase shown is a space filling representation of the ternary1 crystal form. The polymerase contains an unstacked pyrimidine, indicating that the downstream template tunnel does not constrict around pyrimidine bases. (C) A purine base in the unstacked position interacts with residues from the exonuclease domain and the TPR2 subdomain in the ternary2 crystal form. The van der Waals radii of the residues are indicated by dots. (D) A pyrimidine base in the +1 template position in the ternary1 crystal form.

However, in the absence of DNA, such as in the *apo* polymerase (PDBID 1XHX and 1XI1) or the polymerase–TP complex (PDBID 2EX3), Y226 assumes a position that would sterically clash with the path of template DNA emerging from the downstream template tunnel into the active site. These two distinct populations are consistent among 20 of the 21 crystallographically independent copies of phi29 DNAP available (Figure 4B), and may provide insight into the role of this residue in protein-primed initiation of replication.

Downstream template. As expected for a processive replicative polymerase, phi29 DNAP interacts with ssDNA in a sequence nonspecific manner (Figure 5). In all complexes containing ssDNA, residues in the downstream template tunnel interact with the two nucleotides that lie immediately downstream (+1 and +2) of the templating nucleotide (0).

The base of the +1 nucleotide on the template strand is unstacked from the bases of adjacent nucleotides of the single-stranded 5' template overhang (Figure 5A). The base of this unstacked nucleotide fits into a pocket formed by residues V399 and K422 (TPR2 subdomain) and I93 (exonuclease domain) and completed by the nucleotide 5' to the unstacked nucleotide, whereas the sugar stacks on the side chain of Y101 (Figure 5C and D). The downstream template tunnel does not pack tightly around the unstacked pyrimidine

base (Figure 5B) and is large enough to accommodate a purine base, suggesting that during processive synthesis, the size of the downstream template tunnel may remain constant. The +2 nucleotide sits on a hydrophobic surface formed by exonuclease residues M102, I93, and M188. Presumably, the large number of hydrophobic interactions with the bases in the downstream template tunnel compensates for the energy lost by unstacking the +1 nucleotide.

Several hydrophilic residues at the edges of the downstream template tunnel stabilize the polar groups of the nucleotides. Residues Y101, T189, S192, K392, and N396 interact with the backbone through water-mediated and direct hydrogen bonds. The +2 nucleotide interacts through water-mediated hydrogen bonds with D104 and N91. Finally, within the downstream template tunnel, the functional group at the C6 position of a +1 purine interacts with the phosphate of the +2 nucleotide (Figure 5C); no interaction between this phosphate and a +1 pyrimidine is observed (Figure 5D).

Discussion

Those DNAPs that processively replicate long stretches of DNA must maintain intimate, but sequence nonspecific interactions with their substrates. The crystal structures with ssDNA and double-stranded DNA presented here illuminate

how this is achieved by phi29 DNAP. In addition, the structures of the binary and ternary complexes provide insight into the mechanism of translocation in the B-family of polymerases.

Downstream template binding/pre-insertion site

Polymerases universally possess a binding pocket for the nascent base pair that contains a flat interface with which to check for planarity inherent in a correct Watson–Crick nascent base pair (Blasco *et al*, 1993; Freisinger *et al*, 2004). Consequently, the base of the +1 nucleotide is unstacked from the templating nucleotide in polymerase families A, B, X, and Y, reverse transcriptases, and eukaryotic RNAPs (Pelletier *et al*, 1994; Doublé *et al*, 1998; Huang *et al*, 1998; Li *et al*, 1998b; Franklin *et al*, 2001; Gnatt *et al*, 2001). In phi29 DNAP, this is accomplished by a kinking of the template strand that is necessary for it to avoid steric clashes with the C-terminal α -helix of the exonuclease domain and to allow conserved residues from the fingers subdomain (N387 and S388 in phi29), shown to interact with the primer-template and affect polymerization activity (Blasco *et al*, 1993), to stack on the nascent base pair when the fingers close (Figure 2B–D). Comparisons of the structures of phi29

DNAP and other polymerases show that the way in which they interact with +1 nucleotides varies greatly among polymerase families and even within them, although they may all share a common theme of compensating for the loss of base–base stacking energy through van der Waals interactions between the polymerase and this base.

