

Metal-ion dependence of the active-site conformation of the translesion DNA polymerase Dpo4 from *Sulfolobus solfataricus*

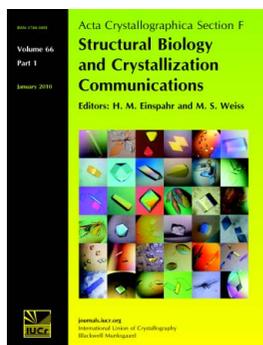
Adriana Irimia, Lioudmila V. Loukachevitch, Robert L. Eoff, F. Peter Guengerich and Martin Egli

Acta Cryst. (2010). **F66**, 1013–1018

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Reproduction of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>



Acta Crystallographica Section F: Structural Biology and Crystallization Communications is a rapid all-electronic journal, which provides a home for short communications on the crystallization and structure of biological macromolecules. It includes four categories of publication: protein structure communications; nucleic acid structure communications; structural genomics communications; and crystallization communications. Structures determined through structural genomics initiatives or from iterative studies such as those used in the pharmaceutical industry are particularly welcomed. *Section F* is essential for all those interested in structural biology including molecular biologists, biochemists, crystallization specialists, structural biologists, biophysicists, pharmacologists and other life scientists.

Crystallography Journals **Online** is available from journals.iucr.org

Adriana Irimia,‡ Lioudmila V. Loukachevitch, Robert L. Eoff, F. Peter Guengerich and Martin Egli*

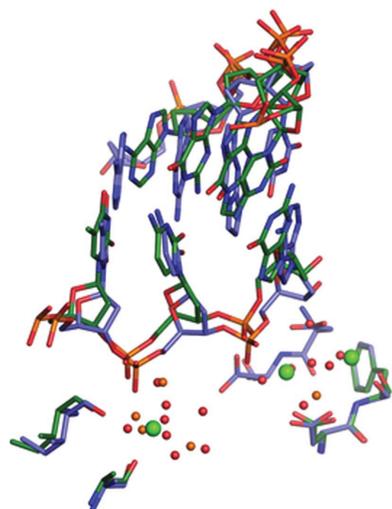
Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University, Nashville, TN 37232, USA

‡ Present address: The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA.

Correspondence e-mail: martin.egli@vanderbilt.edu

Received 18 June 2010
Accepted 23 July 2010

PDB References: Dpo4–Mg²⁺ complex, 2xc9;
Dpo4–dGTP–Mg²⁺ complex, 2xca;
Dpo4–dCTP–Mg²⁺ complex, 2xcp.



© 2010 International Union of Crystallography
All rights reserved

Metal-ion dependence of the active-site conformation of the translesion DNA polymerase Dpo4 from *Sulfolobus solfataricus*

Crystal structures of a binary Mg²⁺-form Dpo4–DNA complex with 1,N²-etheno-dG in the template strand as well as of ternary Mg²⁺-form Dpo4–DNA–dCTP/dGTP complexes with 8-oxoG in the template strand have been determined. Comparison of their conformations and active-site geometries with those of the corresponding Ca²⁺-form complexes revealed that the DNA and polymerase undergo subtle changes as a result of the catalytically more active Mg²⁺ occupying both the A and B sites.

1. Introduction

The Y-class DNA polymerase IV (Dpo4) from *Sulfolobus solfataricus* P2 has served as a useful model system in structure–function studies of damage bypass. Although *S. solfataricus* Dpo4 is considered to be the archaeal homolog of human polymerase κ (hpol κ ; Boudsocq *et al.*, 2001), its bypass behavior differs distinctly from that of pol κ , instead bearing a closer resemblance to that of another of the four eukaryotic translesion polymerases, pol η (Boudsocq *et al.*, 2001; Fiala & Suo, 2004; Prakash *et al.*, 2005). Like pol η , Dpo4 is able to efficiently and accurately bypass template 7,8-dihydro-8-oxo-2'-deoxyguanosine (8-oxoG) by inserting mostly dCTP (Zang *et al.*, 2006), whereas bypass synthesis catalyzed by hpol κ is error-prone, resulting in the incorporation of mostly dATP opposite 8-oxoG (Irimia *et al.*, 2009 and references cited therein).

