Published online 18 June 2018

Nucleic Acids Research, 2018, Vol. 46, No. 14 ***7309–7322***
doi: 10.1093/nar/gky503

Transitions in DNA polymerase β μs-ms dynamics related to substrate binding and catalysis

Eugene F. DeRose, Thomas W. Kirby, Geoffrey A. Mueller, William A. Beard, Samuel H. Wilson and Robert E. London*

Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, USA

Received November 06, 2017; Revised May 15, 2018; Editorial Decision May 18, 2018; Accepted May 24, 2018

ABSTRACT

DNA polymerase β (pol β) plays a central role in the DNA base excision repair pathway and also serves as an important model polymerase. Dynamic characterization of pol β from methyl-TROSY ^{13}C-^{1}H multiple quantum CPMG relaxation dispersion experiments of Ile and Met sidechains and previous backbone relaxation dispersion measurements, reveals transitions in μs-ms dynamics in response to highly variable substrates. Recognition of a 1-nt-gapped DNA substrate is accompanied by significant backbone and sidechain motion in the lyase domain and the DNA binding subdomain of the polymerase domain, that may help to facilitate binding of the apoenzyme to the segments of the DNA upstream and downstream from the gap. Backbone μs-ms motion largely disappears after formation of the pol β–DNA complex, giving rise to an increase in uncoupled μs-ms sidechain motion throughout the enzyme. Formation of an abortive ternary complex using a non-hydrolyzable dNTP results in sidechain motions that fit to a single exchange process localized to the catalytic subdomain, suggesting that this motion may play a role in catalysis.

INTRODUCTION

The base excision repair (BER) pathway contributes to the maintenance of DNA integrity by sensing and removing damaged bases and the repair of apurinic/apyrimidinic (AP) sites. DNA polymerase β (pol β) catalyzes both DNA synthesis (nucleotidyltransferase) and deoxyribose phosphate (dRP) lyase reactions in the BER pathway (1). The nucleotidyltransferase reaction is catalyzed by the 31-kDa polymerase domain of the enzyme while the dRP lyase activity is catalyzed by the 8-kDa amino-terminal lyase domain. The polymerase domain is composed of DNA binding (D), catalytic (C) and nucleotide sensing (N) subdomains, that function in upstream duplex DNA binding, DNA synthesis chemistry, and incoming nucleotide selection, respectively (1–4).

A large number of X-ray crystallographic structures of pol β with a variety of DNA and dNTP substrates have been determined providing a substantial amount of information on the structure and mechanism of pol β (1,2,5). These studies indicate that an elongated form of the enzyme initially re-organizes itself with gapped DNA substrates where the polymerase domain interacts with upstream duplex DNA and the lyase domain interacts with downstream duplex. At this stage, the enzyme is in an open conformation. Upon binding a correct deoxynucleoside triphosphate (dNTP), the N subdomain repositions itself as it transitions from an open binary DNA complex to a closed ternary (i.e., pol β/DNA/dNTP) substrate complex (Figure 1). In the closed complex, the nascent base pair is sandwiched between the growing DNA duplex and the N subdomain (6). Additionally, the amino-terminal lyase domain and carboxyl-terminal N subdomain interact creating a doughnut-like structure. Ligand binding, as well as the open to closed structural change, is expected to alter enzyme sidechains, DNA, and incoming nucleotide conformation/dynamics that can impact polymerase catalysis and fidelity. Since DNA polymerases must select a different/new substrate (DNA and dNTP) with each insertion event, substrate and protein dynamics represent a valuable tactic to optimize each catalytic event. Recent time lapse X-ray crystallographic studies have enabled a more thorough characterization of intermediate structures along the reaction path during correct and incorrect nucleotide insertion (7). In this approach, a stable closed ternary complex crystal is formed with 1-nt-gapped DNA and either a matched or mismatched nucleotide with non-catalytic Ca^{2+} ions, and the reaction is initiated by soaking with Mn^{2+}. Mn^{2+} was used to initiate the reaction because Mg^{2+} does not readily form the closed ternary complex required for catalysis with mismatched nucleotides (8). The *in crystallo* reaction is stopped along the reaction path by rapidly freezing the crystals after various time intervals.

*To whom correspondence should be addressed.Tel: +1 984 287 3573; Fax: +1 301 451 5697; Email: london@niehs.nih.gov

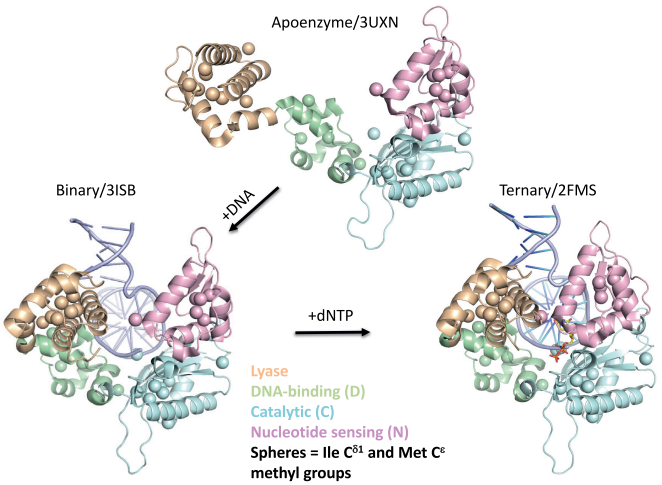

Figure 1. Comparison of X-ray structures of the pol β apoenzyme (3UXN-A (5)), binary complex (3ISB (30)) with 1-nt-gapped DNA, and a ternary complex (2FMS (23)) containing a matched non-hydrolyzable dNTP, dUMPNPP. The structures highlight the transition from fully open apoenzyme to bound and partially closed binary complex to fully closed ternary complex. The lyase domain and polymerase subdomains are color coded as shown in the figure.

Although the X-ray crystallographic studies have provided valuable structural information on substrate binding and enzyme function, there are still significant unanswered questions about substrate recognition and catalysis. For example, how does the enzyme recognize and bind to the site of DNA damage, how does it select the correct dNTP to bind, and what is the nature of the transition state? Some of these processes, such as DNA damage recognition and dNTP selection and binding are inherently dynamic and cannot be completely captured by structural studies alone. Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique for studying molecular processes that are of a dynamic nature over a wide range of time scales, and can provide valuable information about enzymatic function that are of a dynamic nature and not readily accessible by structural studies alone (9).

Some insight into substrate binding and correct nucleotide sensing has been obtained in NMR studies utilizing ^{1}H-^{13}C HSQC experiments of ^{13}C-methyl methionine labeled pol β (10,11). For example, chemical shift and linewidth changes for the Met18 resonance of the lyase domain were observed upon addition of the DNA template, primer and downstream oligonucleotide, and were attributed to proximity to the linker connecting the two pol β domains whose structure is altered in the presence of substrate DNA(10). Changes in the Met155, Met158, Met282 and Met236 resonances, especially broadening of the Met155 and Met282 resonances, were attributed to conformational activation upon formation of an abortive ternary complex of pol β with 1-nt-gapped DNA and dNTP (10).

More recently, Berlow *et al.* (12) characterized the backbone dynamics of unliganded pol β and a binary complex of pol β with a 1-nt-gapped DNA substrate using transverse relaxation optimized spectroscopy (TROSY) ^{15}N Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiments. They found significant ms backbone motion in the lyase domain and the polymerase D subdomain of the enzyme lacking substrates, but only Glu21 in the lyase domain showed evidence of ms backbone dynamics in the binary complex. They concluded that the observed backbone motion in the apoenzyme might facilitate DNA binding by sampling a conformation related to the binary complex. These studies were extended to include ternary complexes with correct and incorrect dNTPs (13). It was reported that a ternary complex with a non-hydrolyzable matched nucleotide substrate showed no evidence of backbone ms motion. A ternary complex formed with a mismatched non-hydrolyzable nucleotide, however, showed some evidence of ms backbone motion even though the amide chemical shifts were not significantly perturbed from the binary complex. These results lead to the conclusion that the enzyme remains in the open state with the mismatch nucleotide. Together, these studies indicate that the most extensive backbone motion on the ms time scale was present in the apoenzyme. Binding of DNA and subsequent formation of a ternary complex largely dampens these motions; however, there was some enhanced backbone flexibility in the ternary complex with the incorrect nucleotide.

Although these studies provide insight on how slow backbone dynamics contributes to substrate recognition, there are still gaps in understanding how the enzyme discriminates among dNTPs, about the nature of the excited state and how sidechain dynamics contributes to substrate recognition and catalysis. The earlier studies of methionine-labeled pol β provided some information about how sidechain motion might contribute to substrate binding and correct nucleotide selection, but a more thorough quantitative investigation of sidechain dynamics is required to obtain a more complete understanding of the role of dynamics on substrate binding and enzyme function. In this paper, we report on the results of ^{13}C-^{1}H multiple-quantum CPMG relaxation dispersion and methyl TROSY HMQC experiments to study isoleucine and methionine methyl μs-ms sidechain dynamics of the apoenzyme and 1-nt-gapped binary and ternary nucleotide complexes of pol β to obtain a more complete understanding of enzyme function. The results demonstrate significant differences in sidechain dynamics in the apoenzyme and each of the complexes that may help to facilitate substrate binding and promote catalysis.

MATERIALS AND METHODS

Materials

Following a strategy used previously (14), we studied a pol β complex with a double-hairpin that forms a 1-nt-gapped DNA substrate with the following sequence: 5′-^{P}GGCGAAGCCTGGTGCGAAGCACC-3′ (underlined nucleotide is in the gap). The oligonucleotide was from IDTDNA, Coralville, IA, USA. The non-hydrolyzable deoxynucleoside triphosphate 2′-deoxyadenosine-5′-[(α,β)-methyleno]triphosphate (dAPCPP) was obtained from Jena Bioscience.

