2948–2957 *Nucleic Acids Research, 2008, Vol. 36, No. 9* *Published online 2 April 2008*
doi:10.1093/nar/gkn138

Mismatched dNTP incorporation by DNA polymerase β does not proceed via globally different conformational pathways†

Kuo-Hsiang Tang[1,2], Marc Niebuhr[3], Chang-Shung Tung[4], Hsiu-chien Chan[2,5], Chia-Cheng Chou[2] and Ming-Daw Tsai[1,2,6,*]

[1]Departments of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210, USA, [2]Genomics Research Center, Academia Sinica, Taiwan, [3]Stanford Synchrotron Radiation Laboratory, MS99, SLAC, Menlo Park, CA 94025, [4]Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, [5]Institute of Bioinformatics and Structural Biology, National Tsing Hua University and [6]Institute of Biological Chemistry, Academia Sinica, Taiwan

Received December 17, 2007; Revised March 7, 2008; Accepted March 12, 2008

ABSTRACT

Understanding how DNA polymerases control fidelity requires elucidation of the mechanisms of matched and mismatched dNTP incorporations. Little is known about the latter because mismatched complexes do not crystallize readily. In this report, we employed small-angle X-ray scattering (SAXS) and structural modeling to probe the conformations of different intermediate states of mammalian DNA polymerase β (Pol β) in its wild-type and an error-prone variant, I260Q. Our structural results indicate that the mismatched ternary complex lies in-between the open and the closed forms, but more closely resembles the open form for WT and the closed form for I260Q. On the basis of molecular modeling, this over-stabilization of mismatched ternary complex of I260Q is likely caused by formation of a hydrogen bonding network between the side chains of Gln^{260}, Tyr^{296}, Glu^{295} and Arg^{258}, freeing up Asp^{192} to coordinate MgdNTP. These results argue against recent reports suggesting that mismatched dNTP incorporations follow a conformational path distinctly different from that of matched dNTP incorporation, or that its conformational closing is a major contributor to fidelity.

High accuracy of DNA replication catalyzed by DNA polymerases is essential for genome stability. DNA polymerase β (Pol β) is an essential enzyme for DNA repair in mammalian cells, and has long served as a model system for studying DNA polymerases (1). Pol β is naturally deficient in 3′–5′ exonuclease proofreading activity and consists of an N-terminal 8 kDa 5′-dRP lyase domain and a C-terminal 31 kDa DNA polymerase (nucleotidyl transferase) domain; the latter includes a DNA binding, a nucleotidyl transferase (catalytic) and a dNTP-binding subdomain (1). Crystallographic studies have suggested that Pol β undergoes sequential conformational changes as it first binds to DNA and then binds a (matched) dNTP (Figure 1A) (2,3). This structural information along with extensive kinetic studies led to elucidation of the detailed mechanism for matched dNTP incorporation by Pol β. In this mechanism, binding of the matched dNTP to the Pol β·DNA binary complex causes a fast dNTP-induced conformational closing of the Pol β·DNA·dNTP ternary complex, followed by a rate-limiting chemical step, subsequent dNTP-binding subdomain reopening and finally pyrophosphate release (4,5).

Despite these progresses, a lack of structural and mechanistic information concerning mismatched dNTP incorporation leaves the molecular mechanism of DNA polymerase fidelity unclear, since fidelity is based upon the comparison between matched and mismatched dNTP incorporations (6,7). The thermodynamic instability inherent to mismatched ternary complexes makes them difficult to crystallize. Even though a few crystal structures containing mismatched DNA have been reported, including *Bacillus* DNA polymerase I fragment with template-primer DNA containing mismatches at the 3′-end of the primer stand (8), Pol β with nicked DNA containing

*To whom correspondence should be addressed. Tel: 886 2 2789 9930 ext. 341; Fax: 886 2 2789 8811; Email: tsai@chemistry.ohio-state.edu
Correspondence may also be addressed to Kuo-Hsiang Tang. Tel: 614 292 6771; Fax: 614 292 6773; Email: jtang@chemistry.ohio-state.edu

†The coordinates of free I260Q have been deposited to PDB (2VAN).

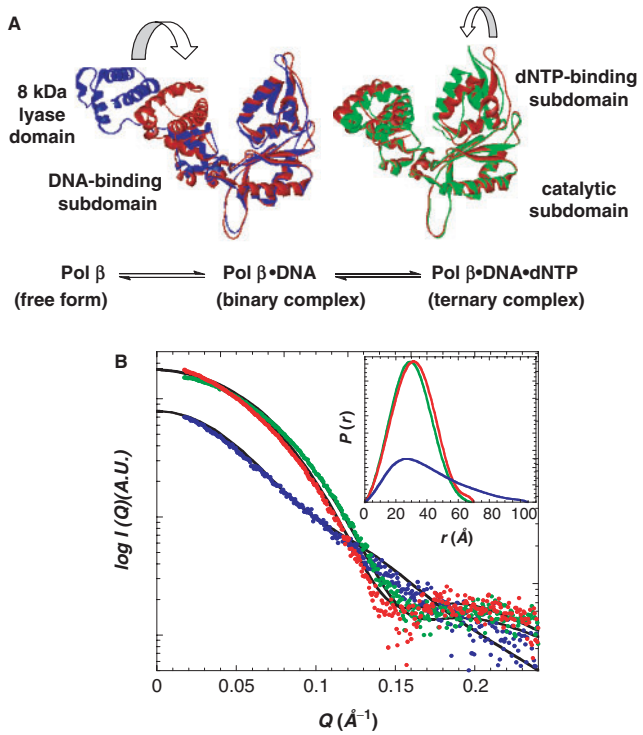

Figure 1. Conformational changes of Pol β. (**A**) Binding of DNA to free Pol β (1BPD, blue) causes rotation of the N-terminal 8 kDa domain in the Pol β·DNA binary complex (1BPX, red) (left), while subsequent binding of matched nucleotide (dNTP) induces the 'open-to-closed' conformational change of the dNTP-binding subdomain in the matched Pol β·DNA·dNTP ternary complex (1BPY, green) (right). For clarity, only the protein moiety in the structures is shown. (**B**) Experimental SAXS data from the free Pol β (blue), binary complex (red) and matched ternary complex (green). The lines represent the calculated SAXS pattern computed from crystal structures fit to the experimental SAXS data. The distance distribution plots computed from the experimental data is shown in the inset. The DNA substrate is the 16/10dd/5-mer 1nt gapped DNA (see Materials and Methods section).

mismatches (9) and ternary complexes of Pol β (10) and Dpo4 (11) accommodating mismatched extension (mutagenic intermediate), none represent a truly functional enzyme·DNA·dNTP ternary complex with mismatched dNTP occupying the dNTP insertion site.

In order to overcome the lack of structural information surrounding polymerase-mismatched dNTP incorporation, we employed small-angle X-ray scattering (SAXS) and molecular modeling to examine the conformational changes of Pol β upon mismatched dNTP incorporation and the results are described in this report. Although SAXS has been used previously to examine the conformations of free Klenow fragment of *Escherichia coli* DNA polymerase I (KF) and *Thermus aquaticus* (*Taq*) DNA polymerase (12), it has not yet been used to probe the ternary complex of a DNA polymerase. We first showed that SAXS can monitor the conformational states of Pol β in agreement with available crystal structures of free Pol β, the Pol β·DNA binary complex, and the Pol β·DNA·dNTP matched ternary complex. SAXS and molecular modeling were then used to monitor the structural change occurring during mismatched dNTP incorporation. Furthermore, we performed parallel studies on the error-prone I260Q of Pol β, which was first identified by a genetic screen (13) and later characterized to show loose binding discrimination among dNTP substrates (14). Structural comparisons between wild-type (WT) Pol β and its I260Q variant allowed us to suggest that the lack of fidelity of I260Q originates from enhanced stabilization of the mismatched ternary complex.

