© 2001 Nature Publishing Group http://structbio.nature.com

The anatomy of infidelity

Tom Ellenberger and Laura F. Silvian

Recent crystal structures of error prone DNA polymerases that bypass damage in DNA templates provide counterexamples to high fidelity polymerases.

© 2001 Nature Publishing Group http://structbio.nature.com

The number of DNA polymerases skyrocketed recently with the discovery of a very large group of error prone DNA polymerases, the Y-superfamily[1], capable of waltzing past DNA lesions that trip up normal replicative polymerases. Structures of two of these remarkable enzymes were recently reported. The crystal structures of the *Saccharomyces cerevisiae* RAD30A polymerase (Pol η)[2] and the *Sulfolobus solfataricus* Dbh polymerase[3] provide the initial glimpses of Y-superfamily polymerases that restart replication at sites of mismatched or damaged bases[4–9] and that contribute to adaptive mutagenesis in cells growing under adverse conditions[10,11]. Error prone polymerases allow cells to cope with unrepaired DNA damage by enabling the completion of replication in the face of otherwise terminal roadblocks. These 'sloppier copiers' lack many of the virtues of other polymerases, including speed, fidelity and a firm grip (processivity) on the DNA template. However, the biological fitness of the error prone DNA polymerases should be judged by a different standard. They are, after all, specialists in resolving messy situations that make respectable polymerases blush.

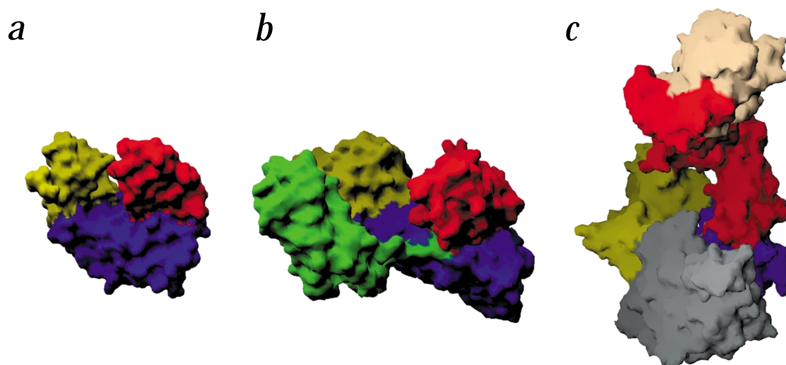

Fig. 1 Comparison of error prone and high fidelity polymerases. There is no sequence similarity between the error prone DNA polymerases, represented here by *Sulfolobus solfataricus* Dbh[3] and *Saccharomyces cerevisiae* Pol η[2], and highly accurate replicative polymerases like the bacteriophage T7 DNA polymerase[31]. Nonetheless, all three polymerases have similar shapes, resembling a right hand composed of fingers (yellow), palm (blue) and thumb (red) subdomains. Dbh and Pol η have finger and thumb domains that are smaller than those of T7 DNA polymerase. This creates a shallow and unencumbered active site in the error prone polymerases. The surfaces of all three enzymes are shown from the vantage point of DNA exiting the polymerase. ***a***, The *S. solfataricus* Dbh catalytic fragment (residues 1–205, PDB accession 1IM4). ***b***, The *S. cerevisiae* pol η polymerase fragment (residues 1–509, PDB accession 1JIH) contains an additional C-terminal polymerase associated domain (green) that is proposed to contact DNA. ***c***, The T7-DNA polymerase (PDB accession 1T7P) is composed of palm, fingers and thumb subdomains with additional proofreading exonuclease (gray) and processivity domains (tan). The extended thumb of T7 DNA polymerase and its more expansive fingers create an enclosed substrate binding site that undoubtably contributes to the high fidelity of this polymerase. The DNA and nucleotide substrates have been omitted from the T7 structure in order to show the active site.

It was unforeseen that cells would be equipped with such a vast number of polymerases to deal with replication mistakes and chemical insults to DNA. The sheer number of lesion bypass polymerases, and in particular the existence of many orthologs in higher organisms, suggests that individual enzymes have highly specialized cellular roles. Additional support for this proposal comes from the unique biochemical characteristics of various Y-superfamily polymerases, including their different efficiencies in bypassing particular lesions and their mutagenic propensities[12–18]. The broad phylogenetic distribution of the lesion bypass polymerases underscores the strategic importance of tolerating DNA damage by replication bypass as a means of survival when DNA repair cannot be completed in a timely manner. Although we do not fully understand the molecular logic behind the decision to either repair a DNA lesion or replicate past it, both options appear to be important. Human patients with a variant form of the inherited disorder xeroderma pigmentosum (XP-V) lack functional Pol η, resulting in the inability to bypass several types of UV-induced lesions in DNA[19]. Although cells from XP-V patients have apparently normal nucleotide excision repair activities that should repair these lesions, they are unable to replicate UV-damaged DNA. This loss of lesion bypass activity is associated with a high incidence of sunlight-induced cancers in XP-V patients.