These differences in the interactions between polymerase and the unstacked nucleotide can be illustrated by comparing examples of A- and B-family polymerases. In the A-family, when the fingers are opened, a tyrosine from the fingers subdomain occupies the templating and insertion sites, precluding the next templating nucleotide from stacking on the upstream duplex (Li *et al*, 1998b; Johnson *et al*, 2003; Yin and Steitz, 2004). The polymerase accommodates the unstacked nucleotide in a hydrophobic pre-insertion site that is a relatively flat surface in T7 RNAP, and a pocket in *B. stearothermophilus* DNAP (Kiefer *et al*, 1998; Johnson *et al*, 2003; Yin and Steitz, 2004). When the fingers close, the unstacked nucleotide moves into templating position as the tyrosine moves out, and residues from the fingers domain move in to collapse the pre-insertion pocket (Figure 6B). These structural observations are consistent with fluorescence studies suggesting the dramatic movement of the

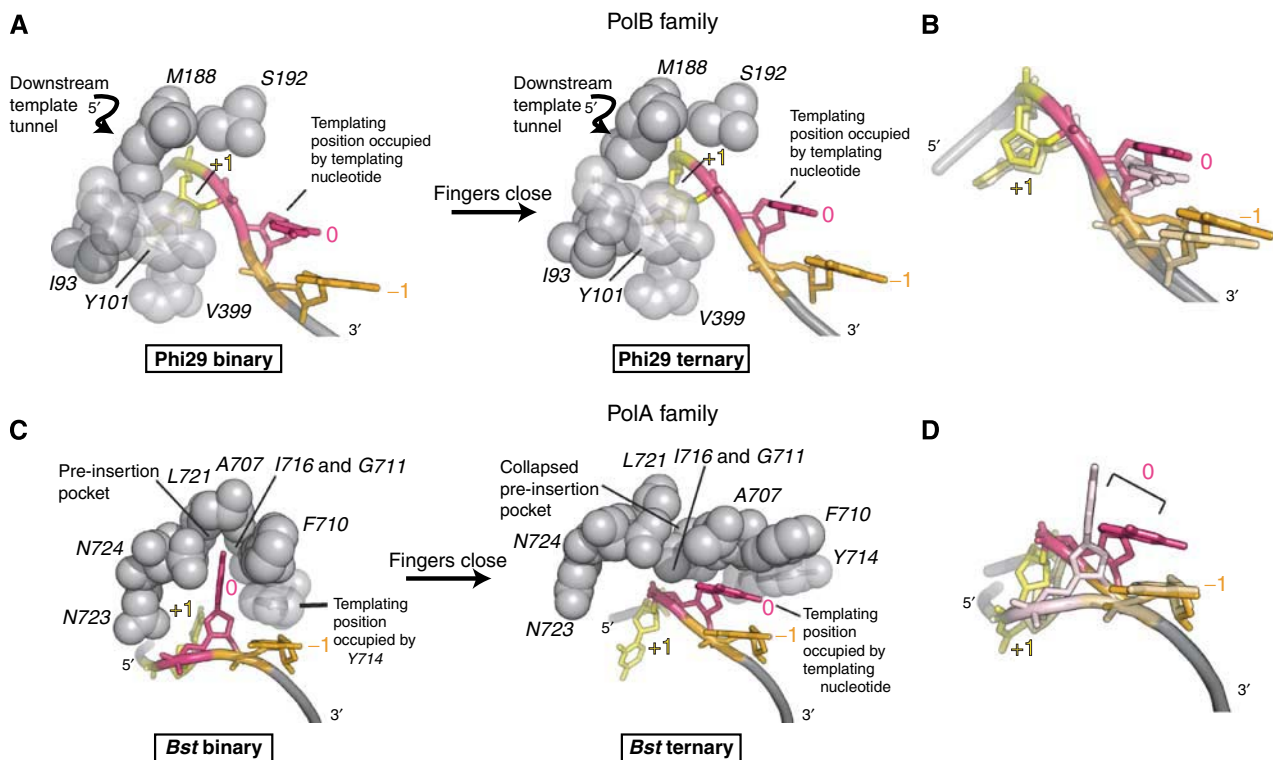


Figure 6 The kink between the +1 and templating nucleotides is a common feature in A- and B-family polymerases. The structures of the *B. stearothermophilus* DNAPs are from 1L3S (binary) and 1L5V (ternary). For consistency, all numbers refer to the templating positions before nucleotide addition. In all panels, the template nucleotide of the nascent base pair is shown in orange (–1), the next templating nucleotide in pink (0), and the one 5' to it in yellow (+1). All residues are shown as spheres, except for the conserved A-family tyrosine (714 in *B. stearothermophilus* DNAP), which is shown as sticks inside spheres to emphasize its movement, and Y101 from phi29 DNAP shown as sticks inside spheres for clarity. (A) When the fingers are opened in a B-family DNAP, the +1 nucleotide is stabilized by nonconserved hydrophobic interactions, and the templating site is occupied by the templating nucleotide from the last round of incorporation. When the fingers close, no significant movements within the DNA occur. (B) A superposition of the DNA from the binary and ternary complexes of phi29 DNAP. The DNA from the binary complex is colored in lighter colors, whereas that of the ternary complex is in darker colors. There is little difference in the overall positioning of the bases. (C) In the A-family, when the fingers are opened, a conserved tyrosine residue occupies the insertion site. The next templating nucleotide is stabilized in a pre-insertion site of hydrophobic residues from the base of the fingers. When the fingers close, the tyrosine moves out of the templating site, and the templating base moves into the templating site, stacking on the template strand of the upstream duplex. (D) A superposition of the DNA from the binary and ternary complexes of *B. stearothermophilus* DNAP. The DNA is colored as in (B). The nucleotide in the 0 position is in a very different location in the two complexes.