Protein expression

The U-[^{2}H], Ile-δ1-[$^{13}CH_3$], Met-[$^{13}CH_3$] pol β C267A variant was prepared as described previously (15) by growth at 37°C of the *Escherichia coli* BL21(DE3) transformant in M9 deuterated (99% D_2O) medium containing U-[^{2}H] glycerol and $^{15}NH_4Cl$ as the sole carbon and nitrogen sources to a 600 nm absorbance of 1. Then 50 mg per liter of α-ketobutyric acid, sodium salt [methyl-^{13}C; 3,3′-$^2H_2$] and 100 mg per liter of L-methionine (methyl-^{13}C; 2,3,3,4,4-$^2H_5$) were added to the deuterated culture 20 min prior to protein induction by addition of IPTG—the expression continued at 30°C overnight. The U-[^{2}H], Ile-δ1-[$^{13}CH_3$], Met-[$^{13}CH_3$], Leu-[$^{13}CH_3$], Val-[$^{13}CH_3$], pol β C267A variant was prepared by similar cell growth, with inclusion of 100 mg per liter of α-ketoisovaleric acid, sodium salt (1,2,3,4-$^{13}C_4$, 99%; 3,4′,4′,4′-$^2H_4$, 98%). The expressed proteins were purified as described previously (15). The protein concentration was determined using a 280 nm extinction coefficient of 20088 $M^{-1}cm^{-1}$.

NMR spectroscopy

NMR samples contained 0.1–0.3 mM protein in a D_2O buffer consisting of 5 mM $MgCl_2$, 50 mM Tris-d_{11} (pH 7.6), 150 mM KCl, 1 mM CDTA, 1 mM dithiothreitol, 0.1 mM 4-(2-aminoethyl)-benzene-suflonyl fluoride (AEBSF), 0.04% NaN_3 and 50 μM DSS as an internal chemical shift standard. $MgCl_2$ was not added to the apoenzyme samples. The CPMG experiments were carried out with ∼175 μM pol β and 10% excess of the double hairpin DNA in both the binary and the dT-dAPCPP ternary complexes. In the ternary complex the nonhydrolyzable dAPCPP concentration was in 40% excess. These concentrations are well above the reported K_d of 1-nt-gapped DNA binding (20 nM) (16) and the K_d of 2 μM for dT-dAPCPP binding (17) to ensure the enzyme complexes were saturated with substrate. All NMR assignment experiments on the pol β complexes were carried out at 35°C on a Varian Inova or Agilent DD2 600 and 800 MHz NMR spectrometers equipped with cryogenic triple resonance probes. The assignment experiments on the apoenzyme were carried out at 25°C. The NMR data were processed using NMRPipe (18) and the chemical shift assignments were made using NMRViewJ (19). The methionine methyl chemical shifts were previously assigned (10,14). The Ile methyl chemical shifts of the binary complex were assigned from (HM)CM(CGCBCA)NH, HM(CMCGCBCA)NH, HMCM[CG]CBCA, HMCM(CGCBCA)CO experiments (20) by measuring correlations of the methyl chemical shifts to the backbone amide, Cα, Cβ and C' chemical shifts that were previously assigned (21). Most of the assignments were made from the methyl detected HMCM[CG]CBCA experiment, which had considerably greater sensitivity than the HMCM(CGCBCA)CO and amide detected experiments. The Ile methyl chemical shifts of the matched dAPCPP ternary complex were assigned by comparison of the shifts with the binary complex and by mapping methyl-methyl NOEs obtained from a 3D methyl-methyl NOESY experiment (22) onto a high-resolution crystal structure (2FMS (23)) of the ternary complex, in a similar fashion to that described by Xaio *et al.* for the MAP kinase ERK2 enzyme (24). The HMCM[CG]CBCA experiment was also used to differentiate between Leu and Val cross-peaks in the 3D methyl-methyl NOESY experiment. The methyl chemical shifts of the apoenzyme were made in a similar fashion to that used to assign the ternary complex and were also compared to previously assigned shifts of methyl resonances in the C and N subdomains (25). The HMCM[CG]CBCA experiment was also used to verify that the Ile Cβ and Cα shifts of the apoenzyme and ternary complex were consistent with previously assigned values (12,21). In addition, HMQC spectra of I106M, I257M and I260M mutants were used to verify some of the Ile assignments in the apoenzyme and ternary complex. All the 3D and 4D pulse sequences were obtained from the laboratory of Dr Lewis Kay, University of Toronto. The Ile and Met methyl chemical shifts of the apoenzyme, binary and dAPCPP ternary complexes were deposited with the Biological Magnetic Resonance Data Bank (26) with accession numbers 27407 (apoenzyme), 27409 (binary complex) and 27410 (ternary complex).

Methyl-TROSY ^{13}C-^{1}H multiple quantum CPMG relaxation dispersion

The methyl-TROSY ^{13}C-^{1}H multiple quantum CPMG relaxation dispersion experiments (27) were carried out on U-[^{15}N,^{2}H] Ile δ1-[^{13}C-^{1}H] and Met ε-[^{13}C-^{1}H] labeled pol β in complex with the double hairpin, 1-nt-gapped DNA substrate described above. The CPMG experiments on the binary and ternary complexes were carried out at 35°C on Agilent DD2 600 and 800 MHz NMR spectrometers. The experiments on the apoenzyme were carried out at 25°C due to the reduced stability of the apoenzyme at the higher temperature. The data were acquired as pseudo 3D datasets using a 40 ms constant time delay with CPMG frequencies ranging from 50 Hz to 1000 Hz at 600 MHz and 50 Hz to 950 Hz at 800 MHz. For each dataset, two duplicate points at 100 and 500 Hz were collected for error estimation. A purge element for smaller proteins and residues with high mobility was used to remove the fast relaxing L1 and L3 contributions to the NMR signal (27). At 800/600 MHz, the data were acquired with acquisition times of 91/122 ms (t2) and 40 ms/35 ms (t1), with 32 scans per FID, using a 1.3 s delay between scans, for a total acquisition time of ∼46/31 h per dispersion. Three complexes were assessed for millisecond time scale motions; the apoenzyme, a binary complex with the 1-nt-gapped DNA, and a matched dT-dAPCPP ternary complex with the non-hydrolyzable nucleotide dAPCPP. Individual and cluster fitting of the CPMG data were accomplished using the Relax software version 3.3.0. (28) Both the analytic MMQ CR72 and numeric NS MMQ two-site exchange models were used to fit the data to obtain the exchange rate constants (k_{ex}), the major state populations (p_A), and the ^{1}H and ^{13}C chemical shift differences ($\Delta\omega_H$ and $\Delta\omega_C$) between the major and minor state populations. For two-site exchange, the population of the minor state is:

$$p_B = (1 - p_A)$$

and the overall exchange rate is equal to the sum of the forward and reverse rates for exchange between the two states:

$$k_{ex} = k_{AB} + k_{BA}$$

with the forward and reverse rates given by,

$$k_{AB} = p_B k_{ex} \text{ and } k_{BA} = p_A k_{ex}$$

respectively. Akaike's Information Criterion (29) model selection was used to select the best model and to determine if statistically significant conformation exchange contributions (R_{ex}) to the total transverse relaxation rates ($R_{2,eff}$) could be obtained from the data. In addition, only the results for methyl groups with dispersion curves with $R_{ex} \geq 3$ Hz are reported

Data analysis

The chemical shift titration data for dT—dGPCPP binding to the pol β 1-nt-gapped DNA complex were analyzed using a three-parameter fit with an in house written Mathematica (Wolfram Research, Champaign, IL, USA http://www.wolfram.com/mathematica/) script. The molecular graphics images were generated using PyMOL (https://pymol.org/2/).

RESULTS

Comparison of $^{1}H/^{13}C$ HMQC spectra of the apoenzyme, binary and ternary complexes of pol β

Crystal structures (4,5,30) indicate that pol β undergoes large conformational changes transitioning from an elongated apoenzyme form to an open binary DNA complex (Figure 1). Figure 1 also shows all the Ile and Met methyl groups that were assigned in this study. A total of 29 methyl groups distributed throughout the protein were assigned, with 8 methyl groups residing in the lyase domain, 4 in the D subdomain, 11 in the C subdomain, and 6 in the N subdomain. Binding of dNTP to form the ternary complex involves repositioning of the N subdomain (Figure 1). Many of the isoleucine and methionine methyl resonances exhibit changes in their HMQC spectra as the result of the apoenzyme to binary to ternary conformational transitions (Figure 2A and B, Figure 3 and Supplementary Table S1A). The largest chemical shift perturbations (CSPs) observed upon DNA binding are located primarily in the lyase domain and the D subdomain (Figure 3A and B). The largest CSPs observed upon nucleotide binding are located primarily in the C and N subdomains (Figure 3E and F). In addition to the residues in the N subdomain, Ile257 and Ile260 in the C subdomain are close to the N subdomain and also appear to be affected by the open to closed transition, while Met155 is closer to the dNTP binding site and is more directly affected by nucleotide binding. Several of the isoleucine δ1-methyl resonances in the binary complex show evidence of conformational exchange line broadening that is not observed in the apoenzyme and disappears when the ternary complex is formed (Figure 2A and B; Supplementary Table S1B). For example, Ile277 that is located in α-helix N of the N subdomain (Figure 1) comes into contact with the lyase domain as a result of the conformational change from binary to ternary complex (Figure 4A). The Ile277 δ1-methyl resonance shows broadening in the binary complex (Figure 2B and Supplementary Table S1B). We attribute this primarily to changes in rotameric distribution, as well as other variations of the local magnetic environment that are associated with the open to closed conformational change. This interpretation is supported by the sharpening of the Ile277 resonance upon formation of the matched dAPCPP ternary complex, so that a single conformation predominates, reducing the exchange broadening (Figure 2B and Supplementary Table S1B). Changes in methyl chemical shifts associated with the open (binary) to closed (ternary) conformational transition were also seen with Ile88, Met155, Ile174, Ile257, Ile260, Met282, Ile293 and Ile323. Several of these residues (e.g., Ile260, Met282, Ile293 and Ile323) are also located near or in the N subdomain (residues 262–335).

In contrast to this picture with the closed, matched (i.e. with a complementary incoming nucleotide) ternary complex, the HMQC spectrum of a 1–1 complex of the enzyme with a non-complementary incoming nucleotide, dGPCPP, was virtually identical to the spectrum of the binary complex (Figure 2C, blue and black spectra). This indicates that the open conformation predominates in the presence of the mismatched nucleotide at least at low concentrations. When the dGPCPP concentration was increased to four times the enzyme concentration, the Ile88 and Ile277 resonances continued to broaden and are slightly shifted (Figure 2C, red spectrum).

