A consensus view of DNA binding by the C family of replicative DNA polymerases

Meindert H. Lamers[a] and Mike O'Donnell[b,1]

[a]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720; and [b]Laboratory of DNA Replication, Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10021

The DNA polymerases that duplicate bacterial chromosomes are members of the C family of polymerases, which shares no sequence homology with any of the other DNA polymerase families. In Gram-negative bacteria the replicative polymerase is polymerase III α-subunit (Pol III), whereas in Gram-positive bacteria the replicative polymerase is called Pol C. Even though the first C family polymerase was discovered ≈40 years ago (1), it is the last polymerase family to be studied structurally. C family polymerases are quite large multidomain proteins, larger than polymerases of other families, which may explain why structural representatives of this class have been solved only recently. In this issue of PNAS, Evans *et al.* (2) present the structure of *Geobacillus kaustophilus* Pol C bound to a DNA substrate at 2.4-Å resolution. This structure provides a high-resolution analysis of a C family replicative DNA polymerase bound to DNA and dNTP. It is also the first structure of a replicative DNA polymerase from a Gram-positive bacterium and, as such, is an interesting drug target for new antibiotics against pathogenic bacteria such as *Staphylococcus aureus*, *Streptococcus pyogenes*, and *Bacillus anthracis*.

The *G. kaustophilus* Pol C, like the Pol III apo enzyme (no DNA or dNTP) structures from *Escherichia coli* (Eco) and *Thermus aquaticus* (Taq) (3, 4), has a folding pattern in the active site that is unlike that of the canonical polymerases such as Pol I and Pol II, and the eukaryotic replicative DNA polymerases Pol δ and Pol ε. Instead, the topology of the C family polymerases is homologous to that of Pol β, a low-fidelity polymerase involved in base excision repair. Pol β belongs to the X family of polymerases that are typically slow enzymes with low fidelity and low processivity in contrast to C family polymerases that are among the most rapid of all DNA polymerases (>700 nt/s) and display high fidelity and exceedingly high processivity (>50 kb). The processivity of C family polymerases is enabled by interaction with a ring-shaped protein, the β-clamp, that encircles DNA and slides along it, thereby acting as a mobile tether. Use of a ring-shaped protein generalizes to the replicases of eukaryotes and archae, which function with the ring-shaped proliferating cell nuclear antigen sliding clamp (PCNA).

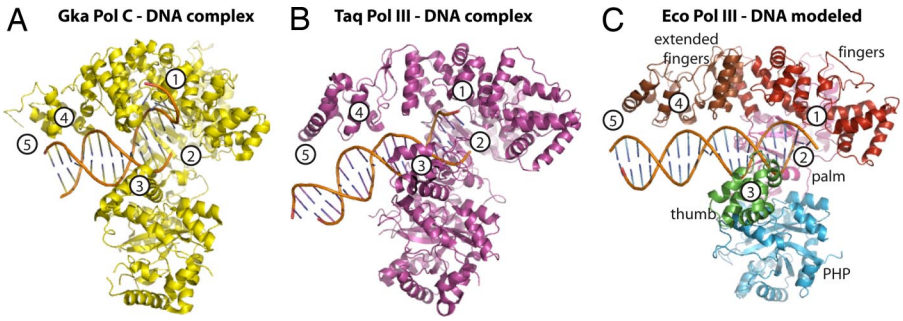

Fig. 1. Comparison of DNA binding in bacterial replicative DNA polymerases. (*A*) *G. kaustophilus* Pol C (2). (*B*) *T. aquaticus* Pol III (5). (*C*) *E. coli* Pol III with DNA modeled (3). N-terminal tail of Pol C and C-terminal tail in Taq Pol III not displayed. In *E. coli* Pol III the different domains are colored and labeled separately. The numbers indicate conserved features of DNA binding in Pol C and Pol III: ①, single-stranded template DNA enters the polymerase guided by the fingers domain; ②, polymerase active site; ③, the thumb domain contacts the DNA minor groove as it leaves the polymerase active site; ④, the extended fingers domain contacts the double-stranded region of the DNA; ⑤, binding site for the DNA sliding clamp.