templating base from the pre-insertion site into the templating site on nucleotide binding by the Klenow fragment of DNAP I (Purohit *et al*, 2003). In contrast, the movements in the B-family are much more subtle and the dislocated base is not stabilized within a pre-insertion site that collapses on fingers closing. Because B-family DNAPs have no residue homologous to the A-family tyrosine that blocks the templating site, the next templating base always occupies the templating site (Figure 6A).

The different interactions that A- and B-family polymerases make with the single-stranded downstream template overhang lead to distinct structural mechanisms of translocation. Both of these mechanisms are based on the steric exclusion of the nascent base pair. In the A-family, the conserved tyrosine responsible for translocation occupies both the insertion and the templating sites; in the B-family, the conserved tyrosine residues involved in translocation occupy only the insertion site. Therefore, it appears that translocation in the B-family originates from the movement of the primer strand and the concomitant movement of the interacting template strand, whereas this movement in the A-family is centered at the nascent base pair.

Conformational equilibrium of the fingers subdomain and the mechanism of translocation

The fingers subdomains of replicative polymerases exist in opened and closed states, with the opened conformation dominating the equilibrium in the absence of incoming nucleotide substrate or pyrophosphate product. A plausible hypothesis for the ability of an incoming nucleotide to shift the equilibrium from the opened to closed conformation is that favorable interactions with both the basic residues in the fingers and the catalytic metal ions chelated by conserved residues of the palm subdomain are possible in the closed conformation, thereby enabling the phosphates of the nucleotide to electrostatically crosslink these two subdomains. After the chemistry step of the polymerization cycle, the dissociation of the pyrophosphate product releases the electrostatic crosslink stabilizing the closed conformation.

In both the A- and B-family polymerases, the conformational change of the fingers subdomain from opened to closed is necessary for the binding of the incoming nucleotide in the insertion site. In the B-family, as seen in the structures we have described above, the closing of the fingers subdomain moves the two conserved tyrosine side chains out of the insertion site into their positions in the nascent base pair-binding pocket. This contrasts with polymerases such as T7 RNAP, where the closing of the fingers moves a tyrosine conserved only within the A-family out of the insertion and templating sites, allowing the templating base to move into templating position and the incoming nucleotide to bind the insertion site.