Among the Y-class translesion DNA polymerases, the structure and mechanism of Dpo4 have arguably been studied in the greatest detail (reviewed in Guengerich, 2006; Yang & Woodgate, 2007; Eoff *et al.*, 2010). Ternary Dpo4–DNA–dNTP complexes have been determined with native template–primer strands (Ling, Boudsocq *et al.*, 2001) and duplexes containing a *cis-syn* thymine dimer (Ling *et al.*, 2003), base mismatches (Trincao *et al.*, 2004), abasic sites (Ling, Boudsocq *et al.*, 2004), benzo[*a*]pyrene (Ling, Sayer *et al.*, 2004), 1,N²-etheno-dG (1,N²- ϵ -G; Zang *et al.*, 2005), 8-oxoG (Zang *et al.*, 2006; Eoff, Irimia, Angel *et al.*, 2007), O⁶-methyl-dG (Eoff, Irimia, Egli *et al.*, 2007), O⁶-benzyl-dG (Eoff, Angel *et al.*, 2007), the hydrophobic T analog 2,4-difluorotoluene (Irimia *et al.*, 2007), 1,N²-propano-dG (Wang *et al.*, 2008), N²-alkyl-dG adducts (Zhang *et al.*, 2009*a*), N²,N²-dimethyl-dG (Zhang *et al.*, 2009*b*), a malondialdehyde–dG adduct (Eoff *et al.*, 2009) and an aminofluorene adduct of dG (Rechkoblit *et al.*, 2010). In all these complexes the active-site metal ions are Ca²⁺ (Vaisman *et al.*, 2005; Rechkoblit *et al.*, 2006). To further ensure the inhibition of primer extension during crystallization, primers with 2',3'-dideoxynucleotides at their 3'-termini were used in some cases (*i.e.* Ling *et al.*, 2001). However, we demonstrated that Ca²⁺ (but not Ba²⁺, Co²⁺, Cu²⁺, Ni²⁺ or Zn²⁺) is a cofactor for Dpo4-catalyzed polymerization opposite both native and 8-oxoG-modified templates and that dNTPs and ddNTPs are polymerase substrates in the presence of Mg²⁺ or Ca²⁺ (Irimia *et al.*, 2006).

To establish putative changes in the Dpo4 active-site geometry as a result of the replacement of Mg²⁺ by Ca²⁺ and/or alterations in the relative orientation of polymerase and DNA, we determined three structures of Dpo4–DNA complexes crystallized with Mg²⁺ and a primer (ternary complexes) featuring 2',3'-dideoxycytidine at its 3'-end.

Table 1

Crystal data and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Binary Dpo4-Mg ²⁺	Dpo4-dGTP-Mg ²⁺	Dpo4-dCTP-Mg ²⁺
Type of lesion	1,N ² -ε-G	8-oxoG	8-oxoG
Crystal data and data collection			
X-ray source	APS	APS	APS
Beamline	SER-CAT 22-BM	DND-CAT 5-ID	SER-CAT 22-ID
Detector	MAR CCD 300	MAR CCD 225	MAR CCD 300
Wavelength (Å)	1.00	1.00	1.00
Temperature (K)	110	110	110
No. of crystals	1	1	1
Space group	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁
Unit-cell parameters			
<i>a</i> (Å)	93.72	94.44	59.75
<i>b</i> (Å)	102.12	103.90	100.42
<i>c</i> (Å)	53.00	52.70	105.87
α (°)	90	90	90
β (°)	90	90	96.07
γ (°)	90	90	90
Resolution range (Å)	50.0–2.2 (2.28–2.2)	29.13–2.5 (2.66–2.5)	49.5–2.6 (2.76–2.6)
No. of measurements	331946	106973	195670
No. of unique reflections	26245 (4241)	18457 (3005)	37629 (5685)
Redundancy	12.6 (10.4)	5.8 (5.3)	5.2 (3.8)
Completeness (%)	99.3 (99.3)	99.6 (99.9)	98.1 (90.1)
<i>R</i> _{merge} [†]	6.2 (48.2)	6.4 (48.6)	6.8 (49.4)
$\langle I/\sigma(I) \rangle$	36.3 (6.2)	21.3 (4.2)	14.8 (3.7)
Solvent content (%)	48.1	50.1	59.6
Refinement			
Model composition (asymmetric unit)			
No. of amino-acid residues	341	341	341/342
No. of water molecules	247	192	305
No. of Mg ²⁺ ions	0	4	3/3
No. of template nucleotides	16	17	17/17
No. of primer nucleotides	14	13	13/13
No. of dGTP	—	1	—
No. of dCTP	—	—	1/1
<i>R</i> ₁ [‡] (%)	22.1	22.0	22.3
<i>R</i> _{free} [§] (%)	25.0	25.9	26.8
Estimated coordinate error (Å)			
Luzzati plot	0.30	0.32	0.34
Luzzati plot (cross-validation)	0.35	0.40	0.42
σ_A plot	0.28	0.35	0.42
σ_A plot (cross-validation)	0.30	0.43	0.49
Temperature factors			
Wilson plot (Å ²)	41.3	57.4	47.0
Mean isotropic (Å ²)	40.7	44.1	48.3
Root-mean-square deviation in temperature factors			
Bonded main-chain atoms (Å ²)	1.4	1.3	1.3
Bonded side-chain atoms (Å ²)	2.0	1.9	1.9
Root-mean-square deviation from ideal values			
Bond lengths (Å)	0.007	0.007	0.007
Bond angles (°)	1.5	1.4	1.3
Dihedral angles (°)	21.3	21.6	22.0
Improper angles (°)	1.1	1.6	1.0