The limited resolution (4.6 Å) of a recent cocrystal structure of Taq Pol III bound to DNA (5) did not permit a detailed evaluation of the interactions between the protein and DNA. The high-resolution structure of *G. kaustophilus* Pol C bound to DNA provides an important improvement over the Taq-DNA structure, because it allows detailed analysis of DNA binding by bacterial replicative DNA polymerases. Comparison of the Pol C-DNA structure with the Taq Pol III-DNA structure and a previously proposed model for DNA binding in *E. coli* Pol III (3) shows that DNA binding is very similar between the different homologs (see Fig. 1) but in addition reveals a remarkable and unexpected degree of conformational flexibility in the polymerase active site associated with DNA binding. In fact, it appears that the active conformation of the catalytic site may not form until the substrate has bound: two of the three absolutely conserved catalytic residues move towards the DNA substrate upon binding, as well as two positively charged residues that move inwards to contact the phosphate tail of the incoming nucleotide. These observations suggest that distortions observed in the active site of the Eco Pol III apo structure may be explained by the absence of bound substrate, needed to elicit the active configuration. Furthermore, the conformational flexibility observed between the different structures extends well beyond the active site, to the extended fingers domain that mediate much of the contacts to DNA. Why would a DNA polymerase undergo these types of conformational changes? Evans *et al.* (2) present an exciting possibility. The flexibility may be part of the switching of the DNA template from the polymerase active site to the exonuclease active site after misincorporation of a wrong nucleotide. They also point out that this conformational flexibility may help replicating machineries in yet another way. Specifically, cells contain "translesion DNA polymerases" that extend DNA over a damaged nucleotide base (6). When the replicative polymerase encounters a damaged base it will stall and a translesion polymerase will be needed for continued synthesis. Translesion DNA polymerases are known to bind the same sliding clamp to which the replicative polymerase is attached, and Evans *et al.* (2) predict that the conformational flexibility inherent in C family polymerases facilitates "polymerase switching" on the sliding

Author contributions: M.H.L and M.O. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 20695.

[1]To whom correspondence should be addressed. E-mail: odonnel@mail.rockefeller.edu.

clamp during translesion synthesis of DNA across chemically-modified bases.

The structure of *G. kaustophilus* Pol C, and its comparison with the Eco and Taq Pol III structures, also helps to resolve a conflicting issue raised by the earlier structure of Taq Pol III (5). In the DNA-bound Pol C structure of Evans *et al.* (2) the planes of the DNA bases are out of alignment with the β-strands of the polymerase active site by $\approx$30°. In the Taq Pol III ternary structure, the planes of the DNA bases are positioned parallel to the active-site β-strands, similar to the way DNA is bound in Pol β (7). In the original model proposed for *E. coli* (3), however, a parallel alignment was not possible because it resulted in several clashes of the DNA with other regions of the polymerase (i.e., the thumb and PHP domains; see Fig. 1). As a result, the DNA was modeled such that the base planes and the active-site β-strands were arranged at an angle of $\approx$30°. Therefore, the Pol C ternary complex structure and Eco Pol III model reveal that the orientation of the base planes at an angle to the β-sheets is accommodated, whereas the Taq Pol III ternary complex structure demonstrates a parallel positioning of the base planes and β-strands [see also figure 4 in Evans *et al*. (2)]. These different conformations of enzyme–substrate complexes may reflect different steps during DNA catalysis or instead may be explained by different methods of DNA binding by C family polymerases of different species.