MATERIALS AND METHODS

Materials

The samples of WT Pol β and I260Q variant of Pol β in this study were over-expressed in *E. coli* and purified as described previously (15) with minor modifications. The concentration of the WT Pol β and I260Q was determined using the extinction coefficient $\varepsilon_{280} = 2.12 \times 10^4\,M^{-1}cm^{-1}$ and the Bradford assay. DNA oligomers were from IDT (Coralville, IA, USA). The dNTPs were from Amersham Biosciences (Pittsburg, PA, USA).

DNA substrates

Studies of WT Pol β and I260Q binary and ternary complexes used (i) a 16/10dd/5-mer 1 nt gapped DNA composed of: primer, 5′-GCTGATGCGC$_{dd}$-3′; template, 5′-CCGACGGCGCATCAGC-3′ and downstream oligomer, 5′-pGTCGG-3′ and (ii) a 23-mer double-hairpin 1 nt gapped DNA with the sequence 5′-pGGCGAAGCC GGGTGCGAAGCACC$_{dd}$-3′. Among the sequences, an underlined template base denotes the templating residue, 5′-p stands for 5′-phosphorylated and 3′-dd stands for 3′-dideoxyterminated. DNA substrates were prepared as reported previously (16).

Sample preparation

Similar reaction buffer, 50 mM MOPS and 100 mM KCl (pH 7.0), was applied in the SAXS and SANS measurements with some variations: 10 mM DTT and 10% glycerol were added in SAXS measurements, whereas 1 mM DTT and no glycerol were included in SANS measurements. The binary and ternary complex of Pol β was prepared by mixing Pol β (2–16 mg/ml for SAXS, 2–5 mg/ml for SANS) with equivalent amount of DNA (binary complex, with or without 10 mM Mg^{2+}) and DNA plus 10 mM dNTP/Mg^{2+} (ternary complex). Our previous studies indicated that under these conditions the binary and ternary complexes of Pol β are in the monomeric form, and that the influence of the 2:1 complex is negligible (16). Analytical ultracentrifugation and concentration-series measurements by SAXS were performed to investigate nonspecific associations (16).

SAXS

SAXS experiments were conducted at BL 4-2 of the Stanford Synchrotron Radiation Laboratory (SSRL) (17) and at BioCAT 18-ID of the Advanced Photon Sources (APS) (18). All of the SAXS data presented in this

report were collected at 20 ± 0.5°C. Data collection and experimental condition optimization are described in Supplementary Data. The optimized buffer condition for SAXS studies was 50 mM MOPS, 10 mM DTT, 10% glycerol and 100 mM KCl (or NaCl) at pH 7.0. To ensure the formation of the binary and ternary complexes of Pol β, we performed DNA and dNTP titration studies and only the data obtained under saturating conditions are presented.

In the comparison of SAXS experimental patterns for binary and various ternary complexes of WT Pol β and I260Q variant, because several factors may contribute to the scattering patterns in the intermediate-to-high *Q*-range data, one should be cautious when interpreting the relative differences in the scattering patterns in the region. In addition to confirming the changes with molecular modeling studies (described in Results section), multiple repetitions with different sample concentrations (2, 4, 8 and 16 mg/ml) were performed to ensure the results being reproducible, suggesting that the experimental scattering differences we observed are presumably not due to the difference in the background subtraction. Different dNTP (dATP, dTTP, dCTP or dGTP) should contribute very similarly to the background. The SAXS data analyses and modeling were described previously (16,19).

To analyze the experimental SAXS data in the low-to-intermediate *Q*-range, data over the interval of 0.04–0.30 $Å^{-1}$ were compared to the predicted SAXS pattern calculated from the crystal structure or the intercalated structure by the program CRYSOL (20). The latter were generated by molecular modeling approaches described in this report. CRYSOL calculates the theoretical SAXS pattern from the atomic coordinates of a crystal structure and then alters the volume and hydration contrast parameters to obtain the best fit to the experimental SASX data. The level of confidence in the data fit was determined by a chi-square analysis as indicated in the program (20).

Small-Angle Neutron Scattering

Small-Angle Neutron Scattering (SANS) studies were conducted at the NCNR using the 30 m NG7 SANS instrument as described in Supplementary Data. The SANS data for the free form, binary and matched ternary complexes of Pol β acquired in 0, 70 and 95% D_2O buffer for contrast variation studies are shown in Supplementary Figure S1.

Crystallization and data collection of free form of I260Q

I260Q was crystallized by hanging-drop vapor diffusion technique at 20 ± 0.2°C with a reservoir solution of 14–18% PEG 3350, 100 mM Bis–Tris at pH 6.5. Crystals appeared in ~5–7 days and were cryoprotected by soaking in precipitant solution with 20% glycerol included and plunging into liquid nitrogen. Diffraction images were collected at 100 K on an R-Axis IV++ with a Rigaku MicroMax007 HF (Rigaku, Tokyo), and were integrated and reduced with the HKL2000 software suite (HKL Research, Inc., Charlottesville, VA, USA) (21). Data collection statistics are summarized in Supplementary Table S1.

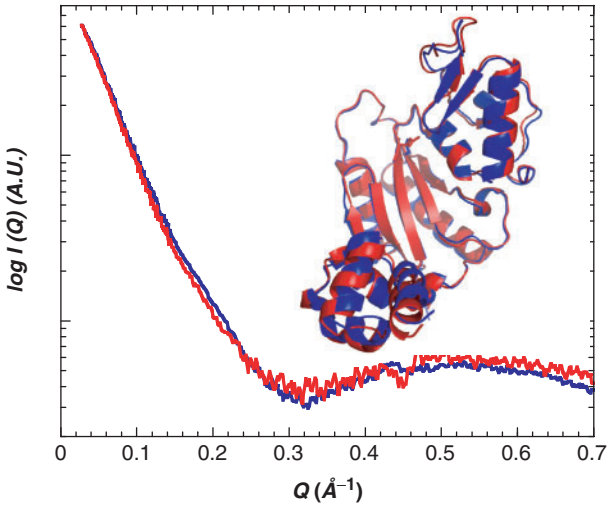

Figure 2. SAXS and crystal structures of WT Pol β and I260Q variant of Pol β in the free form. The merged SAXS curves of WT Pol β (blue) and I260Q (red) in the free form are nearly identical in a large *Q*-range. The corresponding crystal structures (1BPD, blue, from Sawaya *et al.* (2) and 2VAN, red, from this work, respectively) are shown in the inset. The orientation of is chosen to give the best view for comparison. The expanded stereoview is shown in Figure S3B.