Almost 40 years elapsed between the identification of the first DNA polymerase[20] and the discovery of the error prone polymerases[12,19,21–23], which lack the conserved sequence motifs characterizing the DNA polymerases that handle most genomic replication and repair[24,25]. Residues within these conserved motifs interact with bound substrates and have profound effects on the rate and fidelity of DNA synthesis by most polymerases[26–28]. Error prone DNA polymerases have a different set of five sequence motifs (I–V)[6,7,13,29], which are broadly conserved in more than 50 members of the Y-polymerase superfamily[1]. Amino acid substitutions at several of these conserved positions eliminate or greatly diminish polymerase activity and the effects of many others remain to be tested. The crystal structures of Pol η[2] and Dbh[3] reveal the locations of the conserved motifs and suggest what their functions may be, providing a Rosetta Stone for ciphering how DNA lesions are skirted and why these polymerases make so many mistakes.

Like conventional DNA polymerases, Pol η and Dbh have a shape resembling a right hand with fingers, palm, and thumb subdomains[26,28]. The palms of Pol η and Dbh closely resemble the palm of A-family polymerases, such as *Escherichia coli* DNA polymerase I[30] and the bacteriophage T7 DNA polymerase[31]. The palm comprises the floor of the polymerase active site and includes three highly con-

© 2001 Nature Publishing Group http://structbio.nature.com

© 2001 Nature Publishing Group http://structbio.nature.com

served acidic residues that are likely to chelate two catalytically important metals, as in other polymerases that use a two-metal ion mechanism[32]. Other residues from conserved motifs I–V of the Y-polymerase superfamily decorate the surfaces of the cleft between the fingers and thumb. The fingers and thumb of the error prone DNA polymerases are small in comparison to those of other DNA polymerases (Fig. 1). These appendages grip the DNA in other polymerases[27,28]. Their diminutive size in the error prone polymerases may result in fewer interactions with the DNA and perhaps explain the weak, distributive polymerase activity that is characteristic of these enzymes. The structure of Pol η reveals an additional C-terminal domain located opposite the thumb (Fig. 1). This polymerase-associated domain (PAD) is a mixed α/β structure forming one side of a groove that is proposed to bind the DNA product of the polymerase[2]. The shape and dimensions of the PAD match those of the major groove, explaining how it could bind to DNA. The Dbh fragment that was crystallized[3] comprises the fingers, palm and thumb, and it is missing the C-terminal region of Dbh that might be analogous to the PAD of Pol η. The C-terminally truncated Dbh is much less processive than the full-length polymerase[3], demonstrating that the C-terminal region contributes DNA binding function(s).

Sloppiness is a virtue of the error prone DNA polymerases[7]. By playing fast-and-loose with the DNA template, these enzymes can replicate past unrepaired DNA damage or mismatched bases that otherwise bring replication to a standstill. This tolerance of distorted DNA templates comes at a cost[4] — most lesion bypass enzymes studied to date have notoriously high error rates when copying undamaged DNA. Fortunately, these error prone polymerases bind weakly to DNA, incorporating only a few nucleotides before falling off the DNA template[14,18,22,33]. This distributive behavior probably limits their activity *in vivo*, providing the opportunity for a processive and highly accurate replicative polymerase to resume synthesis shortly after a lesion is bypassed. The error prone polymerases also lack intrinsic proofreading activity so that nucleotides inserted opposite a lesioned base are not immediately excised in a futile cycle of synthesis and proofreading.

A comparison of the crystal structures of Pol η and Dbh to those of polymerases with high fidelity suggests a basis for the relaxed templating requirements of the lesion bypass polymerases. The fidelity of template-dependent DNA polymerases is thought to result from a 'geometric selection' imposed by the polymerase active site to accommodate only the correct nucleotide matching the template[34,35]. Geometric selection places strict constraints on the types of base pairs that can be accommodated in the polymerase active site; consequently, bulky adducts and severely distorting chemical modifications are excluded. A substrate-induced change in the conformation of the fingers may further limit the choice of nucleotide substrates[28]. It seems that the ability to bypass chemical lesions is inextricably linked with a low fidelity of DNA synthesis[16]. Replicative DNA polymerases are far more selective than lesion bypass enzymes, and they are unable to transit past many types of chemically damaged bases.

In contrast to higher fidelity polymerases, the fingers of both Pol η and Dbh are small. This gives their active sites an open appearance that is suggestive of fewer interactions with DNA and nucleotide substrates, and less restrictions on mispairing. The fingers of Pol η and Dbh lack the highly conserved residues that make intimate contact with DNA bases in other polymerases[27,28]. Bulky lesions and other distortions of the DNA template might be less encumbered by the small fingers of Pol η and Dbh. The shallowness of the substrate binding pockets of Pol η and Dbh also suggests that water might not be effectively excluded from the active site, depriving the polymerase of an important source of discrimination against mismatched base pairs[35,36]. The high error rates of these polymerases are consistent with a small energetic difference between correctly paired and mispaired substrates bound to their active sites, suggesting they operate with little or no geometric selection. It remains to be seen whether or not the fingers of lesion bypass polymerases change conformation in response to substrate binding. Perhaps these enzymes lack an induced fit mechanism[27,37] for selecting nucleotides that match the DNA template, as was suggested by Steitz and coworkers[3]. It will be fascinating to compare the current structures of Pol η and Dbh with the anticipated structures of other promiscuous DNA polymerases, and to tease out the basis of their unique bypass activities and their mutagenic signatures.