Structural studies of T7 RNAP indicate that the active site residues and substrate remain in virtually the same positions in the post-insertion pre-translocation and post-chemistry pre-translocation states (Yin and Steitz, 2004). On pyrophosphate dissociation, the fingers subdomain opens, moving the aforementioned conserved tyrosine of the fingers subdomain 3.4 Å into the insertion and templating sites. The net consequence is the translocation of the nascent base pair out of the nascent base pair binding pocket into the -1 position (Yin and Steitz, 2004). From these studies, Yin and Steitz

(2004) predicted a similar mechanism for all polymerases that undergo a conformational change in response to nucleotide binding.

Assuming, as mutational data indicate, that the chemical step does not drastically alter the conformations of amino acids at the active site of B-family polymerases (Truniger *et al*, 2002), our structures of the binary and ternary complexes of phi29 DNAP provide a basis for proposing a structural mechanism of translocation for the B-family of polymerases. In this mechanism, in phi29 DNAP, as in T7 RNAP, the dissociation of the pyrophosphate product breaks the electrostatic link between the catalytic magnesium ions and the basic residues of the fingers subdomain that stabilizes and maintains the closed conformation of the fingers subdomain. When the fingers pivot to the opened position, the two conserved tyrosine residues, Y390 and Y254, enter the insertion site (Figure 2). Thus, in the opened conformation, the primer terminus can only be positioned in the post-translocation priming site, as the pre-translocation position is sterically inaccessible (Supplementary Movie). These observations are consistent with the biochemical data obtained with mutants at both tyrosine residues suggesting a direct or indirect interaction with the incoming dNTP (Blasco *et al*, 1992).

Whereas the principle of promoting translocation out of the nascent base pair-binding pocket by the steric exclusion of the nascent base pair seems to be a common theme among replicative polymerases of known structure, the residues involved are only conserved within polymerase families and not among different families. The residues involved in B-family polymerase translocation from sequence motifs 1 and 2a, Y254 and Y390, respectively, in phi29 DNAP, have previously been implicated in maintaining processive replication and in the binding of an incoming dNTP opposite a template base in B-family DNAPs from phi29 and RB69 (Blasco *et al*, 1992; Bonnin *et al*, 1999; Truniger *et al*, 1999; Yang *et al*, 1999, 2002a, 2005). Residue Y390 has no structural homolog in the A-family, whereas Y254, the steric gate residue, is an invariant glutamate in A-family polymerases (Astatke *et al*, 1998). The steric gate residue is not thought to play a role in translocation in the A-family (Yin and Steitz, 2004). Therefore, this study assigns an additional role for the steric gate residue in B-family polymerase translocation.

Although the structures of these binary and ternary substrate complexes with phi29 DNAP provide a structural basis for understanding the mechanism of translocation in B-family polymerases, they cannot, by themselves, address the thermodynamic and kinetic aspects of the translocation process. The structures are consistent with at least two possible kinetic schemes. If, for example, the diffusion of the primer terminus between the pre- and post-translocated positions is slower than the conformational change of the fingers subdomain, then a power stroke mechanism will dominate. However, if the diffusion of the DNA is faster than the conformational change of the fingers (including the changes in the tyrosine positions), a Brownian-ratchet model would better describe the translocation process. Whereas energetic conclusions drawn on the basis of disorder in crystal structures can be unreliable, the observation that two of the copies of primer-template DNA in the binary complex are ordered (Supplementary Figure S3a) and two are disordered is consis-

tent with the possibility that the primer-template diffuses along the helical axis.

Implications for protein priming

The sequence nonspecific interactions of phi29 DNAP with the ssDNA in the downstream template tunnel suggest that it alone cannot establish register of the template. In the context of protein priming, where the polymerase adds the first nucleotide to S232 of TP by pairing it with the second nucleotide from the 3' end of the template (Méndez *et al*, 1992), the priming domain of TP must position the template in the active site by sterically excluding it from the upstream duplex binding region of polymerase. This implies that steric exclusion and end recognition are responsible for the origin specificity.