In order to further evaluate the response of the binary complex to the mismatched nucleotide, a second titration was carried out in which the dGPCPP concentration was increased to ∼20 times the binary complex concentration. Based on the magnitude of the observed shifts and the resolution of the peaks, the response was most readily observed for Ile88 and Ile277 (Supplementary Figure S1A). The mean K_d value for the mismatch dT-dGPCPP binding was 1.2 ± 0.4 mM (Supplementary Figure S1B). This is somewhat larger than the K_d of 211 μM reported recently by Moscato *et al.* (13) for dG-dAPNPP binding and much larger than the previously reported K_d of 2 μM for the matched nucleotide dT-dAPCPP binding (17). Some of the variations in the K_d values for mismatched nucleotides probably arise due to variations in the DNA sequences used. As noted above, Ile88 reports on structural and dynamic changes in the segment linking the two pol β domains, while Ile277 reports on formation of the closed, catalytically activated complex. Since neither of these resonances is positioned near the dNTP binding site, both effects most probably correspond to dynamic equilibria of the enzyme. In both cases, the resonances shifted toward the position of the ternary complex containing the matched dAPCPP; however, analysis of the shift versus concentration curves yielded asymptotic limits that are well short of the shifts observed for the matched nucleotide (Supplementary Figure S1B). We interpret these results to indicate that binding of the mismatched nucleotide does not alter the pol β conformational equilibrium to the extent that the matched nucleotide does. Thus, as the dNTP binding site becomes largely occupied, the conformational response is reduced. Based on the limiting shifts of the Ile277 and Ile88 resonances compared to the shifts characterizing the complex with the matched dNTP, the complex is only closed for ∼ 33% of the time. The dynamic behavior of the mismatched complex is also consistent with the significant broadening of the Ile88 and Ile277 resonances even at the highest dGPCPP concentrations used. This reduces the effective time

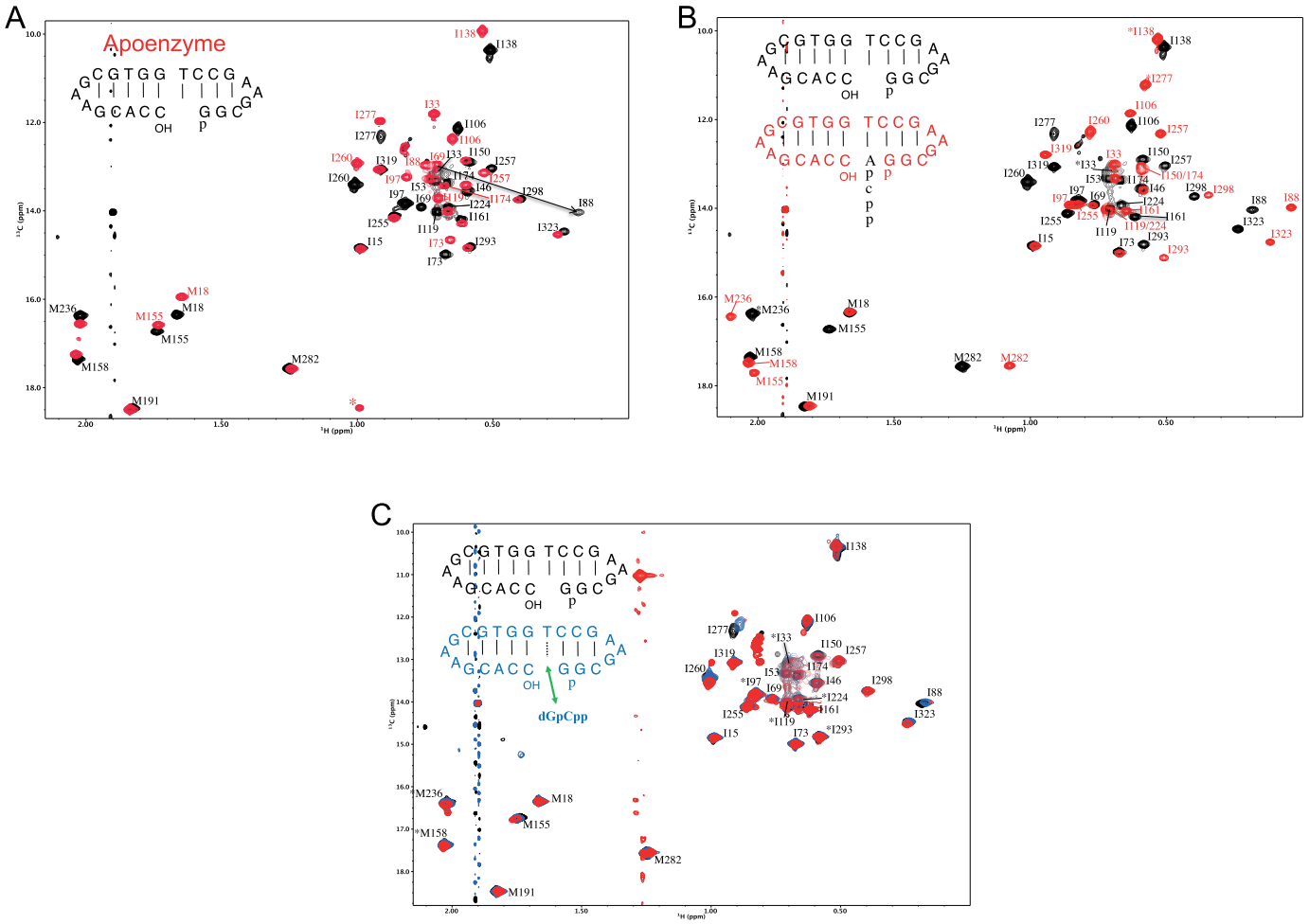

Figure 2. ^{1}H-^{13}C HMQC spectra of the Ile and Met resonances in the pol β complexes. (**A**) Overlay of the ^{1}H/^{13}C HMQC spectra of the apoenzyme and 1-nt-gapped binary DNA complex of pol β. The binary complex spectrum is shown in black while the apoenzyme spectrum is shown in red. The binary spectrum was obtained at 35°C at pH 7.6 in Tris buffer. The apoenzyme spectrum was obtained in the same buffer at 25°C. The peak at ∼1 ppm (^{1}H)/18.5 ppm (^{13}C) in the apoenzyme spectrum arises from a Val methyl group and is marked by an asterisk. This peak appears because the spectrum of the apoenzyme was obtained using an ILVM methyl-labeled sample, whereas the spectra of the binary and ternary complexes were all obtained using IM methyl-labeled samples. (**B**) Overlay of the ^{1}H/^{13}C HMQC spectra of the 1-nt-gapped binary DNA complex of pol β and the ternary complex with the non-hydrolyzable dAPCPP nucleotide. The binary complex spectrum is shown in black while the ternary complex spectrum is shown in red. The spectra were obtained at 35°C at pH 7.6 in Tris buffer. (**C**) Overlay of the ^{1}H/^{13}C HMQC spectra of the 1-nt-gapped binary DNA complex of pol β and the ternary complex with the non-hydrolyzable dGPCPP mismatch dNTP. The binary complex spectrum is shown in black, the spectrum of the ternary complex containing an ∼1–1 ratio (∼250 μM dGPCPP - ∼165 μM pol β) of the mismatched nucleotide is shown in blue and the ternary spectrum containing a 4-fold excess (∼833 μM dGPCPP - ∼200 μM pol β) of the mismatched nucleotide is shown in red. The spectra were obtained at 35°C at pH 7.6 in Tris buffer.

period for the reaction to occur and, along with structural differences, is probably an important limitation on the rate of mismatch incorporation.

The above interpretations are supported by the increasing availability of high-resolution structures for various ternary complexes. Crystallographic analyses show that the Ile277 sidechain adopts the predominant trans-conformation in most pol β binary structures, e.g. 3ISB (30) (Figure 4A). However, in many structures of the closed ternary complex, i.e. 3ISD (30), 3C2M, 3C2K and 3C2L (8), the interaction between Ile277 and residues in the lyase domain stabilizes a *gauche*- (*g*-) conformation (Figure 4A). This change in rotameric preference appears to result from unfavorable contacts between the Ile277 δ1-methyl group with the Ser44 sidechain and the Arg40 backbone carbonyl oxygen (Figure 4A). The large up-field ^{13}C shift of the Ile277 methyl group in the ternary complex (Figures 2B and 3C and E) also indicates that the sidechain distribution is weighted more heavily to the *g*- rotamer in solution (31). Based on the relation between the ^{13}C shift and the sidechain conformation developed by Hansen *et al.* (32), the fraction of Ile277 in the *g*- conformation increases from ∼47% in the binary complex to ∼65% in the ternary complex.

The methyl resonances of Ile33 and Ile88 are also broadened in the HMQC spectra of the binary and mismatched ternary complexes but not in the apoenzyme or the matched ternary complex (Figure 2 and Supplementary Table S1B). In the crystal structures of the binary (3ISB) and ternary (2FMS, 3ISD) complexes, Ile33 is held in a (*gauche*+) *g*+ conformation due to interactions with the sidechain aro-