Structure determination and refinement of free I260Q

The structure of I260Q (residues 91–335) was determined by molecular replacement from the structure of WT Pol β (1BPD) and was refined with CNS (22). Loss of the N-terminal 8 kDa domain occurred during crystallization, which also occurred in most of the crystal structures of free WT Pol β (1NOM, 1ZQU, 1ZQV, 1ZQW, 1ZQX, 1ZQY, 1ZQZ and 2BPC) (23). The model was obtained by annealing with torsion-angle dynamics with an amplitude maximum-likelihood target (24) for all data to 2.1 Å in CNS. The model was improved by subsequent refinement using CNS and model building using the program XtalView/Xfit (25) until the R/R_{free} value was reduced to 21.7%/24.9%. The structural statistics are summarized in Supplementary Table S1. The structural comparisons of free I260Q and free WT are shown in Figure 2 and Supplementary Figure S3B.

Molecular modeling in the ternary complex of Pol β

While the overall root mean square difference (RMSD) is over 2 Å between the two sets of corresponding C_α atoms from the conformations of the binary (1BPX) and the matched ternary complex (1BPY), the pairwise RMSDs for the subdomain regions are all <1 Å. Therefore, each of the sub-domains can be treated as a rigid body upon dNTP-induced conformational change. To develop intercalated structures between the two end structures, we use the approach described subsequently. The structures of the DNA-binding and the dNTP-binding subdomains,

due to the assumption that these regions can be treated as rigid bodies, can be developed using the following transformation matrix:

$$T = \begin{pmatrix} \cos\alpha\cos\beta & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma & x_t \\ \sin\alpha\cos\beta & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma & y_t \\ -\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma & z_t \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where x_t, y_t, z_t are the three translations and γ, β, α are rotational angles with respect to the x-, y- and z-axes, respectively. From the two end structures, the set of corresponding six parameters (x_t, y_t, z_t, γ, β, α) can be derived. The structures of the DNA-binding and the dNTP-binding subdomains in the intercalated conformation can be calculated using the set of parameters $\rho(x_t, y_t, z_t, \gamma, \beta, \alpha)$, with ρ ranging from 0 to 1 to represent structures at the two ends. The structures of the connecting loops (residues 87–90 and 272–275) were generated using an algorithm developed previously that allows the constructing of both protein and RNA loop structures with known ends (26,27).

Molecular modeling in the binary and ternary complexes of I260Q

The I260Q binary structure was generated using the homology modeling approach (28) that includes an energy minimization using the simulation program AMBER (29). The I260Q ternary structure was first generated using the same approach as described in Refs (28) and (29). To explore the potential interactions between Gln^{260} and other residues in the molecule, an annealing procedure was applied. The annealing procedure consists of a minimization followed by a molecular dynamics simulation and minimization using AMBER. Both positioning constraints (for C_α atoms to maintain the protein fold) and distance constraints (for the targeted hydrogen bonds between Gln^{260} and Glu^{295}/Tyr^{296}) were applied during the annealing. After the annealing procedure, side-chain structures of several residues have changed significantly. For generating the intercalated structures between the binary and the ternary complex structures, the side-chain torsional angles were calculated for the two end-structures and the corresponding side-chain torsional angles were scaled and used for generating the side-chain structures accordingly.

RESULTS

Monitoring conformational changes of Pol β by SAXS

We first investigated whether SAXS could discriminate between the different conformational states assumed during the catalytic cycle of Pol β. To appropriately compare the conformational changes shown in Figure 1A, we used the identical DNA substrate as for reported crystal structures (3), the 16/10dd/5-mer 1nt gapped DNA (see Materials and Methods section), in the SAXS measurements shown in Figure 1B. Under the reaction condition reported in this article (10 mM Mg^{2+}, 50 mM MOPS, 100 mM KCl at pH 7.0), the catalytic reaction of the enzyme displayed two-phase stopped-flow fluorescence changes consistent with our previous observations (4).

The experimental SAXS data for free Pol β, Pol β·DNA binary and Pol β·DNA·dNTP ternary complexes are shown in Figure 1B (scattering intensity I versus momentum transfer Q; $Q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle and λ is the wavelength). Formation of the Pol β·DNA binary complex results in smaller values of the radius of gyration (R_g) and the maximum particle dimension (D_{max}) (Table 1). The particle distance distribution function [P(r)] of different states of Pol β (Figure 1B inset) indicates that free Pol β has an elongated shape (R_g = ~29 Å; D_{max} = ~ 105 Å), whereas the binary complex is in a more globular conformation (R_g = ~24 Å; D_{max} = ~70 Å). Subsequent binding of matched dNTP to the binary complex further decreases R_g and D_{max} values, indicating formation of a more compact conformation in the matched ternary complex relative to the binary complex. Similar trend is also observed in contrast variation by SANS (see Supplementary Data), where R_g values for the free form, binary complex and ternary

Table 1. Parameters[a] of WT Pol β and I260Q and their complexes determined by SAXS

Sample	$R_g^{(exp\,tl)}$(Å) (Guinier)	$R_g^{(exp\,tl)}$ (Å) (GNOM)	$R_g^{(calc.)}$ (Å) (CRYSOL)	$D_{max}^{(exp\,tl)}$ (Å) (GNOM)
Pol β	28.7 ± 0.1	29.1 ± 0.1	28.9 (1BPD)	105 ± 2
Pol β·DNA[b]	24.0 ± 0.2	23.8 ± 0.2	23.4 (1BPX)	70 ± 3
G:G mismatched Pol β·DNA·dNTP[b]	23.6 ± 0.2	23.4 ± 0.2	NA[g]	70 ± 2
G:C matched Pol β·DNA·dNTP[b]	23.0 ± 0.1	22.8 ± 0.1	22.4 (1BPY)	65 ± 2
Free I260Q	29.0 ± 0.3	29.4 ± 0.2	NA[c]	105 ± 3
I260Q·DNA[d]	23.6 ± 0.1	23.4 ± 0.2	NA[c]	70 ± 4
G:C matched and G:G mismatched I260Q·DNA·dNTP[d]	23.0 ± 0.2	23.0 ± 0.2	NA[c]	67 ± 2

[a] R_g: the radius of gyration; D_{max}: the maximum particle dimension.
[b] The experimental values of R_g and D_{max} were obtained from both 16/10dd/5-mer DNA and 23-mer double-hairpin 1nt gapped DNA.
[c] NA: not available; due to lack of the corresponding high-resolution structures.
[d] The experimental values of R_g and D_{max} were obtained with 23-mer double-hairpin 1nt gapped DNA.

complex of Pol β in 70% D_2O buffer, in which only protein signal is visible, are 29.0 ± 0.2, 24.7 ± 0.7 and 23.5 ± 0.4 Å, respectively.

Using the program CRYSOL (20) we generated the predicted SAXS patterns of Pol β and its complexes and calculated R_g and D_{max} values from the atomic coordinates of the corresponding crystal structures (2,3). The calculated SAXS patterns correlate well with corresponding experimental data (Figure 1B); and the experimental and calculated R_g and D_{max} values (Table 1) are within experimental error. Furthermore, reconstructed models were generated from the experimental SAXS data using the modeling programs GASBOR (for free form of Pol β) (30) and DAMMIN (for free form and complexes of Pol β) (31), and the structural models of Pol β and its complexes closely resemble the corresponding atomic-resolution structures (Supplementary Figure S2). Together, the data indicate that the polydispersity and heterogeneity of Pol β and its complexes are insignificant for the samples used in this report.