Following incorporation of the first nucleotide, phi29 DNAP exhibits a sliding back motion (Méndez *et al*, 1992). In subsequent rounds of incorporation, this sliding back motion is exchanged for the typical processive, unidirectional movement of replicative polymerases. As described above, the position of Y226 in the *apo* structures clashes with the position of the -2 nucleotide of the template strand in the ternary complex. This supports mutational data implicating this residue in stable duplex binding at the active site (Truniger *et al*, 1996, 1999; Brenkman *et al*, 2001). These observations suggest that the rotameric change of Y226 may be important in the switch from an initially retrograde motion to the typical movement seen in processive DNAPs.

In summary, the structures presented here demonstrate how phi29 DNAP maintains sequence-nonspecific interactions with ssDNA and double-stranded DNA. They also allow us to define how DNA moves through the active site of a B-family DNAP. This mechanism of translocation is conceptually similar, but structurally distinct from that proposed for A-family polymerases. It assigns additional, unpredicted roles for the steric gate residue and a conserved tyrosine from the fingers subdomain in translocation, and has implications for the mechanism of protein-primed initiation.

Materials and methods

Proteins and oligonucleotides

Exonuclease-deficient phi29 DNAP (D12A/D66A) and phi29 TP were expressed and purified as described previously (Prieto *et al*, 1984; Lázaro *et al*, 1995) and stored at -80°C as ammonium sulfate pellets. The pellets were resuspended to ~15–18 mg/ml (polymerase) or ~8–10 mg/ml (TP) in 250 mM NaCl, 50 mM Tris-HCl (pH 7.5), 20 mM ammonium sulfate, and 10 mM MgCl₂ or 10 mM MnCl₂.

The oligonucleotides used are indicated in Table I. The primer:template substrates were annealed by incubation at 80°C (3 mM of each oligonucleotide in 20 mM Tris-HCl (pH 7.5), 10 mM NaCl) for 5 min followed by slow cooling for 2–12 h.

Sample preparation, crystallization, and stabilization

To form the binary complex, polymerase (12 mg/ml) and pre-annealed primer-template (0.6 mM) were incubated at 4°C for 30 min before stepwise dialysis from 250 to 25 mM NaCl in the presence of 10 mM MgCl₂. The mixture was then diluted to 10 mg/ml polymerase + 0.5 mM primer:template with buffer. Crystals were grown by vapor diffusion at 20°C by mixing equal parts of the

diluted incubation mixture and well solution. The binary crystals grew from a well solution of 100 mM Tris-HCl (pH 8.5), 20% PEG 10 000, 200 mM MgCl₂, were stabilized in the presence of 0.5 mM pre-annealed primer-template in 22% PEG 10 000, and cryoprotected by increasing the concentration of ethylene glycol to 20% before freezing in liquid propane.

To obtain the ternary complexes, polymerase and pre-annealed primer:template were incubated at 1.2X(12 mg/ml polymerase, 0.6 mM primer-template) for 30 min at 4°C before the stepwise dialysis down to 50 mM NaCl in the presence of 10 mM MnCl₂. After dialysis, the incubation mixtures were diluted to 1X(10 mg/ml polymerase, 0.5 mM primer-template).

For the ternary1 complex crystals, the 1X mixture of polymerase and primer:template was incubated with 5 mM ddATP for 90 min, followed by a 15 min incubation with 1 mM dTTP. These crystals grew from a well solution of 100 mM CHES (pH 9.5), 15–20% PEG 8000. Typically, 2 µl of the incubation reaction was mixed with an equal volume of well solution. They were stabilized in the presence of 1 mM dTTP in 22% PEG 8000, and cryoprotected by increasing the concentration of ethylene glycol in a stepwise manner to 30% before freezing in liquid propane.