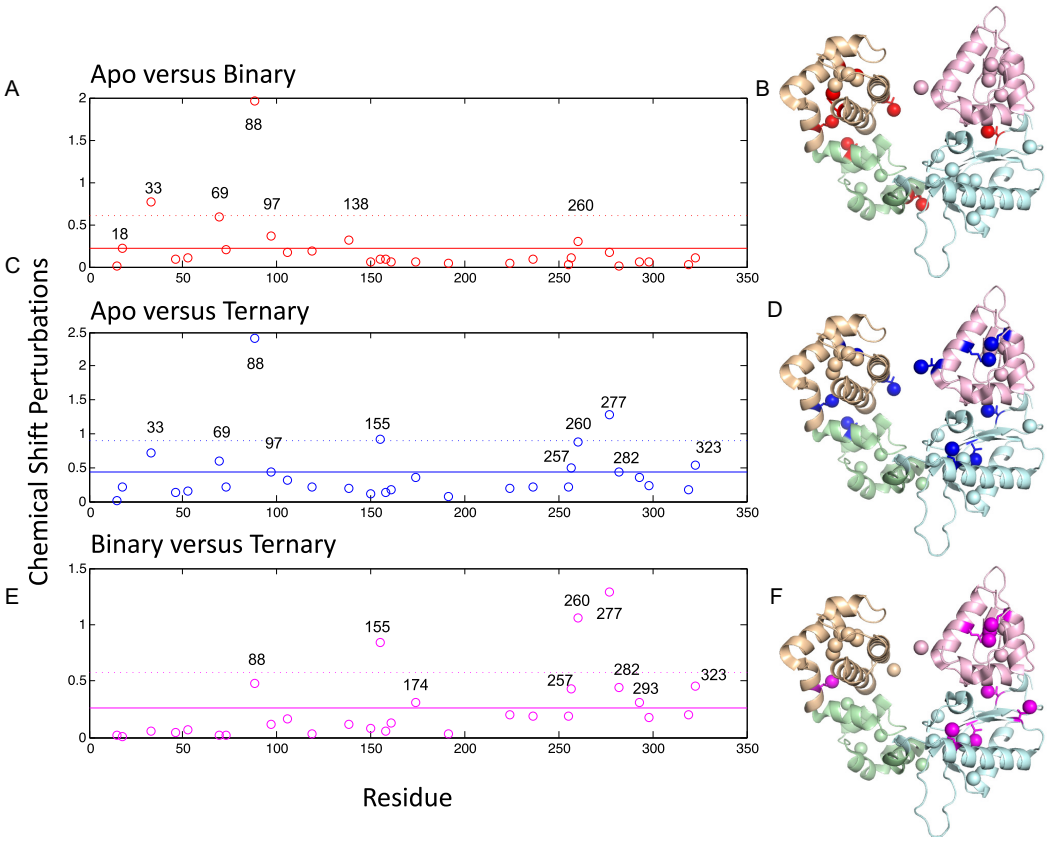

Figure 3. Summary of NMR CSPs of pol β, obtained from the HMQC spectra, for the apoenyzme versus binary DNA complex (**A**), apoenzyme versus matched (dAPCPP) ternary complex (**C**) and the binary versus the matched (dAPCPP) ternary complex (**E**). The horizontal solid line is the mean of the CSPs and the dotted line is the mean plus one standard deviation. The methyl groups with larger CSPs above the solid lines are mapped onto the 2FMS ternary complex structure in (**B**), (**D**) and (**F**). The CSPs were calculated according to the equation: $\sqrt{\left(\frac{\Delta\delta H}{\alpha}\right)^2+\left(\frac{\Delta\delta C}{\beta}\right)^2}$ (40), where $\Delta\delta_X$ is the chemical shift change of the nucleus X (H or C) between the two states, $\alpha = 0.29$ (Ile Hδ1), 0.40 (Met Hε) and $\beta = 1.66$ (Ile Cδ1), 1.82 (Met Cε) are the standard deviations of the ^{1}H and ^{13}C chemical shifts from the Biological Magnetic Resonance Data Bank for the specified atom.

matic rings of His34 and Tyr36 (Figure 4B). The His34 sidechain is constrained by a stacking interaction with a base of the DNA template strand that forces the Ile33 δ1 methyl group to rotate toward the Tyr36 aromatic ring. The His34 sidechain is not constrained by the base stacking interaction in the apoenzyme, allowing the Ile33 sidechain to adopt alternate conformations that can point away from the aromatic ring as seen in the two different molecules in the unit cell of the 3UXN (5) crystal structure (Figure 4B). This results in a significant shift change and narrowing of the apoenzyme resonance (Figures 2A and 3A and C) consistent with fast exchange averaging between conformations. Based on the observed ^{13}C chemical shift of 11.8 ppm, the fraction of the Ile33 sidechain in the *g*- conformation in the apoenzyme is ∼54%, although only the *g*+ and *trans* conformations are observed in the two molecules of the unit cell of the 3UXN crystal structure.

Although the conformations and environments of the Ile33 sidechain in the two complexes are very similar in crystallographic structures (Figure 4B) and demonstrated by the similarity of their δ1 methyl chemical shifts (Figure 2B, Figure 3E and Supplementary Table S1A), the methyl resonance is not exchange broadened in the ternary complex. This and the absence of significant relaxation dispersion (see below) indicates that the sidechain motion of Ile33 is reduced in the ternary complex and that a single conformation predominates. In contrast, although the conformation and chemical environment of the Ile33 sidechain in the binary complex is very similar to that in the ternary complex, the sidechain in the binary complex is undergoing significant μs-ms conformational exchange measured in the relaxation dispersion experiments that also results in exchange broadening in the HMQC spectrum. Since the predominate configuration is tightly constrained to *g*+, the μs-ms dynamics must result from slight variations in rotamer angle that disappear once the closed conformation is formed.

The broadening of the Ile88 methyl resonance in the spectra of the binary and mismatch ternary complexes likely arises from an exchange process that affects its proximity to the Phe25 aromatic ring (Figure 4C). The Ile88 methyl resonance is not exchange broadened in the spectrum of the matched ternary complex, indicating that the sidechain motion of Ile88 is also reduced in the ternary complex when the closed conformation is formed. This is not surprising since the lyase domain makes contact with N subdomain upon formation of the closed complex, reducing the dynam-

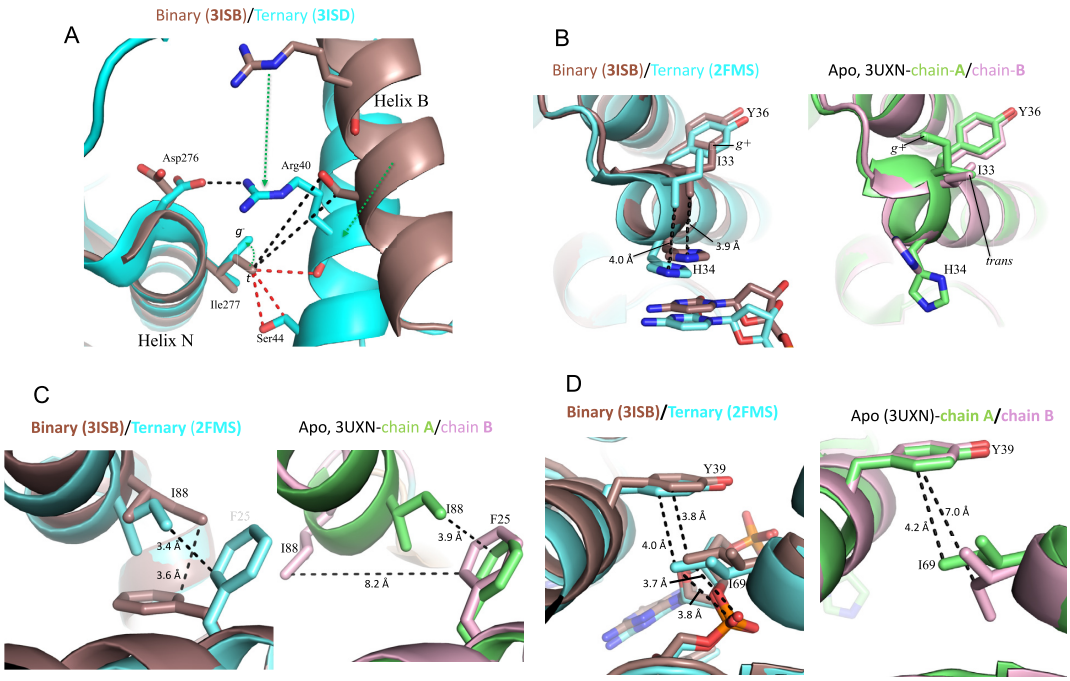

Figure 4. Comparison of binary and ternary structure sidechains related to linewidth and chemical shift changes in the $^{1}H/^{13}C$ HMQC spectra. (**A**) Correlation of Ile277 rotamer and N subdomain conformation. Overlay of pol β region containing helix N in a binary complex (3ISB, brown) and ternary complex (3ISD (30), cyan). Distances between the Ile277 Cδ in the binary complex and nuclei on helix B are indicated with dashed black lines. The steric clash that would result in the ternary complex if the Ile277 sidehchain conformation remained *trans* in the ternary complex is indicated by distances shown as dashed red lines. These unfavorable interactions tend to favor adoption of the *gauche*- conformation by Ile277 in the ternary complex. (**B**) Comparison of the local environment of the δ1-methyl group of Ile33 in the X-ray structures pol β. The overlay of the binary (PDB ID: 3ISB) and ternary complex (PDB ID: 2FMS) shows that the Ile33 sidechain is held in a similar local environment in both structures, due to its interaction with the rings of His34 and Tyr36. The position of the His34 ring is in turn constrained by a stacking interaction with a base on the downstream primer strand of the DNA. The Ile33 sidechains are also in similar chemical environments in solution as demonstrated by the similarity of their δ1-methyl resonances. However, the sidechain is broadened and dynamic in the binary but not the ternary complex, probably due to variability in the position of His34. A structural overlay of the A and B-molecules of the unit cell of the 3UXN X-ray crystal showing that the Ile33 sidechain is not constrained by the His34 ring in the apoenzyme and can sample significantly different chemical environments, resulting in an up-field ^{13}C shift of the δ1-methyl resonance in the HMQC spectrum. (**C**) Orientation of the of Ile88 sidechain with respect to the aromatic ring of Phe25. A structural overlay of the binary complex (3ISB) with the ternary complex (2FMS) shows the close proximity of the sidechain that results in up-field ^{1}H shifts of the δ1-methyl resonances in the HMQC spectra of the both complexes, and line broadening in the spectrum of the binary complex. An overlay of the A- and B-molecules of the unit cell of the 3UXN X-ray crystal showing that the Ile88 sidechain is free to sample significantly different local environments and is not shifted by the ring current effects of the Phe25 aromatic ring. (**D**) Orientation of the of Ile69 sidechain with respect to the aromatic ring of Tyr39. In the overlay of the binary and ternary complexes, the Ile69 δ1-methyl group is sandwiched between the aromatic ring of Tyr39 and the DNA backbone of the downstream DNA strand, resulting in a broadened resonance in the HMQC spectrum of both complexes. In the overlay of the two molecules in the unit cell of the apoenzyme crystal structure, the sidechain can adopt different orientations and is not broadened by its interaction with Tyr39.

ics when the ternary complex is formed. This observation is further supported by the relaxation dispersion results described below. In the spectrum of the apoenzyme, the Ile88 resonance is not broadened or shifted by the ring current effects of Phe25 (Figure 2A), implying that the sidechain is either not interacting with the aromatic ring or that the ring current effects are removed due to rapid conformational averaging. In the 3UXN X-ray crystal structure of the apoenzyme, the Ile88 sidechain of the A-molecule in the unit cell is close to the aromatic ring, while in the B-molecule it is far from the aromatic ring (Figure 4C). This implies that the flexibility in the linker helix (helix E) between the lyase and polymerase domains that gives rise to these different conformations in the crystal structure is also responsible for averaging out the ring current effects on the methyl resonance in solution.