No disruption in the global structure of Pol β upon I260Q modification

To provide a fair comparison between the functional properties of WT and I260Q variant, it is important to show that no global structural perturbation is induced by substitution of Ile^{260} with Gln. Protein concentration series were performed for I260Q, and no aggregation was detected as indicated in the Guinier plot (Supplementary Figure S3A inset). Figure 2 shows only slight deviations beyond $Q > 0.3$ Å^{-1} in the SAXS experimental curves of I260Q relative to that of WT, which can be partially correlated with the differences in local rather than global conformations between these two proteins. Table 1 shows similar R_g and D_{max} values, suggesting that there is no disruption in the global conformation. Additionally, the crystal structure of free I260Q was also acquired. As shown in Figure 2 and Supplementary Figure S3B, the overall structure of free I260Q is similar to that of free WT. The RMSD values between I260Q (2VAN) and WT Pol β (1BPD) for C_α atoms are 1.19, 0.85 and 1.09 Å, respectively, for the following three segments: the DNA-binding/catalytic/dNTP-binding subdomains (residues 91–143, 148–243, 249–335), the DNA binding/catalytic subdomains (residues 91–143, 148–243) and the catalytic/dNTP binding subdomains (residues 148–243, 249–335). The slight hinge movement of the DNA-binding and the dNTP-binding subdomains toward outside the active site cleft is possibly due to the loss of the N-terminal lyase domain and/or modification of residue 260.

SAXS studies of I260Q and mismatched ternary complexes

SAXS was then used to compare the conformational changes induced by matched and mismatched dNTP binding to E·DNA binary complex of WT Pol β and I260Q. Although the dNTP-induced conformational change of Pol β is not as large as that induced by DNA binding, Figures 1B and 3A suggests that there are differences in the SAXS pattern of the binary and ternary complexes in the Q-range between 0.12 and 0.24 Å^{-1}. Also, obtaining better S/N at wider angles requires more concentrated samples (~0.4 mM), so we compared the data in the studies of binary and matched and mismatched ternary complexes only for $Q \geq 0.10$ Å^{-1}, which is less affected by inter-particle interaction effects. The SAXS measurements shown in Figure 3 were performed with the 23-mer double-hairpin 1nt gapped DNA (see Materials and Methods section), because the data quality is better with this DNA substrate compared to the 16/10dd/5-mer 1nt gapped DNA (Supplementary Figure S6). Conformational change of Pol β upon the 23-mer double-hairpin DNA binding and (matched) dNTP-induced conformational changes of the Pol β–DNA binary complex with the 23-mer double-hairpin DNA were suggested by NMR chemical shift perturbation (32). It is possible that the interactions of Pol β with the 23-mer DNA substrate lead to a more stable complex in solution, thereby the SAXS patterns of the binary and various ternary complexes can be better distinguished in the Q-range between 0.12 and 0.24 Å^{-1} with this DNA.

Figure 3 shows the experimental data with 16 mg/ml (~0.4 mM) WT Pol β and I260Q. Figure 3B shows the experimental SAXS patterns of the matched (G:C) and the mismatched (G:A, G:T or G:G) ternary complexes of WT Pol β, and the patterns are well-resolved in the Q-range between 0.12 and 0.24 Å^{-1}. It is interesting to note that the pattern of the G:T (a wobble base pair) mismatched is slightly closer to the G:C matched compared to G:A and G:G mismatched ternary complexes (Figure 3C), which could indicate that less destabilization occurs in the G:T ternary complexes than in the G:A or G:G ternary complex (see Discussion section). Further, all of the WT mismatched complexes give SAXS patterns that lie in-between those of the binary complex and the matched ternary complex (Figure 3B and C) as suggested in the predicted SAXS patterns (described in the next section and also see Figure 4). Similar trend is also observed for the I260Q complexes as for the WT complexes. For clarity, only the data for the G:G mismatched ternary complex of WT and I260Q are shown in Figure 3D.

The SAXS patterns in Figure 3D indicate the following significant points: (i) For both WT and I260Q, the profile of the mismatched ternary complex lies in-between those of the binary complex and the matched ternary complex, but is more similar to the binary complex for WT and to the matched ternary complex for I260Q. These results suggest that WT mismatched ternary complexes exist in 'partially closed' (or partially open) conformations along the pathway of the conformational closing of the dNTP-binding subdomain. Additionally, 2D-[^{1}H, ^{15}N]-TROSY-HSQC NMR spectra (Supplementary Figure S4) reveal a similar qualitative trend. The chemical shifts in the NMR spectrum of the binary complex changed insignificantly when a mismatched dNTP was added to generate the G:G mismatched ternary complex (Supplementary Figure S4C), relative to the changes in the G:C matched ternary complex (Supplementary Figure S4A). (ii) The SAXS profiles of I260Q in the binary and matched