To grow the ternary2 complex crystals, a 1X solution of polymerase and primer:template was incubated with 5 mM ddCTP for 90 min, followed by a 15-min incubation with 1 mM dGTP. These crystals grew from a well solution of 100 mM sodium acetate (pH 4.6), 200 mM ammonium acetate, 15% PEG 4000. Typically, 2 µl of the incubation reaction was mixed with an equal volume of well solution. The ternary2 crystals were stabilized in the presence of 1 mM dGTP in 22% PEG 4000, and cryoprotected in steps up to 25% ethylene glycol before freezing in liquid propane.

To obtain the ssDNA complex form, an equimolar ratio of polymerase (~8 mg/ml) and TP (~4 mg/ml) was dialyzed in buffer containing manganese down to 50 mM NaCl. This protein stock was diluted to 8.5 mg/ml final concentration of total protein and incubated with 1 mM template, 1 mM dATP, 1 mM dGTP, 5 mM ddTTP, 10 mM magnesium acetate, in a buffer of 20 mM Tris-HCl (pH 7.5), 20 mM ammonium acetate, 1 mM DTT, and 50 mM NaCl. These crystals grew from a 1:1 mixture of reaction incubation and well solution consisting of 0.1 M MES (pH 6.5), 12% PEG 20K. The high-resolution polymerase crystal was stabilized in 18% PEG 20K, and cryoprotected in steps up to 22.5% ethylene glycol before freezing in liquid propane.

Structure determination, refinement, and analysis

Diffraction data were integrated and scaled using the HKL software suite (Otwinowski and Minor, 1997). The structures of all complexes were solved by molecular replacement using the *apo* polymerase model (Kamtekar *et al*, 2004) without the fingers subdomain as the search model in the program PHASER (McCoy *et al*, 2005). The models were built using the programs O (Jones *et al*, 1991) and Coot (Emsley and Cowtan, 2004), and refined with the program REFMAC (Murshudov *et al*, 1997). Simulated annealing omit maps were calculated using CNS (Brunger *et al*, 1998). Structure superposition by least square fitting was performed in LSQKAB (Kabsch *et al*, 1976) or in LSQMAN (Kleywegt, 1999). Figures were made using Pymol (DeLano, 2002).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank the staff at NLSL beamline X25 and at APS beamline 24-ID. We thank Cathy Joyce for useful discussions and comments on the manuscript and Scott Bailey and Yong Xiong for help with data processing. We also thank the staff of the CSB Core computational facility. This work was supported by NIH grant GM57510 to TAS, grant BFU2005-00733 from the Spanish Ministry of Education and Science to MS and an institutional grant from Fundación Ramón Areces to the Centro de Biología Molecular Severo Ochoa.