The Ile69 resonance exhibits diminished peak intensity and line broadening in both the binary and ternary HMQC spectra, but not in the apoenzyme (Figure 2A and B; Supplementary Table S1B). In the crystallographic structures of the binary and ternary complexes (Figure 4D), the Ile69 methyl group is constrained to be in close proximity to the aromatic ring of Tyr39 due to its interaction with the phosphate backbone between the first and second nucleotides of the downstream DNA strand. Small fluctuations in the position of the methyl group in the vicinity of the aromatic ring cause the resonance to broaden in the HMQC spectrum. In the apoenzyme the Ile69 is free to adopt multiple conformations (Figure 4D) that remove the broadening effect of the aromatic ring.

Methyl-TROSY ^{13}C-^{1}H multiple quantum CPMG relaxation dispersion

The methyl-TROSY ^{13}C-^{1}H multiple quantum CPMG relaxation dispersion experiments were carried out on U-[^{15}N,^{2}H] Ile δ1-[^{13}C-^{1}H] and Met ε-[^{13}C-^{1}H] labeled pol

β (apoenzyme) and in complex with the 1-nt-gapped DNA substrate shown in Figure 2. Additionally, a ternary complex with a 1-nt-gapped DNA substrate and a non-hydrolyzable nucleotide dAPCPP was examined. The results of these experiments are summarized in Tables 1–3. The 800 MHz dispersion curves of the methyl groups showing the largest exchange contributions (R_{ex}) to the transverse relaxation rates ($R_{2,eff}$) in the apoenzyme (Ile106 and Ile138), the binary complex (Ile33) and the ternary complex (Ile260) are shown in Figure 5A. Dispersion curves of the same methyl groups from the other complexes are also shown for comparison. All the methyl groups exhibiting μs-ms dynamics in the apoenzyme and the two complexes are shown in Figure 5B. The methyl groups are color coded by their individual conformational exchange rates (k_{ex}). The dispersion curves for all the methyl groups in the apoenzyme and the two complexes are shown in Supplementary Figure S2.

Apoenzyme

The methyl TROSY ^{13}C-^{1}H multiple quantum CPMG relaxation dispersion results for the apoenzyme are summarized in Table 1, and the dispersion curves are shown in Supplementary Figure S2A. Only three methyl groups (I106, I119 and I138) of the apoenzyme, out of a total of 29 that were assigned (Figure 1), exhibit significant dispersions (R_{ex} > 3.0/s) indicative of μs-ms sidechain dynamics (Figure 5B). These residues are all located in the D subdomain. The Ile106 and Ile138 methyl resonances in the apoenzyme had the largest of all the R_{ex} values measured in the apoenzyme and in the two complexes (Tables 1–3). The relaxation dispersion data for the three residues are similar and could be fit to a single two-site exchange process, without a large increase in chi-square, with $k_{ex} = 1919 \pm 23\ s^{-1}$ (Table 1) and a ground state population $p_A = 93.9 \pm 1.4\%$. For two-site exchange, $p_B = 1 - p_A$ and $k_{AB} = p_B k_{ex}$, which gives a forward rate of $k_{AB} \sim 117.1 \pm 26.9\ s^{-1}$. The values of the chemical shift differences $\Delta\omega_C$ obtained from the global fit of the CPMG data (Table 1) do not correspond to the shift differences $\Delta\delta_C$ observed between the HMQC spectra of the apoenzyme and the binary or ternary complexes, indicating that the sidechains are not sampling conformations corresponding to the binary or ternary complex, but may be sampling local conformations that allow increased flexibility for DNA binding. These results suggest that the μs-ms sidechain motion observed in the D subdomain may help to facilitate DNA binding of the polymerase domain by allowing the domain to sample different DNA geometries before binding tightly to the gapped DNA.

Binary complex

Ten methyl groups of the binary complex exhibit significant dispersions (R_{ex} > 3.0/s) indicative of μs-ms sidechain dynamics. However, one of these methyl groups, the δ1 methyl group of Ile224, exhibits a small R_{ex} value and yields an unsatisfactory fit to the data with a k_{ex} value of 31 ± 151 s^{-1}. Individual fits of the CPMG dispersion data are given in Table 2, and the dispersion curves are shown in Supplementary Figure S2B. To rule out the possibility that the observed dispersions were due to self-association of the binary enzyme complex, the CMPG experiments were repeated at 800 MHz using another sample at a similar enzyme complex concentration (∼175 μM pol β/193 μM DNA) and a sample at approximately half that concentration (∼87.5 μM pol β/96.6 μM DNA). The dispersion curves obtained from the two complexes at the different concentrations are very similar to each other (Supplementary Figure S3A) and to the curves obtained with the previous sample at 600 and 800 MHz (Supplementary Figure S2B), indicating that the observed μs-ms dynamics is not due to dynamic self-association of the binary complex. The dynamic methyl groups are scattered throughout the enzyme (Figure 5B). The δ1 methyl group of Ile33 in the lyase domain of the binary complex exhibits a large R_{ex} value (12.7/s at 800 MHz) with $k_{ex} = 1907 \pm 150\ s^{-1}$. This μs-ms motion also gives rises to line broadening of the methyl resonance in the HMQC spectrum, due to its interaction with the aromatic ring of Tyr36 as described above, which is not observed in the apoenzyme or ternary complex. The methyl group in the ternary complex is in a very similar conformation and environment, but its methyl resonance is not exchange-broadened, and it exhibits no relaxation dispersion in the CPMG experiments, once the lyase and N subdomain close to form the ternary complex. In the apoenzyme, the sidechain is rapidly averaging between multiple conformations and does not exhibit line broadening or μs-ms dynamics. The isoleucine δ1-methyl groups of residues Ile97 and Ile106 of the D subdomain also have large R_{ex} contributions (≥10/s at 800 MHz) to the transverse relaxation rates, with individual k_{ex} values of $834 \pm 108\ s^{-1}$ and $2369 \pm 97\ s^{-1}$ for Ile97 and Ile106, respectively. The remaining residues exhibiting methyl group dispersion in the binary complex have lower dispersions (R_{ex} < 7/s), and individual k_{ex} values that range from a low of $841 \pm 67\ s^{-1}$ for Ile97, excluding the anomalously low value for Ile224, to a high of $3595 \pm 944\ s^{-1}$ for Ile138. The individual k_{ex} values are mapped onto the structure of the binary complex (3ISB) in Figure 5B. Due to the wide range of k_{ex} and p_A values, a global fit of all the dispersion data was not justified. However, Ile293 and Ile319 in the N subdomain have very similar individual k_{ex} values (Table 2). A cluster fit of the CPMG data for these two residues yields a k_{ex} value of $1901 \pm 405\ s^{-1}$ with $|\Delta\omega_C|$ values of ∼0.3 and ∼0.2 ppm for the Ile293 and Ile319 methyl resonances, respectively. The chemical shift differences for these two residues are similar to the equilibrium chemical shift differences $\Delta\delta_C$ (∼0.3 ppm, Supplementary Table S1A) measured from the HMQC spectra between the binary and ternary complexes. These results, coupled with the observed exchange broadening of Ile277, suggest that the N subdomain of the binary complex may be exchanging with some of the closed form of the enzyme in solution.

Ternary complex

The methyl TROSY ^{13}C-^{1}H multiple quantum CPMG relaxation dispersion results for the matched ternary complex with dT-dAPCPP are summarized in Table 3, and the dispersion curves are shown in Supplementary Figure S2C. Five methyl groups of the ternary complex exhibit significant dispersions ($R_{ex} \geq 3.0$/s) indicative of μs-ms