The recent Pol III and Pol C structures also leave many questions unanswered. One outstanding feature of Pol III is its high speed, not achieved by any other polymerase. In particular, it is puzzling that the active site of Pol III/Pol C is homologous to that of X family polymerases that are characterized by a slow catalytic rate of synthesis. What accounts for the vast difference in rate between C-family and X-family polymerases? The coordination of the active-site residues, magnesium ion, and incoming nucleotide is nearly identical between Pol C and Pol β and therefore would not appear to explain the dramatic difference in activity between the 2 enzymes. One of the striking and unique features of Pol III and Pol C is the extended fingers domain, which is much larger than in any other known polymerase, and may contribute to their rapid catalytic rate. Future studies will be needed to determine whether this may be the case.

Different conformations of enzyme–substrate complexes may reflect different steps during DNA catalysis.

Biochemical studies show that the polymerase binds to the β-sliding clamp through a β-binding peptide sequence at the end of the extended fingers domain, and its position has been modeled similarly in the structure of Eco and Taq Pol III and *G. kaustophilus* Pol C. Future structural studies that include the clamp along with the polymerase are needed to understand their detailed arrangement on DNA and whether they alter their orientation during the catalytic cycle or upon encountering a template lesion. Sliding clamps are homooligomeric rings and therefore contain binding sites for more than 1 DNA polymerase. In fact, 2 different DNA polymerases have been demonstrated to bind 1 clamp simultaneously (8). It would be fascinating to see the structure of binary polymerase/sliding clamp complexes such as this. The DNA polymerase also contacts the clamp loader complex, a pentameric assembly that couples ATP hydrolysis to open and closed sliding clamps onto primed sites. These clamp loaders bind multiple copies of the C family replicative polymerase for simultaneous replication of both strands of duplex DNA. Structures that reveal how multiple C family polymerases are juxtaposed on the clamp loader complex may reveal new secrets at the heart of the replication machine.

The structure of *G. kaustophilus* Pol C bound to DNA described in this issue of PNAS (2) is an exciting and important step forward, but clearly many questions remain to be answered. Despite the fact that Pol III was discovered $\approx$40 years ago (1), its future is looking more exciting than ever before.

ACKNOWLEDGMENTS. We thank Ronald J. Evans, Douglas R. Davies, James M. Bullard, Jeffrey Christensen, Louis S. Green, Joseph W. Guiles, Janice D. Pata, Wendy K. Ribble, Nebojsa Janjic, and Thale C. Jarvis for providing the coordinates of the Pol C-DNA structure. This work was supported in part by National Institutes of Health Grants GM45547 and GM38839.

1. Kornberg T, Gefter ML (1971) Purification and DNA synthesis in cell-free extracts: Properties of DNA polymerase II. *Proc Natl Acad Sci USA* 68:761–764.
2. Evans RJ, *et al.* (2008) Structure of PolC reveals unique DNA binding and fidelity determinants. *Proc Natl Acad Sci USA* 105:20695–20700.
3. Lamers MH, Georgescu RE, Lee SG, O'Donnell M, Kuriyan J (2006) Crystal structure of the catalytic α-subunit of *E. coli* replicative DNA polymerase III. *Cell* 126:881–892.
4. Bailey S, Wing RA, Steitz TA (2006) The structure of *T. aquaticus* DNA polymerase III is distinct from eukaryotic replicative DNA polymerases. *Cell* 126:893–904.
5. Wing RA, Bailey S, Steitz TA (2008) Insights into the replisome from the structure of a ternary complex of the DNA polymerase III α-subunit. *J Mol Biol* 382:859–869.
6. Jarosz DF, Beuning PJ, Cohen SE, Walker GC (2007) Y-family DNA polymerases in *Escherichia coli*. *Trends Microbiol* 15:70–77.
7. 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.
8. Indiani C, McInerney P, Georgescu R, Goodman MF, O'Donnell M (2005) A sliding-clamp tool belt binds high- and low-fidelity DNA polymerases simultaneously. *Mol Cell* 19:805–815.