References

- Astatke M, Ng K, Grindley ND, Joyce CM (1998) A single side chain prevents *Escherichia coli* DNA polymerase I (Klenow fragment) from incorporating ribonucleotides. *Proc Natl Acad Sci USA* **95**: 3402–3407
- Beese LS, Friedman JM, Steitz TA (1993) Crystal structures of the Klenow fragment of DNA polymerase I complexed with deoxynucleoside triphosphate and pyrophosphate. *Biochemistry* **32**: 14095–14101
- Blanco L, Salas M (1996) Relating structure to function in phi29 DNA polymerase. *J Biol Chem* **271**: 8509–8512
- Blasco MA, Lázaro JM, Bernad A, Blanco L, Salas M (1992) Phi 29 DNA polymerase active site. Mutants in conserved residues Tyr254 and Tyr390 are affected in dNTP binding. *J Biol Chem* **267**: 19427–19434
- Blasco MA, Lázaro JM, Blanco L, Salas M (1993) Phi 29 DNA polymerase active site. The conserved amino acid motif 'Kx3NSxYG' is involved in template-primer binding and dNTP selection. *J Biol Chem* **268**: 16763–16770
- Blasco MA, Méndez J, Lázaro JM, Blanco L, Salas M (1995) Primer terminus stabilization at the phi 29 DNA polymerase active site. Mutational analysis of conserved motif KXY. *J Biol Chem* **270**: 2735–2740
- Bonnin A, Lázaro JM, Blanco L, Salas M (1999) A single tyrosine prevents insertion of ribonucleotides in the eukaryotic-type phi29 DNA polymerase. *J Mol Biol* **290**: 241–251
- Brenkman AB, Heideman MR, Truniger V, Salas M, van der Vliet PC (2001) The (I/Y)XGG motif of adenovirus DNA polymerase affects template DNA binding and the transition from initiation to elongation. *J Biol Chem* **276**: 29846–29853
- Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D* **54**: 905–921
- Chelli R, Gervasio FL, Proccacci P, Schettino V (2002) Stacking and T-shape competition in aromatic-aromatic amino acid interactions. *J Am Chem Soc* **124**: 6133–6143
- DeLano WL (2002) *The PyMOL Molecular Graphics System*. San Carlos, CA, USA: DeLano Scientific
- Doublé S, Tabor S, Long AM, Richardson CC, Ellenberger T (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**: 251–258
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**: 2126–2132
- Franklin MC, Wang J, Steitz TA (2001) Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell* **105**: 657–667
- Freisinger E, Grollman AP, Miller H, Kisker C (2004) Lesion (intolerance) reveals insights into DNA replication fidelity. *EMBO J* **23**: 1494–1505
- Gao G, Orlova M, Georgiadis MM, Hendrickson WA, Goff SP (1997) Conferring RNA polymerase activity to a DNA polymerase: a single residue in reverse transcriptase controls substrate selection. *Proc Natl Acad Sci USA* **94**: 407–411
- Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**: 1876–1882
- Guajardo R, Sousa R (1997) A model for the mechanism of polymerase translocation. *J Mol Biol* **265**: 8–19
- Hanson J, Huxley HE (1955) The structural basis of contraction in striated muscle. *Symp Soc Exp Biol* **9**: 228–264
- Huang H, Chopra R, Verdine GL, Harrison SC (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**: 1669–1675
- Johnson SJ, Taylor JS, Beese LS (2003) Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc Natl Acad Sci USA* **100**: 3895–3900
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47** (Part 2): 110–119
- Kabsch W, Kabsch H, Eisenberg D (1976) Packing in a new crystalline form of glutamine synthetase from *Escherichia coli*. *J Mol Biol* **100**: 283–291
- Kamtekar S, Berman AJ, Wang J, Lázaro JM, de Vega M, Blanco L, Salas M, Steitz TA (2004) Insights into strand displacement and processivity from the crystal structure of the protein-primed DNA polymerase of bacteriophage phi29. *Mol Cell* **16**: 609–618
- Kamtekar S, Berman AJ, Wang J, Lázaro JM, de Vega M, Blanco L, Salas M, Steitz TA (2006) The phi29 DNA polymerase:protein-primer structure suggests a model for the initiation to elongation transition. *EMBO J* **25**: 1335–1343
- Kiefer JR, Mao C, Braman JC, Beese LS (1998) Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* **391**: 304–307
- Kleywegt GJ (1999) Experimental assessment of differences between related protein crystal structures. *Acta Crystallogr D* **55**: 1878–1884
- Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA (1992) Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**: 1783–1790
- Lázaro JM, Blanco L, Salas M (1995) Purification of bacteriophage phi 29 DNA polymerase. *Methods Enzymol* **262**: 42–49
- Li Y, Kong Y, Korolev S, Waksman G (1998a) Crystal structures of the Klenow fragment of *Thermus aquaticus* DNA polymerase I complexed with deoxyribonucleoside triphosphates. *Protein Sci* **7**: 1116–1123
- Li Y, Korolev S, Waksman G (1998b) Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J* **17**: 7514–7525
- McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D* **61**: 458–464
- Méndez J, Blanco L, Esteban JA, Bernad A, Salas M (1992) Initiation of phi 29 DNA replication occurs at the second 3' nucleotide of the linear template: a sliding-back mechanism for protein-primed DNA replication. *Proc Natl Acad Sci USA* **89**: 9579–9583
- Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53**: 240–255
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**: 307–326
- Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J (1994) Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* **264**: 1891–1903
- Pelletier H, Sawaya MR, Wolfle W, Wilson SH, Kraut J (1996) Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. *Biochemistry* **35**: 12742–12761
- Prieto I, Lázaro JM, Garcia JA, Hermoso JM, Salas M (1984) Purification in a functional form of the terminal protein of *Bacillus subtilis* phage phi 29. *Proc Natl Acad Sci USA* **81**: 1639–1643
- Purohit V, Grindley ND, Joyce CM (2003) Use of 2-aminopurine fluorescence to examine conformational changes during nucleotide incorporation by DNA polymerase I (Klenow fragment). *Biochemistry* **42**: 10200–10211
- Rodríguez I, Lázaro JM, Blanco L, Kamtekar S, Berman AJ, Wang J, Steitz TA, Salas M, de Vega M (2005) A specific subdomain in phi29 DNA polymerase confers both processivity and strand-displacement capacity. *Proc Natl Acad Sci U S A* **102**: 6407–6412
- Rothwell PJ, Waksman G (2005) Structure and mechanism of DNA polymerases. *Adv Protein Chem* **71**: 401–440
- Salas M (1991) Protein-priming of DNA replication. *Annu Rev Biochem* **60**: 39–71
- Saturno J, Lázaro JM, Esteban FJ, Blanco L, Salas M (1997) o29 DNA polymerase residue Lys383, invariant at motif B of DNA-dependent polymerases, is involved in dNTP binding. *J Mol Biol* **269**: 313–325
- Simon SM, Peskin CS, Oster GF (1992) What drives the translocation of proteins? *Proc Natl Acad Sci USA* **89**: 3770–3774
- Steitz TA, Smerdon S, Jager J, Wang J, Kohlstaedt LA, Friedman JM, Beese LS, Rice PA (1993) Two DNA polymerases: HIV reverse transcriptase and the Klenow fragment of *Escherichia coli* DNA polymerase I. *Cold Spring Harb Symp Quant Biol* **58**: 495–504