[†] $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where the outer sum (*hkl*) is taken over the unique reflections. [‡] $R = \sum_{hkl} (|F_{\text{obs}}| - |F_{\text{calc}}|) / \sum_{hkl} |F_{\text{obs}}|$, where $|F_{\text{obs}}|$ and $|F_{\text{calc}}|$ are the observed and calculated structure-factor amplitudes, respectively. [§] R_{free} is the same as *R* but calculated for a set of reflections (5% of the total) omitted from the refinement process.

2. Methods

2.1. Protein expression and purification

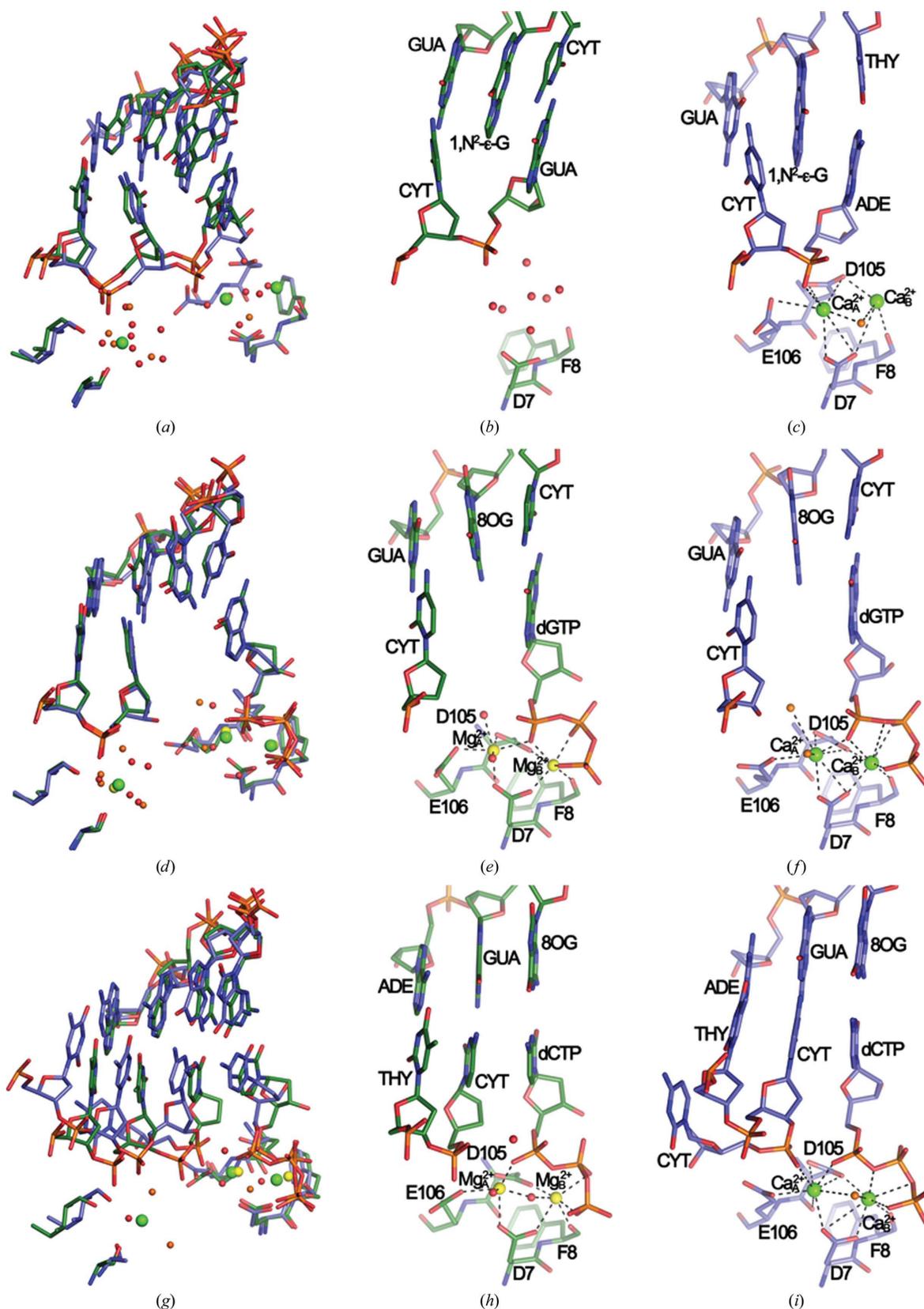
Wild-type Dpo4 DNA polymerase from *S. solfataricus* was expressed in *Escherichia coli* and purified to electrophoretic homogeneity as described previously (Zang *et al.*, 2005).