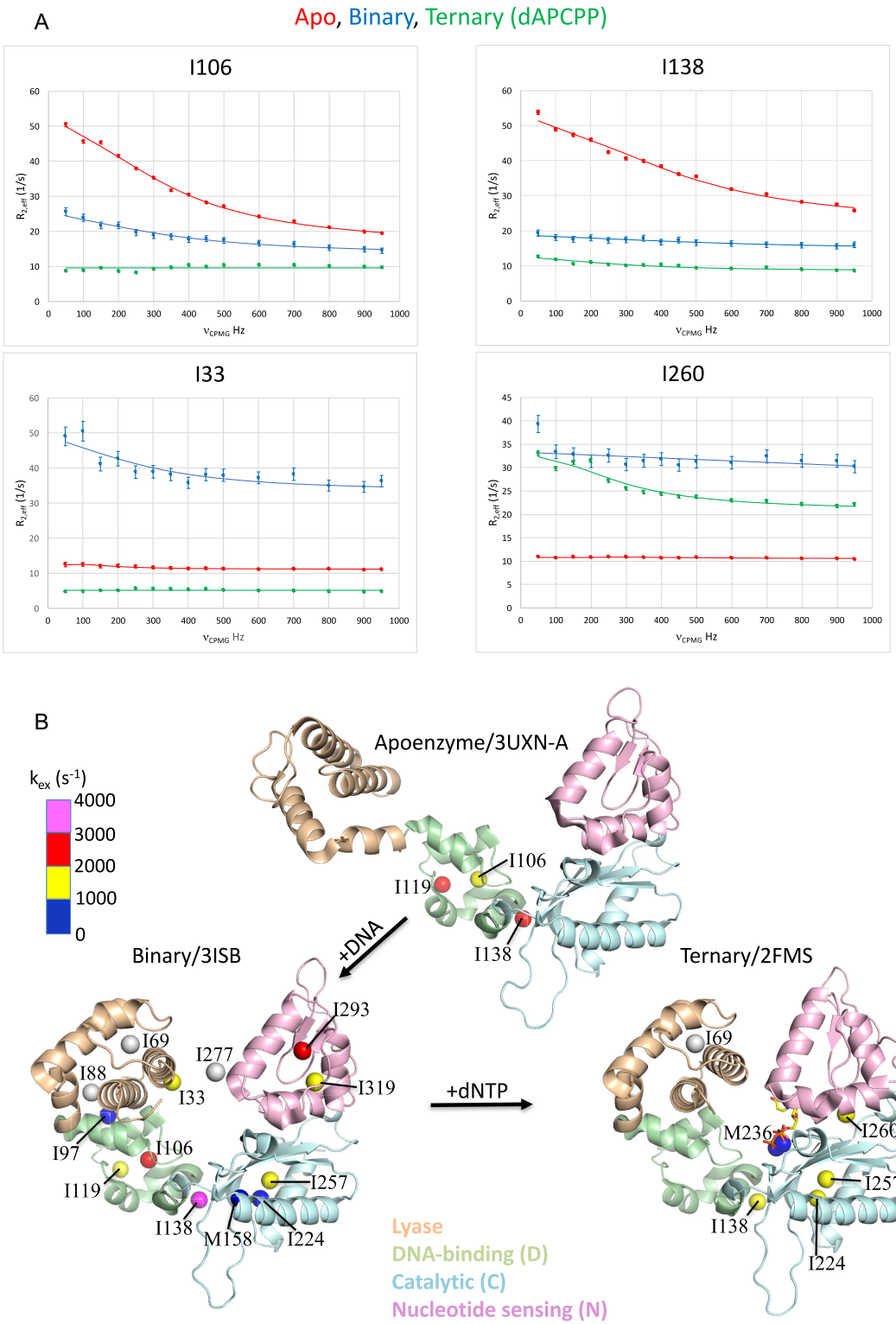

Figure 5. Summary of Ile and Met methyl groups exhibiting µs-ms dynamics in the ^{13}C-^{1}H multiple-quantum CPMG relaxation dispersion experiments. (**A**) Comparison of methyl dispersion curves obtained at 800 MHz for Ile106, Ile138, Ile33 and Ile260. The curves shown in red, blue and green correspond to the curves obtained for the apoenzyme, binary and ternary (dAPCPP) complexes, respectively. Ile106 and Ile138 in the apoenzyme exhibited the largest dispersions of all the methyl groups in the apoenzyme as well as in the binary and ternary complexes. Ile33 had the largest dispersion of all the methyl groups in the binary complex, and Ile206 exhibited the largest dispersion of all the methyl groups in the ternary complex. Ile260 in the binary complex exhibits an elevated relaxation rate, but the exchange contribution to the relaxation rate (R_{ex}) is <3 Hz. (**B**) The methyl groups in the apoenzyme (3UXN-A), the binary complex (3ISB) and the ternary complex (2FMS) showing µs-ms dynamics are color coded by the magnitude of their exchange (k_{ex}) rates obtained from individual fits of the CPMG dispersion curves. The methyl groups that showed significant exchange line broadening are shown as light-gray spheres.

Table 1. Apo pol β methyl TROSY CPMG

Residue	k_{ex}(/s)	p_A(%)	$\|\Delta\omega_H$(ppm)$\|$	$\|\Delta\omega_C$(ppm)$\|$	R_{ex}(/s)*	χ^2/DF**	$\Delta\omega_C$(ppm) global fit	Domain/ Subdomain
I106	1876 ± 14	87.7 ± 0.4	0.1 ± 0.0	0.8 ± 0.0	19.8/30.4	2.9	1.0 ± 0.1	D
I119	2258 ± 126	94.4 ± 0.2	0.1 ± 0.0	0.5 ± 0.0	3.7/5.4	0.5	0.5 ± 0.1	D
I138	2040 ± 26	94.0 ± 0.6	0.1 ± 0.0	1.2 ± 0.0	18.9/24.3	6.0	1.2 ± 0.1	D
Global fit***	1919 ± 23	93.9 ± 1.4				2.6		

* R_{ex} at 600/800 MHz = $R_{2,eff}$(50 Hz) – $R_{2,eff}$(1000 Hz)/$R_{2,eff}$(950 Hz).
**χ^2/DF is the reduce χ2 value for the fit of the CPMG data. The degrees of freedom DF = # data points - # of fit parameters - 1. In this study, DF is 30 - 5 - 1 = 24 for the individual fits, and 3 × 30 - 3 × 3 - 2 - 1 = 78 for the global fit.
***All global fit $\Delta\omega_H$(ppm) values were ≤0.1.

Table 2. Pol β binary complex methyl TROSY CPMG results

Residue	k_{ex}(/s)	p_A(%)	$\|\Delta\omega_H$(ppm)$\|$	$\|\Delta\omega_C$(ppm)$\|$	R_{ex}(/s)*	χ^2/DF**	$\|\Delta\omega_C\|$(ppm) cluster fit	Domain/ subdomain
I33	1907 ± 150	68.7 ± 1.3	0.0 ± 0.0	0.3 ± 0.0	9.2/12.7	1.9		Lyase
I97	834 ± 108	50.5 ± 2.8	0.0 ± 0.0	0.1 ± 0.0	3.8/10.0	0.7		D
I106	2369 ± 97	60.7 ± 2.1	0.1 ± 0.0	0.3 ± 0.0	6.0/11.2	1.2		D
I119	1060 ± 202	86.5 ± 2.5	0.0 ± 0.0	0.2 ± 0.0	2.8/6.1	0.6		D
I138	3595 ± 944	95.4 ± 0.0	0.1 ± 0.0	0.5 ± 0.1	2.0/4.2	0.3		D
M158	841 ± 67	98.0 ± 0.6	0.0 ± 0.0	0.6 ± 0.1	3.9/6.4	0.5		C
I224***	31 ± 151	94.9 ± 6.1	0.0 ± 0.0	1.0 ± 0.4	2.2/3.5	0.7		C
I257	1396 ± 235	54.7 ± 2.5	0.1 ± 0.0	0.1 ± 0.2	2.8/5.4	1.0		C
I293	2093 ± 151	87.1 ± 3.3	0.1 ± 0.0	0.2 ± 0.1	2.3/3.0	0.3	0.3 ± 0.0	N
I319	1889 ± 315	79.8 ± 0.9	0.0 ± 0.0	0.2 ± 0.0	2.0/4.0	0.1	0.2 ± 0.0	N
Cluster 293–319	1901 ± 450	89.0 ± 0.5				0.4		

* R_{ex} at 600/800 MHz = $R_{2,eff}$(50 Hz) – $R_{2,eff}$(1000 Hz)/$R_{2,eff}$(950 Hz).
**χ^2/DF is the reduce χ2 value for the fit of the CPMG data. The degrees of freedom DF = # data points - # of fit parameters - 1. In this study, DF is 30 - 5 - 1 = 24 for the individual fits, and 2 × 30 - 2 × 3 - 2 - 1 = 51 for the cluster fit.
*** The error in k_{ex} is larger than the value.

sidechain dynamics (Figure 5B). To rule out the possibility that the observed dispersions were due to exchange of the nucleotide, the CMPG experiments were repeated at 800 MHz with another sample with a similar ∼1:1 concentration ratio of nucleotide to enzyme–DNA complex (∼245 μM dAPCPP:175 μM pol β-193 μM DNA) and an ∼4:1 concentration ratio of nucleotide to enzyme complex (∼692 μM dAPCPP:173 μM pol β-191 μM DNA). The dispersion curves obtained from the two ternary complexes at the different concentrations are very similar to each other (Supplementary Figure S3B) and to the curves obtained with the previous sample at 600 and 800 MHz (Supplementary Figure S2C), indicating the that the observed μs-ms dynamics is not due to the exchange of the nucleotide, but results from intrinsic dynamic behavior of the pol β–DNA–nucleotide complex. Most of these residues are located in the C subdomain while Ile138 is positioned at the interface between the D and C subdomains. These could be fit to a single two-site exchange process, without a large increase in chi-square, with k_{ex} = 1239 ± 165 s^{-1} (Table 3), and a ground state population p_A = 98.2 ± 0.2%. For two-site exchange, p_B = 1 – p_A and k_{AB} = $p_B k_{ex}$, which gives a forward rate of k_{AB} ∼ 22.3 ± 3.9 s^{-1}. This rate is similar to reported rates of k_{pol} for dT-A insertion (33,34). The values of the chemical shift differences $\Delta\omega_C$ obtained from the global fit of the CPMG data (Table 3) do not correspond to the differences in chemical shifts $\Delta\delta_C$ observed between the HMQC spectra of the ternary complex and the binary complex or apoenzyme. Overall, the most striking difference between the binary and ternary complex is that the μs-ms sidechain dynamics is distributed throughout the enzyme and is characterized by multiple independent exchange processes in the binary complex (Figure 5B and Table 2), but becomes localized to the C subdomain in the ternary complex and can be fit to a single exchange process (Figure 5B and Table 3). This conclusion is further supported by the observation that while several of the methyl resonances throughout the enzyme, including those of Ile33, Ile69, Ile88 and Ile277, show significant line broadening in the spectrum of the binary complex there is much less line broadening in the spectrum of the ternary complex (Figure 2B). Ile69 exhibits diminished peak intensity and line broadening in both the binary and ternary HMQC spectra due to its interaction with Try39 and the downstream primer DNA strand as discussed above. In comparison, the apoenzyme also exhibits μs-ms dynamics characterized by a single exchange process that is localized in the D subdomain (Figure 5B and Table 1).

Changes in rotamer preferences rather than rotamer specificity

A general feature of the chemical shift and CPMG data presented here is that there are few examples of isoleucine or methionine environments that exhibit sufficient bias to select pure rotameric forms. Rotamer populations can be related to sidechain chemical shift data (31,32,35). For studies of the methionine-labeled enzyme, the Met χ3 values were estimated from the chemical shift data summarized in Supplementary Table S1A using Equation 20 in Butterfoss *et al.* (35). Upon formation of the pol β-gapped DNA bi-

Table 3. Pol β dT-dAPCPP ternary complex methyl TROSY CPMG

Residue	k_{ex}(/s)	p_A(%)	$\|\Delta\omega_H$(ppm)$\|$	$\|\Delta\omega_C$(ppm)$\|$	R_{ex}(/s)*	χ^2/DF**	$\Delta\omega_C$(ppm) global fit	Domain/ subdomain
I138	1616 ± 224	95.3 ± 1.1	0.1 ± 0.0	0.4 ± 0.2	2.9/4.0	1.7	0.6 ± 0.1	D
M158***	638 ± 234	98.1 ± 0.3	0.1 ± 0.0	0.4 ± 0.2	2.4/2.7	1.5		C
I224	1742 ± 74	98.7 ± 0.0	0.1 ± 0.1	1.2 ± 0.4	4.5/6.3	2.0	1.0 ± 0.1	C
M236	603 ± 190	98.6 ± 1.7	0.1 ± 0.0	0.9 ± 0.2	6.3/9.3	5.7	1.0 ± 0.1	C
I257	1959 ± 2341	99.5 ± 0.7	0.3 ± 0.0	2.4 ± 0.5	2.3/3.0	1.2	0.6 ± 0.1	C
I260	1226 ± 200	98.1 ± 0.1	0.1 ± 0.0	0.6 ± 0.2	9.1/10.9	3.6	1.1 ± 0.1	C
Global fit****	1239 ± 165	98.2 ± 0.2				1.2		