- Temiakov D, Patlan V, Anikin M, McAllister WT, Yokoyama S, Vassilyev DG (2004) Structural basis for substrate selection by τ 7 RNA polymerase. *Cell* **116**: 381–391
- Truniger V, Blanco L, Salas M (1999) Role of the 'YxGG/A' motif of Phi29 DNA polymerase in protein-primed replication. *J Mol Biol* **286**: 57–69
- Truniger V, Lázaro JM, Esteban FJ, Blanco L, Salas M (2002) A positively charged residue of phi29 DNA polymerase, highly conserved in DNA polymerases from families A and B, is involved in binding the incoming nucleotide. *Nucleic Acids Res* **30**: 1483–1492
- Truniger V, Lázaro JM, Salas M (2004) Two positively charged residues of phi29 DNA polymerase, conserved in protein-primed DNA polymerases, are involved in stabilisation of the incoming nucleotide. *J Mol Biol* **335**: 481–494
- Truniger V, Lázaro JM, Salas M, Blanco L (1996) A DNA binding motif coordinating synthesis and degradation in proofreading DNA polymerases. *EMBO J* **15**: 3430–3441
- Yang G, Franklin M, Li J, Lin TC, Konigsberg W (2002a) A conserved Tyr residue is required for sugar selectivity in a Pol alpha DNA polymerase. *Biochemistry* **41**: 10256–10261
- Yang G, Franklin M, Li J, Lin TC, Konigsberg W (2002b) Correlation of the kinetics of finger domain mutants in RB69 DNA polymerase with its structure. *Biochemistry* **41**: 2526–2534
- Yang G, Lin T, Karam J, Konigsberg WH (1999) Steady-state kinetic characterization of RB69 DNA polymerase mutants that affect dNTP incorporation. *Biochemistry* **38**: 8094–8101
- Yang G, Wang J, Konigsberg W (2005) Base selectivity is impaired by mutants that perturb hydrogen bonding networks in the RB69 DNA polymerase active site. *Biochemistry* **44**: 3338–3346
- Yin YW, Steitz TA (2004) The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* **116**: 393–404