2.2. Crystallization of Dpo4–DNA and Dpo4–DNA–dNTP Mg²⁺-form complexes

The sequences of the 18-mer template and 14-mer primer strands used for the binary Dpo4 complex with Mg²⁺ are similar to those used for the binary complex obtained in the presence of Ca²⁺ (Zang *et al.*, 2005). However, the base pair at the replicative position is G₁₄:C₄^T in the present study instead of A₁₄:T₄^T in the Ca²⁺-form complex (Zang *et*

al., 2005). Specifically, the primer was 5'-GGG GGA AGG ATT CG-3' and the template was 5'-TCA C(1,N²-ε-G)G AAT CCT TCC CCC-3'. The sequences of the 18-mer template and 13-mer primer strands in the ternary Dpo4–DNA–dGTP and Dpo4–DNA–dCTP Mg²⁺-form complexes are identical to those in the corresponding Ca²⁺-form complexes (Zang *et al.*, 2005), except that the primer strands contain a 2',3'-dideoxy residue (C*) in the present study. Specifically, the primer was 5'-GGG GGA AGG ATT C*-3' and the template was 5'-TCA C(8-oxoG)G AAT CCT TCC CCC-3'.

All DNA duplexes were obtained by annealing primer and template strands. The duplexes were then incubated for 30 min on ice with the Dpo4 polymerase in a 1:1.2 (protein:DNA) ratio in 20 mM Tris–HCl buffer pH 7.5 containing 60 mM NaCl. In the case of the ternary complexes, Dpo4 and DNA were incubated for 10 min with 1 mM dGTP (Dpo4–dGTP–Mg²⁺ complex) or 1 mM dCTP

**Figure 1**

Dpo4 Mg^{2+} -form versus Ca^{2+} -form complexes. (a) Superimposition of the binary Dpo4- Mg^{2+} (green) and Dpo4- Ca^{2+} (blue; PDB code 2bq3) complexes. The active sites of the binary Dpo4- Mg^{2+} (b) and binary Dpo4- Ca^{2+} (c) complexes are shown. (d) Superimposition of the Dpo4-dGTP- Mg^{2+} (green) and Dpo4-dGTP- Ca^{2+} (blue; PDB code 2c22) complexes. The active sites of the Dpo4-dGTP- Mg^{2+} (e) and Dpo4-dGTP- Ca^{2+} (f) complexes are shown. (g) Superimposition of the Dpo4-dCTP- Mg^{2+} (green) and Dpo4-dCTP- Ca^{2+} (blue; PDB code 2c2e) complexes. The active sites of the Dpo4-dCTP- Mg^{2+} (h) and Dpo4-dCTP- Ca^{2+} (i) complexes are shown. Mg^{2+} ions are shown as yellow spheres and Ca^{2+} ions are shown as green spheres. Water molecules in the Dpo4- Mg^{2+} and Dpo4- Ca^{2+} complexes are shown as small red and orange spheres, respectively.

(Dpo4-dCTP-Mg²⁺ complex) in the presence of 5 mM MgCl₂. Crystals were grown using the sitting-drop vapor-diffusion method by mixing equal amounts of Dpo4-DNA complex solution and of a reservoir solution containing 12–20% polyethylene glycol 3350, 0.2 M ammonium acetate, 0.1 M magnesium acetate and 20 mM Tris pH 7.5. All crystals were flash-frozen and stored in liquid nitrogen prior to data collection.

2.3. X-ray data collection, processing, structure determination and refinement

X-ray diffraction data were collected on beamlines at the Advanced Photon Source (Argonne National Laboratory, Argonne, Illinois, USA) at a wavelength of 1.00 Å (Table 1). Individual diffraction data sets were processed using *HKL-2000* (Otwinowski & Minor, 1997). *CCP4* programs (Collaborative Computational Project, Number 4, 1994), including the *TRUNCATE* procedure (French & Wilson, 1978), were used to further process the data. Selected crystal data and statistics relating to data processing and quality are listed in Table 1.