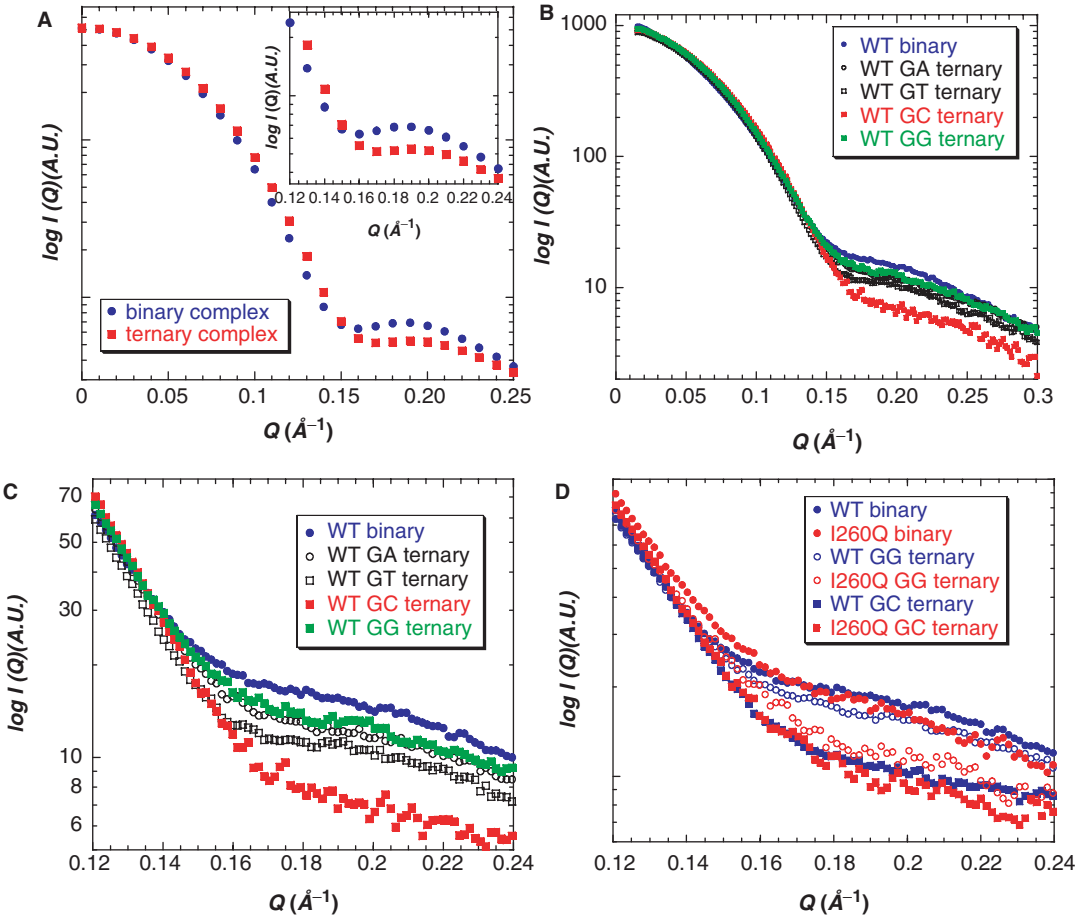

Figure 3. SAXS measurements of WT Pol β and I260Q variant of Pol β. (**A**) The predicted SAXS patterns of the binary (open form) and matched ternary (closed form) complexes calculated from the crystal structures. The data in the Q-range between 0.12 and 0.24 Å^{-1} is shown in the inset. (**B**) The experimental SAXS data for the binary (filled circle), mismatched ternary [G:A (open circle), G:T (open square) and G:G (filled square)] and matched (G:C, filled square) ternary complexes of WT Pol β in the Q-range up to 0.30 Å^{-1}. (**C**) Same as panel B, only the data in the Q-range between 0.12 and 0.24 is shown. (**D**) The experimental SAXS data for the binary, mismatched (G:G) ternary and matched (G:C) ternary complexes of WT Pol β (blue) and I260Q (red) in the Q-range between 0.12 and 0.24 Å^{-1} is shown. The DNA substrate is the 23-mer double-hairpin 1nt gapped DNA (see Materials and methods section).

ternary complexes differ from those of the corresponding complexes of WT. The data suggest that the conformations of the I260Q binary and matched ternary complexes are slightly more compact than the corresponding complexes of WT. Full understanding of these small differences will require high-resolution structural studies. The present study focuses on the differences between matched and mismatched ternary complexes for both WT and I260Q.

Further interpretation of SAXS data by molecular modeling

We used molecular modeling to support the possibility that the mismatched ternary complex of WT lies on the conformational path between the open and the closed forms. Figure 4A (inset) shows a set of four intercalated structures (gray) between the open crystal structure (1BPX, blue) and the G:C matched ternary complex crystal structure (1BPY, red). The structures of the 23-mer double-hairpin 1nt gapped DNA in both the open and the closed complexes were modeled (Supplementary Figure S5B) according to structure of the 16/10dd/5-mer DNA (Supplementary Figure S5A) in either 1BPX or 1BPY (3). The rigid body approaches for generating the modeled structures are described in Materials and methods section. Analogous to the experimental SAXS patterns (Figure 3), the predicted SAXS curves calculated from the crystal and modeled structures by CRYSOL (20) can be best distinguished in the Q-range between 0.12 and 0.24 Å^{-1} (Figure 4A and B). Figure 4C indicates that the experimental SAXS curve of the WT G:G mismatched ternary complex shown in Figure 3B fits well with the predicted SAXS curve (the green curve in Figure 4B) calculated from the first intercalated structure (chi-squared value for the fit was 3.02), which is ~25% on the conformational path between 1BPX and 1BPY. The chi-squared values for the fit to neighboring curves are substantially higher (>7.0).

As crystal structures are unavailable for complexes of I260Q, the coordinates of the binary (1BPX) and matched ternary (1BPY) complexes of the WT were used for the homology modeling approach (28). An annealing procedure was used to generate the structures of the binary and ternary complexes of I260Q, which were subsequently used to generate intercalated structures as shown in Figure 4D.