* R_{ex} at 600/800 MHz = $R_{2,eff}$(50 Hz) – $R_{2,eff}$(1000 Hz)/$R_{2,eff}$(950 Hz).
**χ^2/DF is the reduce χ2 value for the fit of the CPMG data. The degrees of freedom DF = # data points - # of fit parameters - 1. In this study, DF is 32 - 5 - 1 = 24 for the individual fits, and 5 × 30 - 5 × 3 - 2 - 1 = 132 for the global fit.
*** Not used in global fit because R_{ex} < 3 (/s).
**** All global fit $\Delta\omega_H$(ppm) values were ≤ 0.1.

nary complex, the χ3 trans rotamer fraction p_t of Met18 increases by ∼0.1, while upon formation of the abortive ternary complex, the Met155 p_t value increases by ∼0.2. Met282 and Met191 exhibit less ligand dependence, and are predicted to exhibit the highest χ3 *trans* fractions, with p_t = 0.67 for Met282 and 0.85 for Met191. Examination of pol β crystal structures indicates that the χ3 *trans* conformation is frequently observed for Met282, while the *trans* conformation for Met191 is observed only in a few examples that correspond to the highest resolution structures (PDB ID: 4KLI (7); PDB ID: 2FMP (23)). Interestingly, despite the rarity of pol β crystal structures containing a *trans* Met191 conformation, the NOE data for apo pol β show strong interactions between Met191 Cε and Val162 Cγ2, Val193 Cγ2 and Ile255 Cδ1—all more consistent with the *trans* than with the *gauche* conformation for χ3 identified in the crystal structure (PDB ID: 3UXN (5)). Crystallographic determinations of methionine χ3 values are in general limited by the difficulty of defining the position of the terminal methyl group due to the much larger electron density of the bonded sulfur.

A similar analysis of the isoleucine χ2 values based on the ^{13}C shift data in Supplementary Table S1A using the relation given by Hansen *et al.* (32) indicates a strong preference of Ile138 χ2 for the *g*- conformation in both the binary and matched ternary complexes, while residues Ile73, Ile15, Ile293 and Ile323 exhibit strong χ2 *trans* preferences. Importantly, in all of the cases exhibiting significant exchange contributions to R_2, the fitted $\Delta\omega$ term (Table 1) is well below the *gauche*$^-$/*trans* shift differences of 4.5 ppm for methionine χ3 (35) or 5.5 ppm for isoleucine χ2 (32). The interconverting states monitored correspond instead to relatively small changes in preferences for the different conformational states.

DISCUSSION

The binding of the apoenzyme to gapped DNA results in a large reorganization of the enzyme. In solution, the apoenzyme is elongated (36), but wraps around a 1-nt-gapped DNA substrate forming a doughnut like structure (Sawaya *et al.* (4)). This is facilitated by Helix-hairpin-Helix (HhH) motifs found in the lyase domain (residues 55–79) and the D subdomain (residues 92–118) (Pelletier and Sawaya, (3)). These motifs interact with the DNA backbone of the broken strand of the downstream and upstream duplex, respectively. The formation of the binary complex results in changes in the chemical shifts of many of the Ile and Met methyl resonances throughout the enzyme and the broadening of several resonances in the lyase domain and N subdomain. This line broadening is not observed in the apoenzyme. In the apoenzyme, sidechain methyl μs-ms dynamics is restricted to the D subdomain and can be characterized by a single exchange process. This motion may allow the enzyme to sample alternate substrate geometries to facilitate binding of the polymerase domain to the upstream side of the gap, after the initial binding of the lyase domain to the downstream portion of the gap. Previous ^{15}N CPMG relaxation dispersion experiments of the apoenzyme showed that a large number of resonances in both the lyase domain and D subdomain backbone exhibited μs-ms dynamics (12). The timescale of the backbone motion (k_{ex} ∼ 1900 s^{-1}) was similar to the timescale of the sidechain motion found in the D subdomain (k_{ex} ∼ 1400 s^{-1}), although no μs-ms sidechain dynamics was observed for any of the methyl resonances in the lyase domain. The residues that exhibited μs-ms backbone motion in the previous study (12) and μs-ms sidechain motion in the current study are shown in Figure 6. The dynamics observed in the lyase domain suggests that the backbone dynamics may facilitate initial sampling and binding to gapped DNA, and that both sidechain and backbone μs-ms dynamics in the D subdomain contribute to sampling and binding of the polymerase domain to the upstream DNA.

A large transition from backbone to sidechain μs-ms motion occurred upon formation of the binary complex. The number of sidechains observed to be undergoing μs-ms dynamics increased from three in the apoenzyme to 10 in the binary complex (Figure 5B and Table 1), while only Glu21 was observed to be exhibiting backbone μs-ms dynamics in a previous study of a binary complex (12). Two additional sidechain resonances in the lyase domain and Ile277 in the N subdomain of the binary complex exhibited line broadening (Figure 5B and Supplementary Table S1B) due to intermediate conformational exchange. The sidechain dynamics also became distributed throughout the enzyme in the binary complex with minimal coupling of the motion between the different sidechains (Figure 5B). The changes in overall μs-ms dynamics in going from the apoenzyme to the binary complex are summarized in Figure 6. The transition

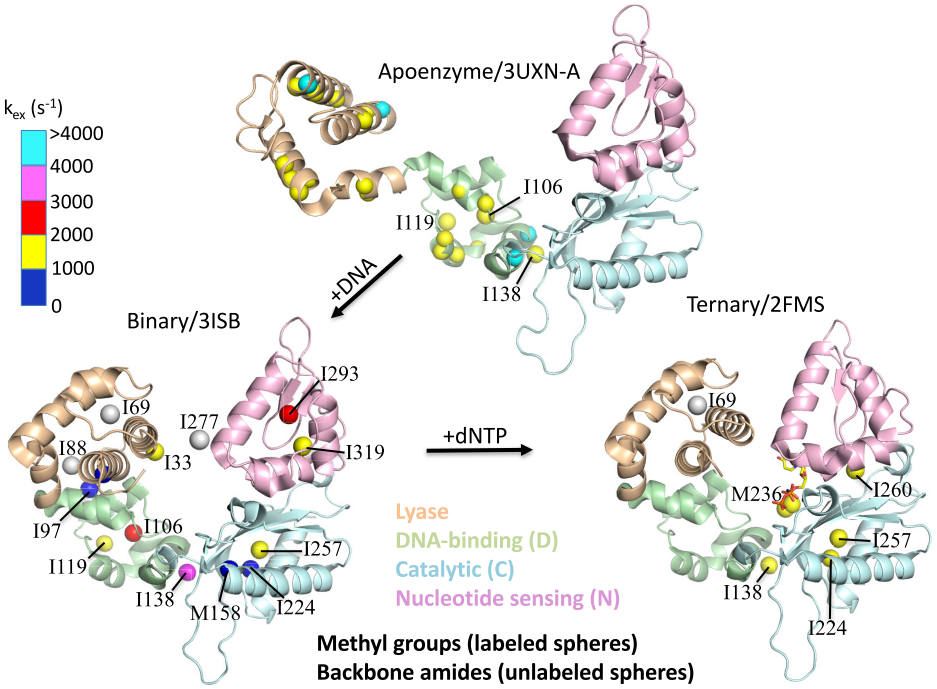

Figure 6. Summary of transitions in pol β μs-ms backbone and sidechain dynamics. The sidechain Ile and Met methyl groups from this study and the backbone amide nitrogen atoms from a previous study (12) that were found to be exhibiting μs-ms dynamics are shown as labeled (methyl) and unlabeled (amide nitrogen) spheres color coded by the magnitude of their exchange rates (k_{ex}). In the previous study of a binary complex, Glu21 was the only backbone residue observed to be undergoing μs-ms motion. In this study of the apoenzyme and ternary complex sidechain dynamics and the previous study of the backbone dynamics of the apoenzyme, the exchange rates were obtained from global fits of the CPMG data, with the exception of four residues (cyan) that showed elevated backbone exchange rates ($k_{ex} > 4000\ s^{-1}$) in the apoenzyme. The exchange rates for the methyl groups in the binary complex were obtained from individual fits of the dispersion curves. The methyl group resonances exhibiting significant exchange line broadening in the binary and ternary complexes are shown as light gray spheres.

from backbone to sidechain μs-ms motion upon formation of the binary complex may be due to a reduction in the overall flexibility of the backbone of the enzyme, on the μs-ms time scale sampled in these experiments, that gives rise to an increased level of sidechain motion. The large reduction in backbone motion and increase in sidechain motion also suggests that the backbone motion is primarily useful for initial DNA binding. Once the DNA complex is formed the sidechain motion may then allow greater flexibility for nucleotide selection and binding. Additional experiments will be required to provide more details.

Formation of a ternary complex involving a correct base pair is associated with conformational activation in which the N subdomain closes around the nascent base-pair, making contact with the amino-terminal lyase domain. (6) This transition is accompanied by shift and linewidth changes of specific methionine and isoleucine residues located primarily in the lyase domain and N and C subdomains, and significant changes in their μs-ms sidechain dynamics. The direct contact between the amino-terminal lyase domain and the carboxyl-terminal N subdomain in the closed complex with the matched nucleotide directly influences the conformation of Ile277, located at this interface, leading to an increased preference for the g^- conformation in order to minimize steric conflicts (Figures 2B and 4A). The formation of the closed ternary complex is also accompanied by a reduction in exchange broadening of the Ile33 and Ile88 methyl resonances in the lyase domain (Figure 2B). Ile33 also experiences a loss in μs-ms sidechain dynamics upon formation of the ternary complex even though it is in similar local environments in both the binary and ternary complexes. These changes to the Ile33, Ile88 and Ile277 resonances all imply that the lyase and N subdomain become more rigid upon closing to form the ternary complex.