Because the binary Dpo4-Mg²⁺ and ternary Dpo4-dGTP-Mg²⁺ crystals in the present study exhibited the same space group (*P*₂₁₂₁) and similar unit-cell parameters as the corresponding Ca²⁺-form complexes (Zang *et al.*, 2005, 2006), the refined models of the latter complexes (PDB codes 2bq3 and 2c22, respectively) without solvent molecules were used as starting models. The initial models were refined using several rounds of rigid-body refinement with increasing resolution until the diffraction limit was reached. Map visualization and model rebuilding were performed with the program *TURBO-FRODO* (Cambillau & Roussel, 1997) and all refinements were conducted with the program *CNS* (Brünger *et al.*, 1998).

The Dpo4-dCTP-Mg²⁺ complex and the corresponding Ca²⁺-form complex (Zang *et al.*, 2006) crystallized in different space groups with deviating unit-cell parameters. Therefore, the refined structure of the latter complex (PDB code 2c2e) was used as a model for molecular-replacement searches with the program *Phaser* (McCoy *et al.*, 2007). Refinement and model building were carried out with the programs *CNS* and *TURBO-FRODO*, respectively, and final data statistics and refinement parameters are summarized in Table 1. All figures were prepared with the program *PyMOL* (DeLano, 2002).

2.4. Data deposition

Crystallographic models of the three complexes have been deposited in the Protein Data Bank (PDB; <http://www.rcsb.org>). The PDB codes are 2xc9 for the binary Dpo4-Mg²⁺ complex, 2xca for the ternary Dpo4-dGTP-Mg²⁺ complex and 2xcp for the ternary Dpo4-dCTP-Mg²⁺ complex.

3. Results and discussion

The structure of a binary complex (binary Dpo4-Mg²⁺) with an 18-mer template containing 1, *N*²-ε-G at the -1 position and a 14-mer primer was determined at 2.2 Å resolution. The structures of ternary complexes with 13-mer primers and 18-mer templates containing 8-oxoG either at the replicating (paired to dCTP, termed Dpo4-dCTP-Mg²⁺) or the -1 position (unopposed by a primer base and dC-dGTP at the replicating position, termed Dpo4-dGTP-Mg²⁺) were determined at 2.6 and 2.5 Å resolution, respectively. All relevant crystal data and diffraction data-collection and refinement parameters are listed in Table 1. The binary Dpo4-Mg²⁺ (Figs. 1a–1c) and Dpo4-dGTP-Mg²⁺ (Figs. 1d–1f) complexes represent so-called type II structures (Ling *et al.*, 2001), with two template bases residing

at the polymerase active site (1, *N*²-ε-dG and dC, and 8-oxoG and dC, respectively). The Dpo4-dCTP-Mg²⁺ complex (Figs. 1g–1i) is representative of a type I structure, with only the replicating template base (8-oxoG) paired to the incoming nucleotide (dCTP) at the active site.

In the binary Dpo4-Mg²⁺ complex the primer strand is detached from the cluster of Glu and Asp residues that mark the active site and its orientation is not compatible with metal ions engaged in inner-sphere coordination to both the terminal phosphate and acidic side chains (Figs. 1a–1c). We modeled solvent peaks in the vicinity of the terminal phosphate (*n*), formerly occupied by the Ca_A²⁺ and Ca_B²⁺ ions (Fig. 1c), and near the *n* - 1 phosphate as water molecules and alternatively as Mg²⁺ or Na⁺ hexahydrates (Supplementary Fig. S1¹). Of the three scenarios, that with only water molecules appeared to be the most plausible on the basis of the observed distances between the backbone phosphates and the conserved Asp (Asp7, Asp105) and Glu (Glu106) residues that are involved in metal-ion coordination at the active site (Supplementary Fig. S1). Thus, the average distance between the Asp OD1/2 and Glu OE1/2 atoms and the phosphate of the 3'-terminal G in the primer strand is 6.8 Å (Supplementary Fig. S1a) and none of the modeled Mg²⁺ or Na⁺ ions engage in outer sphere contacts to the *n* - 1 or *n* phosphate groups (Supplementary Figs. S1b and S1c). An overall view of the superimposition of the binary Dpo4-Mg²⁺ and Dpo4-Ca²⁺ complexes revealed only minor deviations between the two apart from the distinct orientation of the primer strand in the former (Supplementary Figs. S2a–S2c). We believe that the change from an A·T (Ca²⁺-form) to a G·C pair (Mg²⁺-form) at the active site is unlikely to be the cause of the observed conformational and metal-ion coordination differences.

