

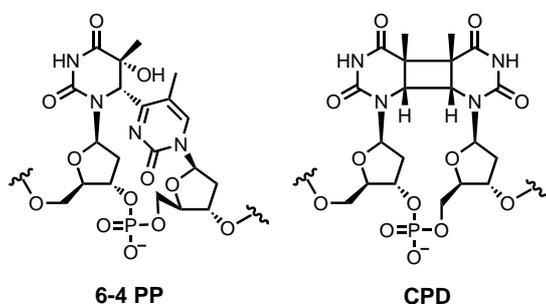
Molecular Insights into Error-Prone DNA Replication and Error-Free Lesion Bypass

Andreas Marx* and Daniel Summerer^[a]

KEYWORDS:

DNA damage · DNA polymerases · DNA replication · thymidine dimers · UV light

Ultraviolet (UV) light causes a variety of damage in DNA. The most abundant lesions are pyrimidine dimers such as the pyrimidine (6-4) pyrimidone photoproduct (6-4PP) and the *cis,syn* cyclobutane–pyrimidine dimer (CPD)^[1] (Scheme 1). These lesions are often repaired sluggishly and remain in the DNA,



Scheme 1. UV-light-induced thymidine dimer DNA lesions. 6-4PP = pyrimidine (6-4) pyrimidone photoproduct, CPD = *cis,syn* cyclobutane pyrimidine dimer.

where they considerably impede the DNA replication machinery.^[1] How cells perform DNA synthesis past these kind of lesions has for a long time remained obscure. Only recently, it has been discovered that several specialized DNA polymerases which belong to the new superfamily Y are clearly involved in translesion synthesis.^[2] These DNA polymerases (pols) are found in a variety of organisms ranging from *Escherichia coli* to humans. One of their most prominent functional characteristics is their high

error propensity when dealing with undamaged DNA, which distinguishes them from known high-fidelity DNA polymerases. Most of their biological roles and functions have not been elucidated yet. Nevertheless, recent investigations suggest the involvement of pol ι together with pol ζ in error-prone translesion DNA synthesis.^[3] Furthermore, it has been shown that pol ι possesses lyase activity, which suggests that the enzyme may be involved in DNA repair processes.^[4] DNA polymerase η is unique among the eukaryotic DNA polymerases since it has the ability to perform error-free bypass synthesis through CPD lesions, which are poorly removed by repair processes.^[5] It has been shown that cells from patients with the inherited disorder of a variant form of xeroderma pigmentosum (XP-V) are unable to perform CPD bypass synthesis as a result of expression of severely truncated or functionally inactive forms of pol η .^[5e, f] These observations strongly suggest that pol η is the XP-V gene product and that the polymerase is involved in suppression of sunlight-induced skin cancer. Yeast and human pol η replicate through a *cis,syn* thymidine–thymidine dimer by inserting two deoxyadenosine residues opposite the lesion. Although, pol η misinserts noncomplementary nucleotides opposite the first thymidine (T) in the lesion with high frequency, extension from mismatched substrates is greatly impeded. Presumably, this feature allows the polymerase to dissociate from the primer template complex and recruit exonucleases to remove the mispaired nucleotide. Overall, the UV-damaged lesion would be more selectively bypassed than

predicted from the misinsertion frequency alone. In general, pol η appears to be able to replicate DNA accurately through the lesions that are most frequently formed as a result of exposure to UV light.

These unique properties of pol η instantly triggered interest in its molecular and structural basis. Recently, several crystal structures of error-prone Y-family DNA polymerases have been solved from which valuable new insights into their molecular action can be drawn.^[6] Trincão et al. determined the crystal structure of an N-terminal fragment of *Saccharomyces cerevisiae* pol η that includes the active site.^[6c] This structure reveals that pol η retains the overall shape of a right hand composed of thumb, palm, and finger domains, as seen in other known DNA polymerases^[6–9] (Figure 1). However, pol η contains a novel polymerase-associated domain (PAD), which mimics an extra set of fingers. The structure shows that the palm domain of pol η is nearly superimposable with that of high-fidelity DNA polymerases.^[7, 8] In contrast, the finger and thumb domains, which make extensive contacts with the primer template and incoming triphosphate in other DNA polymerases, are unusually small and stubby. In particular, the “O helices”, which are believed to play a pivotal role in fidelity checking, are absent in pol η . Further insights into error-prone DNA synthesis were gained through the crystal structures of *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4) reported by Ling et al.^[6d] This enzyme is able to bypass *cis-syn* CPD lesions efficiently and thus, with respect to translesion synthesis, the enzymatic properties of Dpo4 are similar to that of the distantly related eukaryotic pol η . Ling et al. successfully obtained Dpo4 structures in ternary complexes with the DNA primer template and either a canonical or noncanonical incoming nucleoside triphosphate. The overall structure of the Dpo4–DNA ternary complex showed the familiar shape of a half-