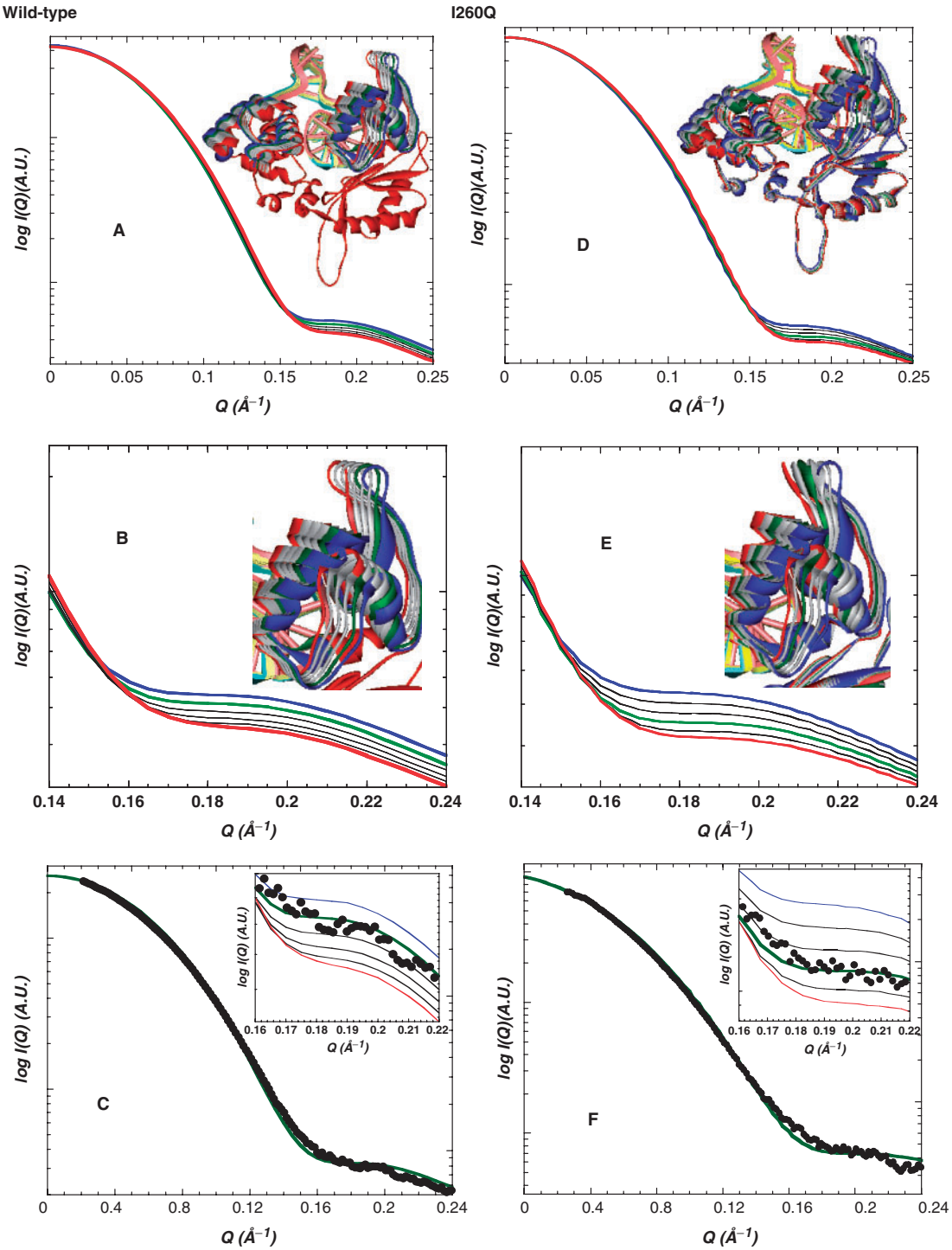

Figure 4. SAXS and the molecular modeling studies of the binary and ternary complexes of WT Pol β (panels A–C) and I260Q (panels D–F). (**A**) A set of intercalated structures (inset), including the binary complex (blue), mismatched ternary complex (green) and matched ternary complex (red) of WT Pol β, along with the predicted SAXS patterns, including the binary (blue), mismatched (green) and matched ternary (red) complexes, calculated from the crystal and intercalated structures shown in the Q-range up to 0.25 $Å^{-1}$. (**B**) Same as panel A, only the structures of the dNTP-binding subdomain (inset) and the predicted SAXS pattern of WT complexes in the Q-range between 0.12 and 0.24 $Å^{-1}$ are shown. (**C**) Superimposition of the experimental SAXS data (filled circle) for the G:G mismatched complex of WT Pol β in Figure 3B and the calculated SAXS patterns in panel B. The experimental data fit best with the first intercalated curve (the green SAXS curve in panel B). Chi-squared value for the goodness of fit was 3.02. The data fit in the Q-range between 0.16 and 0.22 $Å^{-1}$ are shown in the inset. (**D**) Similar to panel A, the modeled structures (inset) and predicted SAXS patterns, including the binary (blue), G:G mismatched ternary (green) and matched ternary (red) complexes of I260Q, calculated form the modeled structures in the Q-range up to 0.25 $Å^{-1}$ are shown. (**E**) Same as panel D, only the structures of the dNTP-binding subdomain (inset) and the predicted SAXS pattern of I260Q complexes in the Q-range between 0.12 and 0.24 $Å^{-1}$ are shown. (**F**) Superimposition of the experimental SAXS data (filled circle) for the G:G mismatched complex of I260Q in Figure 3C and the calculated SAXS patterns in panel E. The experimental data fit best with the first intercalated curve (the green SAXS curve in panel E). Chi-squared value for the data fit was 4.87. The data fit in the Q-range between 0.16 and 0.22 $Å^{-1}$ are shown in the inset.

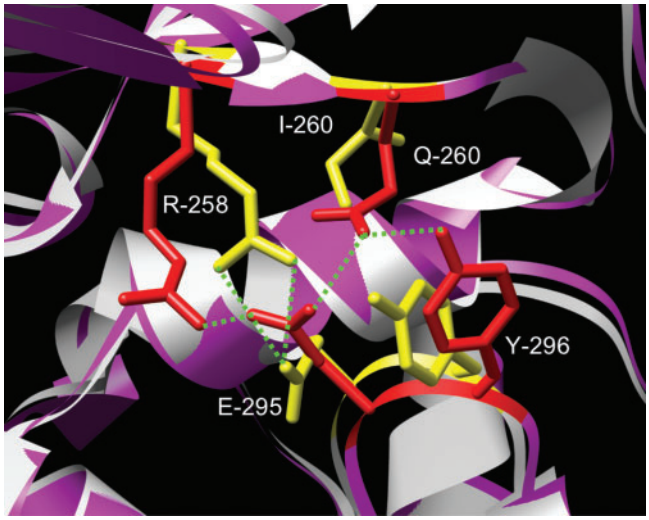

Figure 5. The relative positions of Asp^{192}, Arg^{258}, Ile^{260} (or Gln^{260} in I260Q), Glu^{295} and Tyr^{296} at the active site of the crystal structure of WT Pol β matched ternary complex (residues shown in yellow) and modeled structure of the I260Q matched ternary complex (residues shown in red).

Similar to the studies of WT Pol β, the predicted SAXS curves calculated from the modeled structures of the I260Q complexes can be also differentiated in the *Q*-range between 0.12 and 0.24 $Å^{-1}$ (Figure 4D and E). Figure 4F shows that the experimental SAXS curve of the I260Q G:G mismatched ternary complex shown in Figure 3D fits reasonably well (chi-squared value for the fit was 4.87) with the predicted SAXS patterns (the green curve in Figure 4E) calculated from the third intercalated structure, which is ~60% on the conformational path between the binary and matched ternary complexes. The chi-squared values for the fit to neighboring curves are higher (>7.0). The 60% homology model is consistent with the experimental observation. While we cannot prove that the 60% model structure is the definitive structure for the I260Q complex, it is a good estimate of the I260Q G:G mismatched complex structure. Examination of the modeled structures of I260Q indicates that the closed form of I260Q could be stabilized by additional hydrogen bonds involving Gln^{260}, Glu^{295}, Tyr^{296} and Arg^{258} as shown in Figure 5.

DISCUSSION

This work addresses two major questions: how WT Pol β differentiates matched from mismatched dNTP, and how a single I260Q mutation turns Pol β into an error-prone polymerase. For some enzymes it may not be essential to investigate the mechanism of reaction of poor substrate analogs. For DNA polymerases this is an interesting and important issue for the following reasons: (i) mismatched dNTPs are also natural nucleotides that will compete with the matched dNTP; (ii) to achieve high fidelity, a high-fidelity polymerase not only has to have high activity for the matched dNTP, but also has to avoid the incorporation of mismatched dNTP and (iii) it is equally important to understand how the mismatched dNTP can be incorporated by an error-prone polymerase. As we have indicated previously (7), understanding of the mechanism of fidelity of DNA polymerases requires understanding of the mechanism of mismatched dNTP incorporation in addition to that of matched dNTP incorporation.

It was recently proposed by several research groups (9,33–37) that a perturbed conformation of the ternary complex or even a different reaction pathway may be engaged by a mismatched dNTP incorporation in DNA polymerases. X-ray structural analysis of Pol β with mismatched base pairs situated in the polymerase active site suggested that the mismatched dNTP insertion could occur through partially open conformation (9). Our SAXS and molecular modeling studies presented here also indicate that the conformation of the mismatched ternary complex of WT Pol β lies on the conformational pathway from the open to the closed forms. Alternatively, fluorescence studies of KF (35) and T7 DNA polymerase (34) suggested that the mismatched dNTP incorporation occurs through a distinct conformation which does not lie on the 'open-to-closed' pathway.

Our results in Figure 4 indicate that the conformation of the mismatched ternary complex of WT Pol β is close to that of the binary complex (open form), and that the conformation of the mismatched ternary complex of I260Q is akin to the matched ternary complex (closed form) of WT Pol β and close to the matched ternary complex (closed form) of I260Q. The significance of such structural information in the catalytic mechanism of WT and I260Q are further addressed in the next two sections. Together, our studies suggest that a common conformational pathway is followed by matched and mismatched dNTP incorporation by WT Pol β and its error-prone I260Q variant. The similar results are also suggested in the recent kinetic studies (unpublished data in Tsai laboratory). The putative 'misaligned' state, distinct from the open and closed states, suggested in the recent report of T7 DNA polymerase (34) is not detected in our studies of WT Pol β and I260Q. It remains to be established whether such discrepancy reflects differences in the mismatched discrimination mechanism employed by two polymerases, or whether it results from different experimental systems and conditions. As a way of reconciliation, one can consider the WT mismatched ternary complex as existing in a 'local misaligned state' in the same global conformational path. The possible structural distinctions associated with the mismatched complex that SAXS cannot resolve will require high-resolution structural studies.