Comparison of the ternary Dpo4-dGTP-Mg²⁺ and Dpo4-dGTP-Ca²⁺ complexes revealed only very minor deviations in the relative orientations of the polymerase, template–primer duplex and dGTP (Figs. 1d–1f). Similarly, the conformations of the three phosphates of the incoming nucleotide display a close resemblance in the two structures. Owing to the different coordination geometries and ionic radii of Mg²⁺ and Ca²⁺ (Irimia *et al.*, 2006, and references cited therein), the side chains of Asp7, Asp105 and Glu106 at the active site adopt slightly different conformations in the two complex forms (Figs. 1d–1f). Both Mg_A²⁺ and Mg_B²⁺ exhibit a virtually ideal octahedral geometry (six ligands), whereas the Ca²⁺ ions are engaged in seven or eight contacts. The actual locations of the two metal ions in the structures as well as that of a third ion near the *n* - 1 phosphate are rather similar (Fig. 1d). The sugar of the 3'-terminal primer residue in both complexes (2',3'-ddC in the Mg²⁺-form and 2'-dC in the Ca²⁺-form) adopts a *C2'-endo* conformation. However, the puckers of the dGTPs differ: *C3'-endo* in the Mg²⁺-form and *C2'-endo* in the Ca²⁺-form (Fig. 1d). However, this change is of no consequence in terms of the spacing between the C2'–C3' sugar edge of the 3'-terminal residue of the primer and the α-phosphate of dGTP in the two structures. In the Mg²⁺-form complex C3' (ddC) lies at 7.2 Å from Pα; this relatively large distance is a characteristic feature of type II Dpo4 complexes in which two primer residues wrap around an unpaired template residue (8-oxoG; Figs. 1e and 1f). Whereas the two ternary dGTP complexes are similar despite harboring distinct alkaline-earth metal ions at their active sites (Supplementary Figs. S2d–S2f), an overlay of the Dpo4-dCTP-Mg²⁺ and binary Dpo4-Mg²⁺ complexes again demonstrates the distinct shift of the primer strand in the latter away from the active site (Fig. 2a).

¹ Supplementary material has been deposited in the IUCr electronic archive (Reference: BE5151).

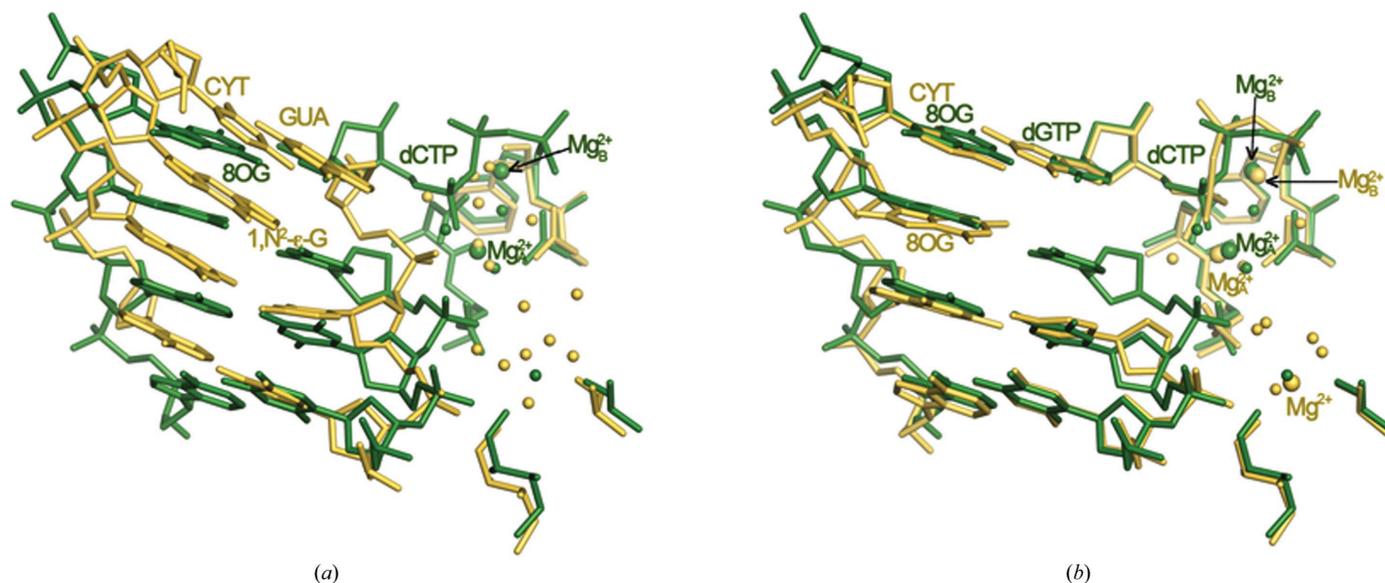


Figure 2 Comparisons between binary and ternary Mg^{2+} -form Dpo4 complexes. (a) The binary Dpo4- Mg^{2+} complex (yellow) superimposed on the ternary Dpo4-dCTP- Mg^{2+} complex (green). (b) The binary Dpo4-dGTP- Mg^{2+} complex (yellow) superimposed on the Dpo4-dCTP- Mg^{2+} complex (green). Active-site nucleotides and ions are labeled with colors matching those of the nucleotides and amino acids. Water molecules are shown as small spheres; larger spheres represent Mg^{2+} ions.