[a] Dr. A. Marx, Dipl.-Chem. D. Summerer
Kekulé-Institut für Organische Chemie
und Biochemie
Universität Bonn
Gerhard-Domagk-Strasse 1
53121 Bonn (Germany)
Fax: (+49) 228-73-5388
E-mail: a.marx@uni-bonn.de

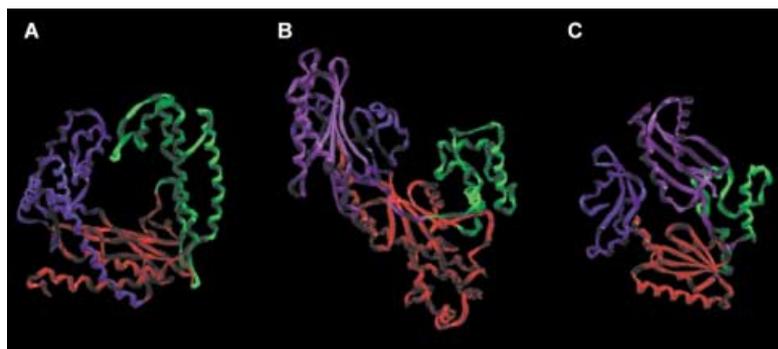


Figure 1. DNA polymerase structures. A: ternary complex of Taq DNA polymerase (Protein Data Bank (PDB) entry code 2KTQ); B: DNA polymerase η (coordinates provided by the authors); C: the ternary complex of Dpo4 (PDB entry code 1JX4). The palm (red), fingers (blue), thumb (green), and polymerase-associated domain/little finger (purple) are shown.

open right hand found in several DNA polymerases in which parts of the primer template are embedded.^[7, 8] Analogous to pol η , Dpo4 contains a domain termed the little finger in addition to the finger, palm, and thumb domains. Again, the palm domain is structurally similar to those of other DNA polymerases and the essential carboxylates in Dpo4 are in identical positions to those in T7, Klenow, *Bacillus*, Taq, and RB69 DNA polymerases, as well as in HIV reverse transcriptase.^[8] However, the finger and thumb domains of Dpo4 that surround the incoming triphosphate and template nucleotide are unusually small. The O helices that are believed to be involved in selection for the right nucleobase geometry and are present in all high-fidelity DNA polymerases are absent in Dpo4 as well as in pol η .^[7] Overall, the nascent base pair between the template and incoming nucleotide in Dpo4 is less tightly surrounded in the vicinity of the enzyme than it is in high-fidelity DNA polymerases. The open and solvent-accessible active site may be one structural basis for the error-prone replication of undamaged DNA by Y-family DNA polymerases. These findings further support the current concept that geometrical constraints are the major causative of DNA polymerase fidelity.^[10]

The primer template complexed by Dpo4 is nearly standard B-form DNA and undergoes protein interactions with eight base pairs through both hydrogen bonds and van der Waals contacts. These interactions are mostly with the phosphodiester and sugar moieties. In contrast to high-fidelity DNA polymerases, no hydrogen

bonds with donor or acceptor sites of the nucleobases through the minor groove are observed in Dpo4. Checking of hydrogen-bonding patterns by high-fidelity DNA polymerases is believed to contribute to fidelity through prevention of inadvertent sealing of mismatched base pairs into the nascent DNA.^[7c, 10]

The second crystal structure obtained by Ling et al. contains a mismatched dideoxyguanosine triphosphate (ddGTP) in the active site of Dpo4.^[6d] In the active site, conformations of the sugar phosphate moieties of the primer, template, and nucleoside triphosphate differ significantly from that found when a canonical nucleotide is bound. Translocation of the template without replication of the first template base (guanine) leads the incoming ddGTP to form a canonical base pair with the next template base (cytosine; Figure 2). This structure shows that the

active site of Dpo4 is able to admit two adjacent template bases simultaneously; this suggests possible mechanisms for thymidine dimer translesion synthesis. However, since the second template base can form a canonical base pair with the incoming nucleotide, it is unclear what structure would result if the second template base were noncanonical. Structural data for such DNA polymerases bound to DNA that includes a CPD lesion would be extremely helpful to further clarify this point.