Tom Ellenberger and Laura Silvian are in the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA. Correspondence should be addressed to T.E. email: tome@hms.harvard.edu or L.F.S. email: lsilvian@hms.harvard.edu.

1. Ohmori, H. *et al. Mol. Cell* **8**, 7–8 (2001).
2. Trincao, J. *et al. Mol. Cell* **8**, 417–426 (2001).
3. Zhou, B.-L., Pata, J.D. & Steitz, T.A. *Mol. Cell* **8**, 427–437 (2001).
4. Bridges, B.A. *Curr. Biol.* **8**, R886–R888 (1998).
5. Friedberg, E.C. & Gerlach, V.L. *Cell* **98**, 413–416 (1999).
6. Woodgate, R. *Genes Dev.* **13**, 2191–2195 (1999).
7. Goodman, M.F. & Tippin, B. *Curr. Opin. Genet. Dev.* **10**, 162–168 (2000).
8. Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S. & Prakash, L. *Nature* **406**, 1015–1019. (2000).
9. Cox, M.M. *et al. Nature* **404**, 37–41 (2000).
10. Foster, P.L. *Cold Spring Harbor Symposia on Quantitative Biology* **65**, 21–29 (2000).
11. McKenzie, G.J., Lee, P.L., Lombardo, M., Hastings, P.J. & Rosenberg, S.M. *Mol. Cell* **7**, 571–579. (2001).
12. Tang, M. *et al. Proc. Natl. Acad. Sci. USA* **96**, 8919–8924 (1999).
13. Reuven, N.B., Arad, G., Maor-Shoshani, A. & Livneh, Z. *J. Biol. Chem.* **274**, 31763–31766 (1999).
14. Johnson, R.E., Washington, M.T., Prakash, S. & Prakash, L. *J. Biol. Chem.* **275**, 7447–7450 (2000).
15. Washington, M.T., Johnson, R.E., Prakash, S. & Prakash, L. *J. Biol. Chem.* **274**, 36835–36838 (1999).
16. Matsuda, T., Bebenek, K., Masutani, C., Hanaoka, F. & Kunkel, T.A. *Nature* **404**, 1011–1013 (2000).
17. Ohashi, E. *et al. J. Biol. Chem.* **275**, 39678–39684 (2000).
18. Tang, M. *et al. Nature* **404**, 1014–1018 (2000).
19. Masutani, C. *et al. Nature* **399**, 700–704 (1999).
20. Lehman, I.R., Bessman, M.J., Simms, E.S. & Kornberg, A. *J. Biol. Chem.* **233**, 163–170. (1958).
21. Nelson, J.R., Lawrence, C.W. & Hinkle, D.C. *Nature* **382**, 729–731 (1996).
22. Wagner, J. *et al. Mol. Cell* **4**, 281–286 (1999).
23. Johnson, R.E., Prakash, S. & Prakash, L. *J. Biol. Chem.* **274**, 15975–15977 (1999).
24. Delarue, M., Poch, O., Tordo, N., Moras, D. & Argos, P. *Protein Eng.* **3**, 461–467 (1990).
25. Braithwaite, D.K. & Ito, J. *Nucleic Acids Res.* **21**, 787–802 (1993).
26. Joyce, C.M. & Steitz, T.A. *Annu. Rev. Biochem.* **63**, 777–822 (1994).
27. Kunkel, T.A. & Bebenek, K. *Annu. Rev. Biochem.* **69**, 497–529 (2000).
28. Patel, P.H. & Loeb, L.A. *Nature Struct. Biol.* **8**, 656–659 (2001).
29. Johnson, R.E., Prakash, S. & Prakash, L. *Proc. Natl. Acad. Sci. USA* **97**, 3838–3843 (2000).
30. Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. & Steitz, T.A. *Nature* **313**, 762–766. (1985).
31. Doublie, S., Tabor, S., Long, A.M., Richardson, C.C. & Ellenberger, T. *Nature* **391**, 251–258 (1998).
32. Steitz, T.A. *J. Biol. Chem.* **274**, 17395–17398 (1999).
33. Tang, M. *et al. Proc. Natl. Acad. Sci. USA* **95**, 9755–9760 (1998).
34. Goodman, M.F. *Proc. Natl. Acad. Sci. USA* **94**, 10493–10495 (1997).
35. Kool, E.T. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 1–22 (2001).
36. Petruska, J., Sowers, L.C. & Goodman, M.F. *Proc. Natl. Acad. Sci. USA* **83**, 1559–1562 (1986).
37. Johnson, K.A. *Annu. Rev. Biochem.* **62**, 685–713 (1993).