Compared with the binary and ternary dGTP complexes, more significant differences are seen between the overall structures (Supplementary Figs. S2g–S2i) of the ternary Dpo4-dCTP- Mg^{2+} and Dpo4-dCTP- Ca^{2+} complexes as well as between their active sites (Figs. 1g–1i). Both are of type I, but in the Ca^{2+} -form the primer was extended by a single residue during crystallization. The primer strand curls around between dC at position 13 and dC at position 14 and the latter is accommodated inside the minor groove (Fig. 1i, lower left). A superimposition of the two complexes reveals that equivalent base pairs are staggered and that the penultimate residue of the primer in the Ca^{2+} -form complex is farther removed from $P\alpha$ of dCTP relative to the corresponding situation in the Mg^{2+} -form (3'-terminal ddC · $P\alpha$; Fig. 1g). Such a difference is not observed when the ternary Dpo4-dGTP- Mg^{2+} and Dpo4-dCTP- Mg^{2+} complexes are compared (Fig. 2b). The two complexes also crystallize in different space groups ($P2_12_12_1$ for the Ca^{2+} -form *versus* $P2_1$ for the Mg^{2+} -form; Table 1). Crystals of the first form contain one complex per asymmetric unit and crystals of the second feature two complexes per asymmetric unit. Interestingly, the sugar conformations of the 3'-terminal ddC residues differ in the two independent Mg^{2+} -form complexes: C3'-*endo* (Fig. 1h) *versus* C2'-*endo* (the corresponding dC in the Ca^{2+} -form complex exhibits a C2'-*endo* pucker; Fig. 1i). The conformational flip critically affects the distance between C3' (or O3' in dC) and $P\alpha$: 3.7 Å (C3'-*endo*) *versus* 4.3 Å (C2'-*endo*). The C3' · $P\alpha$ distance is 5.2 Å in the Ca^{2+} -form complex. The relatively tight distance between the terminal residue of the primer and the α -phosphate of dCTP in Dpo4-dCTP- Mg^{2+} is noteworthy. The more optimal line-up of nucleophile and incoming dNTP compared with the Ca^{2+} -form complex may support the conclusion that the use of magnesium allows more optimal structural studies of precatalytic polymerase complexes and mechanistic insights into accurate replication and bypass. However, the comparison is complicated by the fact that the complexes crystallized in different space groups and the presence of an additional nucleotide at the 3'-end of the primer in the Ca^{2+} -form complex that was added during crystallization. Additionally, the resolutions of around 2.5 Å of the crystal structures of the ternary Mg^{2+} -form complexes and the absence of the 3'-hydroxyl group at the primer terminus limit to some degree definitive state-

ments about the sugar conformation and therefore the proximity to the incoming nucleotide.

4. Conclusions

In summary, we describe the first crystal structures obtained in the presence of Mg^{2+} instead of Ca^{2+} of binary Dpo4-DNA and ternary Dpo4-DNA-dC(G)TP complexes with 1, N^2 - ϵ -G and 8-oxoG, respectively, in the template strand. Comparisons of their conformations and active-site geometries with those of the corresponding Ca^{2+} -form complexes revealed that the DNA and polymerase undergo subtle changes as a result of the catalytically more active Mg^{2+} occupying both the A and B sites. The observation, at least in the structures of the ternary Mg^{2+} -form and Ca^{2+} -form complexes, of relatively small changes in the overall and active-site conformations is in line with our previous finding that Ca^{2+} is a cofactor (albeit a poor one compared with Mg^{2+}) of Dpo4-catalyzed DNA polymerization.

This work was supported by US NIH Grants P01 ES05355 (to ME), P30 ES000267 (to FPG and ME), R01 ES010375 (to FPG), F32 CA119776 and K99 GM084460 (to RLE). We are grateful to Drs P. S. Pallan, R. Pattanayek and Z. Wawrzak for assistance with X-ray data collection. Vanderbilt University is a member institution of LS-CAT at the APS, Argonne, Illinois, USA. Use of the APS was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-06CH11357.

References

- Boudsocq, F., Iwai, S., Hanaoka, F. & Woodgate, R. (2001). *Nucleic Acids Res.* **29**, 4607–4616.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
- Cambillau, C. & Roussel, A. (1997). *TURBO-FRODO*, version OpenGL.1. University Aix-Marseille II, Marseille, France.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.