In conclusion, several novel DNA polymerases with new functions and properties have been discovered during the past few years. Most of their functional properties and roles in which they interplay with other enzymes and are involved in replication, repair, or other fundamental cell processes remain to be elucidated before the question of why there are so many DNA polymerases and how they are recruited at the exact moment required can finally be answered. On the molecular level recent functional and structural investigations of "novel" DNA polymerases have added significantly to our molecular understanding of how cells deal with damaged DNA in translesion synthesis. The structural data suggest less tight and more open active sites as a putative reason for the error propensity observed in low-fidelity DNA polymerases. We await further multidisciplinary approaches that should lead to a wealth of new insights and significantly contribute to our understanding of these complex processes.

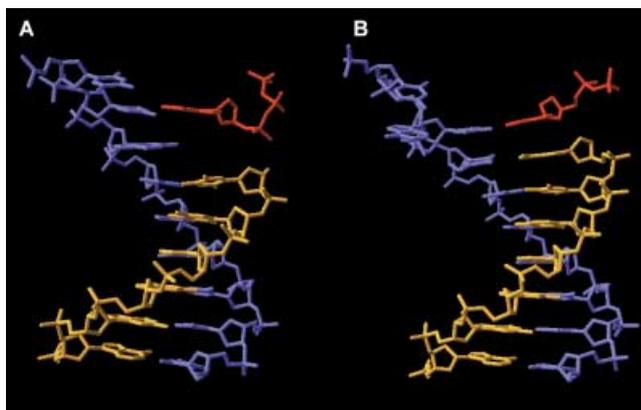


Figure 2. DNA bound in the active site. Protein elements are removed for clarity (PDB entry codes 1JX4 and 1JXL). A: primer (gold) and template (blue) in Dpo4 with incoming noncanonical triphosphate (ddGTP, red) exhibiting translocation of the template that results in a base pair between ddGTP and the second nucleobase in the template. B: primer (gold) and template (blue) in Dpo4 with incoming canonical triphosphate ddATP (red).

We thank A. K. Aggarwal, New York, for kindly providing the coordinates of DNA polymerase η and M. Strerath, Bonn, for his assistance in preparing the graphical material.