Figure 6 shows the current knowledge of the kinetic scheme of matched dNTP incorporation catalyzed by Pol β (7). All of the intermediate species of this scheme have been characterized structurally and/or kinetically, but not for mismatched dNTP incorporation. The use of mismatched dNTP, error-prone I260Q of Pol β and solution structural studies by SAXS allowed us to characterize the intermediate species in the catalytic reaction of mismatched dNTP. While the SAXS data in Figure 3 suggest that the G:G mismatched ternary complex of WT is ~25% on the conformational pathway from the open binary form to the closed ternary

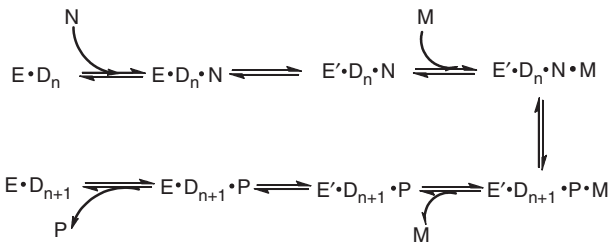

Figure 6. Kinetic scheme for single-nucleotide incorporation by Pol β. E·D_n = Pol β·DNA binary complex; E: Pol β in open conformation; E′: Pol β in closed conformation; D_n/D_{n+1}: DNA before/after the nucleotide is incorporated; *N*: MgdNTP; E·D_n·*N*: open Pol β·DNA·dNTP ternary complex; E′D_n·*N*: closed Pol β·DNA·dNTP ternary complex; M: catalytic Mg^{2+}; P: $MgPP_i$.

form, this alone cannot rule out an alternative interpretation—that the G:G mismatched ternary complex consists of a 3:1 mixture of open and closed conformations which are in equilibrium. However, the former explanation fits well with the results of molecular modeling (Figure 4). Furthermore, the qualitative structural information provided by chemical shift perturbation in the NMR spectra indicate that most of the peaks that are shifted in Supplementary Figure S4A (the G:C matched ternary complex) remain unshifted or only slightly shifted in Supplementary Figure S4C (the G:G mismatched ternary complex). On the other hand a few peaks in the G:G ternary complex are shifted to different directions, giving a pattern inconsistent with the possibility that the mismatched ternary complex represents a mixture of 1/4 of the closed form and 3/4 of the open form. Since both NMR and SAXS were performed under saturating concentrations of dNTP and Mg^{2+} (10 mM of each), lack of saturation should not be a cause for the absence of changes in the NMR spectra of mismatched ternary complex.

Unlike Pol β and many other high-fidelity polymerases, the low fidelity Y-family DNA polymerase DinB has been shown to exist in a closed conformation without binding of DNA and dNTP (38,39). Pol λ, another family X DNA polymerase, showed no significant conformational change from the free form to the ternary complex (40). These studies raised the possibility that the subdomain closing conformational change may not be required for error-prone enzymes, or that the lack of it could be a cause for the low fidelity. This hypothesis was adopted to explain the low-fidelity of I260Q variant (14), as the authors proposed that the I260Q variant hinders the dNTP-binding subdomain closure, therefore its dNTP insertion accuracy is lower.

However, our results indicate the contrary—the main difference leading to the low fidelity of I260Q appears to be a reduction in ability to destabilize the mismatched ternary complex. Based on the crystal structures of the binary (1BPX) and matched ternary (1BPY) complex of Pol β, the modification of Ile by Gln at residue 260 should not interfere with the dNTP-binding subdomain closure. Instead, residue 260 is in the proximity of Glu^{295} and Tyr^{296}. As demonstrated in Figure 5, when residue 260 is Gln (polar) instead of Ile (non-polar), it has the potential to form hydrogen bonds with Glu^{295} and Tyr^{296} in the closed from of the ternary complex, freeing up Asp^{192} to coordinate MgdNTP. The distances are greater between these residues in the open form, which explains why the free form of I260Q would not pre-exist in the closed form. If these hydrogen bonds form as predicted by the molecular modeling studies, the I260Q ternary complex will prefer to be in the closed form. Thus, the I260Q will change the conformational preference of the molecule. As predicted, our SAXS results indicate that I260Q undergoes conformational changes for both matched and mismatched dNTP incorporations.

CONCLUSION

We report herein our studies on the structural mechanism of mismatched dNTP incorporations by WT Pol β and I260Q. The analyses would not be possible if mismatched dNTP led the enzyme to proceed by a globally different conformational path. Thus, the elucidation of the mechanism of misincorporation described in this study has enhanced our understanding of the mechanism of fidelity of Pol β, and also affirms the conclusions of our previous studies on this enzyme (7). Full understanding of the mechanism of fidelity requires comparison of the kinetic schemes of intermediate structures between the reactions of matched and mismatched dNTP. While the reactions of matched dNTP have been studied extensively, the present study represents a first step toward detailed understanding of the reactions of mismatched dNTP.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Liang Guo (APS) for SAXS data collection at BioCAT 18-ID, Dr Ann Aulabaugh (Wyeth Research Biophysics/Enzymology-Screening Science) for SV, Dr Susan Krueger and Dr Boualem Hammouda (NCNR) for assisting SANS measurements and Michelle Roetteger and Dr Marina Bakhtina for useful discussion. This work was supported by NIH Grant GM43268, NSC Grant 95-2745-B and GRC intramural funding (to M.-D.T.). K.-H.T. was partially supported by the Academia Sinica Postdoctoral Fellowship. SAXS studies were conducted at BL 4-2 of the SSRL and BioCAT 18-ID of the APS, supported by the NIH and the DOE. SANS were conducted at the NCNR using the 30 m NG7 SANS instrument. X-ray diffraction data was performed at GRC and NSRRC (Taiwan). NMR experiments were performed at the High-Field Biomacromolecular NMR Core Facility supported by the National Research Program for Genomic Medicine of Taiwan. This is paper 15 in the series 'DNA Polymerase β' from Tsai laboratory. For paper 14, see Ref. [16]. Funding to pay the Open Access publication charges for this article was provided by the Genomics Research Center.

Conflict of interest statement. None declared.