- DeLano, W. L. (2002). *PyMOL Molecular Viewer*. DeLano Scientific, San Carlos, California, USA.
- Eoff, R. L., Angel, K. C., Egli, M. & Guengerich, F. P. (2007). *J. Biol. Chem.* **282**, 13573–13584.
- Eoff, R. L., Egli, M. & Guengerich, F. P. (2010). *The Chemical Biology of DNA Damage*, edited by N. E. Geacintov & S. Broyde, pp. 299–330. Weinheim: Wiley-VCH.
- Eoff, R. L., Irimia, A., Angel, K. C., Egli, M. & Guengerich, F. P. (2007). *J. Biol. Chem.* **282**, 19831–19843.
- Eoff, R. L., Irimia, A., Egli, M. & Guengerich, F. P. (2007). *J. Biol. Chem.* **282**, 1456–1467.
- Eoff, R. L., Stafford, J., Szekely, J., Rizzo, C. J., Egli, M., Guengerich, F. P. & Marnett, L. J. (2009). *Biochemistry*, **48**, 7079–7088.
- Fiala, K. A. & Suo, Z. (2004). *Biochemistry*, **43**, 2116–2125.
- French, S. & Wilson, K. (1978). *Acta Cryst.* **A34**, 517–525.
- Guengerich, F. P. (2006). *Chem. Rev.* **106**, 420–452.
- Irimia, A., Eoff, R. L., Guengerich, F. P. & Egli, M. (2009). *J. Biol. Chem.* **284**, 22467–22480.
- Irimia, A., Eoff, R. L., Pallan, P. S., Guengerich, F. P. & Egli, M. (2007). *J. Biol. Chem.* **282**, 36421–36433.
- Irimia, A., Zang, H., Loukachevitch, L. V., Eoff, R. L., Guengerich, F. P. & Egli, M. (2006). *Biochemistry*, **45**, 5949–5956.
- Ling, H., Boudsocq, F., Plosky, B. S., Woodgate, R. & Yang, W. (2003). *Nature (London)*, **424**, 1083–1087.
- Ling, H., Boudsocq, F., Woodgate, R. & Yang, W. (2001). *Cell*, **107**, 91–102.
- Ling, H., Boudsocq, F., Woodgate, R. & Yang, W. (2004). *Mol. Cell*, **13**, 751–762.
- Ling, H., Sayer, J. M., Plosky, B. S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D. M. & Yang, W. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 2265–2269.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Prakash, S., Johnson, R. E. & Prakash, L. (2005). *Annu. Rev. Biochem.* **74**, 317–353.
- Rechkoblit, O., Kolbanovskiy, A., Malinina, L., Geacintov, N. E., Broyde, S. & Patel, D. J. (2010). *Nature Struct. Mol. Biol.* **17**, 379–388.
- Rechkoblit, O., Malinina, L., Cheng, Y., Kuryavyy, V., Broyde, S., Geacintov, N. E. & Patel, D. J. (2006). *PLoS Biol.* **4**, 25–42.
- Trincao, J., Johnson, R. E., Wolffe, W. T., Escalante, C. R., Prakash, S., Prakash, L. & Aggarwal, A. K. (2004). *Nature Struct. Mol. Biol.* **11**, 457–462.
- Vaisman, A., Ling, H., Woodgate, R. & Yang, W. (2005). *EMBO J.* **24**, 2957–2967.
- Wang, Y., Musser, S., Saleh, S., Marnett, L. J., Egli, M. & Stone, M. P. (2008). *Biochemistry*, **47**, 7322–7334.
- Yang, W. & Woodgate, R. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 15591–15598.
- Zang, H., Goodenough, A. K., Choi, J.-Y., Irimia, A., Loukachevitch, L. V., Kozekov, I. D., Angel, K. C., Rizzo, C. J., Egli, M. & Guengerich, F. P. (2005). *J. Biol. Chem.* **280**, 29750–29764.
- Zang, H., Irimia, A., Choi, J.-Y., Angel, K. C., Loukachevitch, L. V., Egli, M. & Guengerich, F. P. (2006). *J. Biol. Chem.* **281**, 2358–2372.
- Zhang, H., Eoff, R. L., Kozekov, I. D., Rizzo, C. J., Egli, M. & Guengerich, F. P. (2009a). *J. Biol. Chem.* **284**, 3563–3576.
- Zhang, H., Eoff, R. L., Kozekov, I. D., Rizzo, C. J., Egli, M. & Guengerich, F. P. (2009b). *J. Biol. Chem.* **284**, 17687–17699.