- [1] Recent summaries: a) T. Carell, L. T. Burgdorf, L. M. Kundu, M. Cichon, *Curr. Opin. Chem. Biol.* **2001**, *5*, 491–498; b) T. Lindahl, R. D. Wood, *Science* **1999**, *286*, 1897–1905; c) E. C. Friedberg, G. C. Walker, W. Siede, *DNA Repair and Mutagenesis*, American Society for Microbiology Press, Washington DC, **1995**.
- [2] Recent reviews and commentaries: a) E. C. Friedberg, P. L. Fischhaber, C. Kisker, *Cell* **2001**, *107*, 9–12; b) Z. Livneh, *J. Biol. Chem.* **2001**, *276*, 25639–25642; c) M. F. Goodman, B. Tippin, *Nat. Rev. Mol. Cell. Biol.* **2000**, *1*, 101–109; d) U. Hübscher, H. P. Nasheuer, J. E. Syvaaja, *Trends Biochem. Sci.* **2000**, *25*, 143–147; e) E. C. Friedberg, W. J. Feaver, V. L. Gerlach, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5681–5683; f) E. C. Friedberg, V. L. Gerlach, *Cell*, **1999**, *98*, 413–416; g) R. E. Johnson, M. T. Washington, S. Prakash, L. Prakash, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12224–12226.
- [3] a) A. Vaisman, A. Tissier, E. G. Frank, M. F. Goodman, R. Woodgate, *J. Biol. Chem.* **2001**, *276*, 30615–30622; b) R. E. Johnson, M. T. Washington, L. Haracska, S. Prakash, L. Prakash, *Nature* **2000**, *406*, 1015–1019.
- [4] K. Bebenek, A. Tissier, E. G. Frank, J. P. McDonald, R. Prasad, S. H. Wilson, R. Woodgate, T. A. Kunkel, *Science* **2001**, *291*, 2156–2159.
- [5] a) E. Glick, K. L. Vigna, L. A. Loeb, *EMBO J.* **2001**, *20*, 7303–7312; b) M. T. Washington, R. E. Johnson, L. Prakash, S. Prakash, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8355–8360; c) R. E. Johnson, M. T. Washington, S. Prakash, L. Prakash, *J. Biol. Chem.* **2000**, *275*, 7447–7450; d) C. Masutani, R. Kusumoto, S. Iwai, F. Hanaoka, *EMBO J.* **2000**, *19*, 3100–3109; e) R. E. Johnson, C. M. Kondratik, S. Prakash, L. Prakash, *Science* **1999**, *285*, 263–265; f) C. Masutani, R. Kusumoto, A. Yamada, N. Dohmae, M. Yokoi, M. Yuasa, M. Araki, S. Iwai, K. Takio, F. Hanaoka, *Nature* **1999**, *399*, 700–704; g) R. E. Johnson, S. Prakash, L. Prakash, *Science* **1999**, *283*, 1001–1004.
- [6] a) L. F. Silvan, E. A. Toth, P. Pham, M. F. Goodman, T. Ellenberger, *Nat. Struct. Biol.* **2001**, *8*, 984–989; b) B.-L. Zhou, J. D. Pata, T. A. Steitz, *Mol. Cell* **2001**, *8*, 427–437; c) J. Trincão, R. E. Johnson, C. R. Escalante, S. Prakash, L. Prakash, A. K. Aggarwal, *Mol. Cell* **2001**, *8*, 417–426; d) H. Ling, F. Boudsocq, R. Woodgate, W. Yang, *Cell* **2001**, *107*, 91–102.
- [7] Recent summaries: a) P. H. Patel, L. A. Loeb, *Nat. Struct. Biol.* **2001**, *8*, 656–659; b) P. H. Patel, M. Suzuki, E. Adman, A. Shinkai, L. A. Loeb, *J. Mol. Biol.* **2001**, *308*, 823–837; c) T. A. Kunkel, K. Bebenek, *Annu. Rev. Biochem.* **2000**, *69*, 497–529.
- [8] a) M. C. Franklin, J. Wang, T. A. Steitz, *Cell* **2001**, *98*, 413–416; b) S. Doublé, S. Tabor, A. M. Long, C. C. Richardson, T. Ellenberger, *Nature* **1998**, *391*, 251–258; c) Y. Li, S. Korolev, G. Waksman, *EMBO J.* **1998**, *17*, 7514–7525; d) J. R. Kiefer, C. Mao, J. C. Braman, L. S. Beese, *Nature* **1998**, *391*, 304–307; e) H. F. Huang, R. Chopra, G. L. Verdine, S. C. Harrison, *Science* **1998**, *282*, 1669–1675; f) D. L. Ollis, P. Brick, R. Hamlin, N. G. Xuong, T. A. Steitz, *Nature* **1985**, *313*, 762–766.
- [9] For DNA polymerases that are not handlike, see: a) A. K. Showalter, I.-J. L. Byeon, M.-I. Su, M.-D. Tsai, *Nat. Struct. Biol.* **2001**, *8*, 942–946; b) M. W. Maciejewski, R. Shin, B. Pan, A. Marintchev, A. Denninger, M. A. Mullen, K. Chen, M. R. Gryk, G. P. Mullen *Nat. Struct. Biol.* **2001**, *8*, 936–941.
- [10] a) E. T. Kool, J. C. Morales, K. M. Guckian, *Angew. Chem.* **2000**, *112*, 1046–1068; *Angew. Chem. Int. Ed.* **2000**, *39*, 991–1009; b) T. A. Kunkel, S. H. Wilson, *Nat. Struct. Biol.* **1998**, *5*, 95–99; c) U. Diederichsen, *Angew. Chem.* **1998**, *110*, 1745–1747; *Angew. Chem. Int. Ed.* **1998**, *37*, 1655–1657; d) M. F. Goodman, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10493–10495.

Received: November 22, 2001 [H 323]