REFERENCES

1. Beard,W.A. and Wilson,S.H. (2006) Structure and mechanism of DNA polymerase Beta. *Chem. Rev.*, **106**, 361–382.
2. Sawaya,M.R., Pelletier,H., Kumar,A., Wilson,S.H. and Kraut,J. (1994) Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. *Science*, **264**, 1930–1935.
3. Sawaya,M.R., Prasad,R., Wilson,S.H., Kraut,J. and Pelletier,H. (1997) Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry*, **36**, 11205–11215.
4. Bakhtina,M., Lee,S., Wang,Y., Dunlap,C., Lamarche,B. and Tsai,M.D. (2005) Use of viscogens, dNTPalphaS, and rhodium(III) as probes in stopped-flow experiments to obtain new evidence for the mechanism of catalysis by DNA polymerase beta. *Biochemistry*, **44**, 5177–5187.
5. Dunlap,C.A. and Tsai,M.D. (2002) Use of 2-aminopurine and tryptophan fluorescence as probes in kinetic analyses of DNA polymerase beta. *Biochemistry*, **41**, 11226–11235.
6. Showalter,A.K. and Tsai,M.D. (2002) A reexamination of the nucleotide incorporation fidelity of DNA polymerases. *Biochemistry*, **41**, 10571–10576.
7. Showalter,A.K., Lamarche,B.J., Bakhtina,M., Su,M.I., Tang,K.H. and Tsai,M.D. (2006) Mechanistic comparison of high-fidelity and error-prone DNA polymerases and ligases involved in DNA repair. *Chem. Rev.*, **106**, 340–360.
8. Johnson,S.J. and Beese,L.S. (2004) Structures of mismatch replication errors observed in a DNA polymerase. *Cell*, **116**, 803–816.
9. Krahn,J.M., Beard,W.A. and Wilson,S.H. (2004) Structural insights into DNA polymerase beta deterrents for misincorporation support an induced-fit mechanism for fidelity. *Structure*, **12**, 1823–1832.
10. Batra,V.K., Beard,W.A., Shock,D.D., Pedersen,L.C. and Wilson,S.H. (2005) Nucleotide-induced DNA polymerase active site motions accommodating a mutagenic DNA intermediate. *Structure*, **13**, 1225–1233.
11. Trincao,J., Johnson,R.E., Wolfle,W.T., Escalante,C.R., Prakash,S., Prakash,L. and Aggarwal,A.K. (2004) Dpo4 is hindered in extending a G.T mismatch by a reverse wobble. *Nat. Struct. Mol. Biol.*, **11**, 457–462.
12. Joubert,A.M., Byrd,A.S. and LiCata,V.J. (2003) Global conformations, hydrodynamics, and X-ray scattering properties of *Taq* and *Escherichia coli* DNA polymerases in solution. *J. Biol. Chem.*, **278**, 25341–25347.
13. Starcevic,D., Dalal,S. and Sweasy,J. (2005) Hinge residue Ile260 of DNA polymerase beta is important for enzyme activity and fidelity. *Biochemistry*, **44**, 3775–3784.
14. Starcevic,D., Dalal,S., Jaeger,J. and Sweasy,J.B. (2005) The hydrophobic hinge region of rat DNA polymerase beta is critical for substrate binding pocket geometry. *J. Biol. Chem.*, **280**, 28388–28393.
15. Werneburg,B.G., Ahn,J., Zhong,X., Hondal,R.J., Kraynov,V.S. and Tsai,M.D. (1996) DNA polymerase beta: pre-steady-state kinetic analysis and roles of arginine-283 in catalysis and fidelity. *Biochemistry*, **35**, 7041–7050.
16. Tang,K.H., Niebuhr,M., Aulabaugh,A. and Tsai,M.D. (2008) Solution structures of 2 : 1 and 1 : 1 DNA polymerase-DNA complexes probed by ultracentrifugation and small-angle X-ray scattering. *Nucleic Acids Res.*, **36**, 849–860.
17. Smolsky,I.L., Liu,P., Niebuhr,M., Ito,K., Weiss,T.M. and Tsuruta,H. (2007) Biological small-angle x-ray scattering facility at the Stanford synchrotron radiation laboratory. *J. Appl. Cryst.*, **40**, S453–S458.
18. Fischetti,R., Stepanov,S., Rosenbaum,G., Barrea,R., Black,E., Gore,D., Heurich,R., Kondrashkina,E., Kropf,A.J., Wang,S. *et al.* (2004) The BioCAT undulator beamline 18ID: a facility for biological non-crystalline diffraction and X-ray absorption spectroscopy at the Advanced Photon Source. *J. Synchrotron. Radiat.*, **11**, 399–405.
19. Tang,K.H., Guo,H., Yi,W., Tsai,M.D. and Wang,P. (2007) Investigation of the conformational states of Wzz and the Wzz•O-antigen complex under near-physiological conditions. *Biochemistry*, **46**, 11744–11752.
20. Svergun,D.I., Barberato,C. and Koch,M.H.J. (1995) CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Cryst.*, **28**, 768–773.
21. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
22. Brunger,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. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
23. Pelletier,H. and Sawaya,M.R. (1996) Characterization of the metal ion binding helix-hairpin-helix motifs in human DNA polymerase beta by X-ray structural analysis. *Biochemistry*, **35**, 12778–12787.
24. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. D*, **53**, 240–255.
25. McRee,D.E. (1999) XtalView/Xfit–A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.*, **125**, 156–165.
26. Tung,C.S. (1996) A geometrical approach in folding a pseudoknot motif within the *E. coli* 16S-RNA. *J. Biomol. Struct. Dyn.*, **14**, 153–161.
27. Ryu,K., Gilchrist,R.L., Tung,C.S., Ji,I. and Ji,T.H. (1998) High affinity hormone binding to the extracellular N-terminal exodomain of the follicle-stimulating hormone receptor is critically modulated by exoloop 3. *J. Biol. Chem.*, **273**, 28953–28958.
28. Tung,C.S. (1999) Structural study of homeodomain protein-DNA complexes using a homology modeling approach. *J. Biomol. Struct. Dyn.*, **17**, 347–354.
29. Case,D.A., Cheatham,T.E.,3rd, Darden,T., Gohlke,H., Luo,R., Merz,K.M.,Jr., Onufriev,A., Simmerling,C., Wang,B. and Woods,R.J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.*, **26**, 1668–1688.
30. Svergun,D.I., Petoukhov,M.V. and Koch,M.H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.*, **80**, 2946–2953.
31. Svergun,D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.*, **76**, 2879–2886.
32. Kirby,T.W., DeRose,E.F., Beard,W.A., Wilson,S.H. and London,R.E. (2005) A thymine isostere in the templating position disrupts assembly of the closed DNA polymerase beta ternary complex. *Biochemistry*, **44**, 15230–15237.
33. Joyce,C.M. and Benkovic,S.J. (2004) DNA polymerase fidelity: kinetics, structure, and checkpoints. *Biochemistry*, **43**, 14317–14324.
34. Tsai,Y.C. and Johnson,K.A. (2006) A new paradigm for DNA polymerase specificity. *Biochemistry*, **45**, 9675–9687.
35. Purohit,V., Grindley,N.D. and Joyce,C.M. (2003) Use of 2-aminopurine fluorescence to examine conformational changes during nucleotide incorporation by DNA polymerase I (Klenow fragment). *Biochemistry*, **42**, 10200–10211.
36. Florian,J., Goodman,M.F. and Warshel,A. (2005) Computer simulations of protein functions: searching for the molecular origin of the replication fidelity of DNA polymerases. *Proc. Natl Acad. Sci. USA*, **102**, 6819–6824.
37. Xiang,Y., Oelschlaeger,P., Florian,J., Goodman,M.F. and Warshel,A. (2006) Simulating the effect of DNA polymerase mutations on transition-state energetics and fidelity: evaluating amino acid group contribution and allosteric coupling for ionized residues in human pol beta. *Biochemistry*, **45**, 7036–7048.
38. Zhou,B.L., Pata,J.D. and Steitz,T.A. (2001) Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain. *Mol. Cell*, **8**, 427–437.
39. Silvian,L.F., Toth,E.A., Pham,P., Goodman,M.F. and Ellenberger,T. (2001) Crystal structure of a DinB family error-prone DNA polymerase from *Sulfolobus solfataricus*. *Nat. Struct. Biol.*, **8**, 984–989.
40. Garcia-Diaz,M., Bebenek,K., Krahn,J.M., Kunkel,T.A. and Pedersen,L.C. (2005) A closed conformation for the Pol lambda catalytic cycle. *Nat. Struct. Mol. Biol.*, **12**, 97–98.