In contrast, the HMQC spectrum of the mismatch dT-dGPCPP ternary complex with ∼1 equivalent of dGPCPP is very similar to the spectrum of the binary complex, indicating that the mismatch nucleotide does not readily bind to form the closed ternary complex. Further titration with the mismatch dNTP yielded a large average K_d of ∼1.2 mM (Figure S1A and S1B), with asymptotic limits for the Ile277 and Ile88 methyl resonances that fall well short of the values observed in the matched ternary complex (Figure 2B and Figure S1A). These results indicate that the mismatch nucleotide binds very weakly and does not alter the open to closed conformational equilibrium to the extent that the matched nucleotide does. Based on the limiting shifts of the Ile277 and Ile88 resonances compared to the shifts observed with the matched dNTP, the complex may be closed for ∼33% of the time, effectively reducing the time for the nucleotidyl transfer reaction to occur and reducing the likelihood of mismatch incorporation.

Significant differences between the μs-ms dynamics of the isoleucine and methionine sidechains of the binary and the matched ternary complex are also observed. In the binary complex, the μs-ms sidechain dynamics and line broaden-

ing is distributed throughout the enzyme and characterized by multiple localized exchange processes (Figure 6 and Table 2). When the enzyme adopts the closed ternary complex conformation with the correct incoming nucleotide, the ms sidechain dynamics becomes localized primarily in the C subdomain and can be characterized by a single exchange process (Figure 6 and Table 3). This dynamic behavior is similar to that observed with dihydrofolate reductase (DHFR), in which more widespread uncoupled μs-ms backbone motion is observed in the DHFR–NADPH binary complex, that transitions to more localized coupled motion around the methotrexate or trimethoprim inhibitor binding site upon formation of the ternary complexes (37). There may be other enzymes that exhibit similar dynamic transitions as well. Importantly, the results for pol β indicate that the closed ternary complex does not exist as a single stable state. Interestingly, a simple two-state model imposes restrictions on the observed rate of catalysis. If the major state ($p_A \sim 98\%$) of the wild-type enzyme is in the active conformation, then changes to the dynamic behavior of the enzyme through mutagenesis should lead to a decrease in activated complex (i.e. decrease in k_{pol}) since the system is highly optimized. However, if the minor population (p_B) corresponds to the catalytically activated enzyme, then altered enzyme dynamics could result in a decrease or increase in the activated population resulting in a proportional change in k_{pol}. The observation that an active site mutation (R258A) leads to an increase in k_{pol} (6) is consistent with the latter interpretation. The similarity of the forward rate of exchange measured from the CPMG experiments ($k_{AB} \sim 22\ s^{-1}$) and reported rates of nucleotide incorporation k_{pol} (33,34), and the localization of the μs-ms dynamics in the C subdomain are also consistent with this interpretation.

Although the forward rate k_{AB} of the exchange process obtained from a global fit of the ternary complex CPMG data is similar to reported values of k_{pol}, there is not enough information available to ascertain whether the dominant higher energy state sampled by this exchange process is related to the transition state. However, the localization of the exchange process to the catalytic domain is suggestive that the dynamics may be related to catalysis. The role of dynamic fluctuations in promoting catalysis has been the subject of extensive investigations over the past several years, many of which utilized the *E. coli* enzyme DHFR as a model system (38). Recently, it has also been shown that a Michaelis complex of arginine kinase exchanges with the transition state at a rate that matches the catalytic rate (39). In this case, the $\Delta\omega_C$ values obtained from the CPMG data also matched the equilibrium chemical shift differences $\Delta\delta_C$ measured between the Michaelis complex and a transition state analog, demonstrating that the Michaelis complex samples a higher energy conformation that is similar to the transition state analog bound state. Alternatively, dynamics may play an indirect role where alternate conformational transitions improve the ability of the polymerase to productively interact with different dNTP molecules that vary with each catalytic event.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Lalith Perera and Dr. Jessica L. Wojtaszek for careful reading of the manuscript and Lois Wyrick for help in preparing the figures.

FUNDING

Intramural Research Program of the NIH, National Institute of Environmental Health Sciences [ZIA ES050111 to R.E.L., Z01 ES050158, ES050159 to S.H.W.]; NIH and NIEHS [HHSN273200700046U to E.F.D.]. Funding for open access charge: Intramural Research Program of the NIH, National Institute of Environmental Health Sciences [ZIA ES050111].
Conflict of interest statement. None declared.

REFERENCES

1. Beard,W.A. and Wilson,S.H. (2014) Structure and mechanism of DNA polymerase beta. *Biochemistry*, **53**, 2768–2780.
2. Beard,W.A. and Wilson,S.H. (2006) Structure and mechanism of DNA polymerase beta. *Chem. Rev.*, **106**, 361–382.
3. 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.
4. 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.
5. Gridley,C.L., Rangarajan,S., Firbank,S., Dalal,S., Sweasy,J.B. and Jaeger,J. (2013) Structural changes in the hydrophobic hinge region adversely affect the activity and fidelity of the I260Q mutator DNA polymerase beta. *Biochemistry*, **52**, 4422–4432.
6. Beard,W.A., Shock,D.D., Batra,V.K., Prasad,R. and Wilson,S.H. (2014) Substrate-induced DNA polymerase beta activation. *J. Biol. Chem.*, **289**, 31411–31422.
7. Freudenthal,B.D., Beard,W.A., Shock,D.D. and Wilson,S.H. (2013) Observing a DNA polymerase choose right from wrong. *Cell*, **154**, 157–168.
8. Batra,V.K., Beard,W.A., Shock,D.D., Pedersen,L.C. and Wilson,S.H. (2008) Structures of DNA polymerase beta with active-site mismatches suggest a transient abasic site intermediate during misincorporation. *Mol. Cell*, **30**, 315–324.
9. Palmer,A.G. (2004) NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.*, **104**, 3623–3640.
10. Bose-Basu,B., DeRose,E.F., Kirby,T.W., Mueller,G.A., Beard,W.A., Wilson,S.H. and London,R.E. (2004) Dynamic characterization of a DNA repair enzyme: NMR studies of [methyl-C-13]methionine-labeled DNA polymerase beta. *Biochemistry*, **43**, 8911–8922.
11. Kirby,T.W., DeRose,E.F., Cavanaugh,N.A., Beard,W.A., Shock,D.D., Mueller,G.A., Wilson,S.H. and London,R.E. (2012) Metal-induced DNA translocation leads to DNA polymerase conformational activation. *Nucleic Acids Res.*, **40**, 2974–2983.
12. Berlow,R.B., Swain,M., Dalal,S., Sweasy,J.B. and Loria,J.P. (2012) Substrate-Dependent millisecond domain motions in DNA polymerase beta. *J. Mol. Biol.*, **419**, 171–182.
13. Moscato,B., Swain,M. and Loria,J.P. (2016) Induced fit in the selection of correct versus incorrect nucleotides by DNA polymerase beta. *Biochemistry*, **55**, 382–395.
14. 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.

15. Kirby,T.W., DeRose,E.F., Beard,W.A., Shock,D.D., Wilson,S.H. and London,R.E. (2014) Substrate rescue of DNA polymerase beta containing a catastrophic L22P mutation. *Biochemistry*, **53**, 2413–2422.
16. Vande Berg,B.J., Beard,W.A. and Wilson,S.H. (2001) DNA structure and aspartate 276 influence nucleotide binding to human DNA polymerase beta - Implication for the identity of the rate-limiting conformational change. *J. Biol. Chem.*, **276**, 3408–3416.
17. Upton,T.G., Kashemirov,B.A., McKenna,C.E., Goodman,M.F., Prakash,G.K., Kultyshev,R., Batra,V.K., Shock,D.D., Pedersen,L.C., Beard,W.A. *et al.* (2009) Alpha,beta-difluoromethylene deoxynucleoside 5′-triphosphates: a convenient synthesis of useful probes for DNA polymerase beta structure and function. *Org. Lett.*, **11**, 1883–1886.
18. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) Nmrpipe - a multidimensional spectral processing system based on unix pipes. *J. Biomol. NMR.*, **6**, 277–293.
19. Johnson,B.A. and Blevins,R.A. (1994) Nmr View - a computer-program for the visualization and analysis of Nmr data. *J. Biomol. NMR*, **4**, 603–614.
20. Tugarinov,V. and Kay,L.E. (2003) Side chain assignments of Ile delta 1 methyl groups in high molecular weight proteins: An application to a 46 ns tumbling molecule. *J. Am. Chem. Soc.*, **125**, 5701–5706.
21. Mueller,G.A., DeRose,E.F., Kirby,T.W. and London,R.E. (2007) NMR assignment of polymerase beta labeled with H-2, C-13, and N-15 in complex with substrate DNA. *Biomol. NMR Assign.*, **1**, 33–35.
22. Tugarinov,V., Kay,L.E., Ibraghimov,I. and Orekhov,V.Y. (2005) High-resolution four-dimensional H-1-C-13 NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. *J. Am. Chem. Soc.*, **127**, 2767–2775.
23. Batra,V.K., Beard,W.A., Shock,D.D., Krahn,J.M., Pedersen,L.C. and Wilson,S.H. (2006) Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure*, **14**, 757–766.
24. Xiao,Y., Lee,T., Latham,M.P., Warner,L.R., Tanimoto,A., Pardi,A. and Ahn,N.G. (2014) Phosphorylation releases constraints to domain motion in ERK2. *Protein Sci.*, **23**, 263–264.
25. Gryk,M.R., Maciejewski,M.W., Robertson,A., Mullen,M.A., Wilson,S.H. and Mullen,G.P. (2002) Letter to the Editor: H-1, C-13 and N-15 resonance assignments for the perdeuterated 22 kD palm-thumb domain of DNA polymerase beta. *J. Biomol. NMR*, **22**, 197–198.
26. Markley,J.L., Ulrich,E.L., Berman,H.M., Henrick,K., Nakamura,H. and Akutsu,H. (2008) BioMagResBank (BMRB) as a partner in the Worldwide Protein Data Bank (wwPDB): new policies affecting biomolecular NMR depositions. *J Biomol NMR*, **40**, 153–155.
27. Korzhnev,D.M., Kloiber,K., Kanelis,V., Tugarinov,V. and Kay,L.E. (2004) Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: Application to a 723-residue enzyme. *J. Am. Chem. Soc.*, **126**, 3964–3973.
28. Morin,S., Linnet,T.E., Lescanne,M., Schanda,P., Thompson,G.S., Tollinger,M., Teilum,K., Gagne,S., Marion,D., Griesinger,C. *et al.* (2014) relax: the analysis of biomolecular kinetics and thermodynamics using NMR relaxation dispersion data. *Bioinformatics*, **30**, 2219–2220.
29. Akaike,H. (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov,B.N. and Csaki,F. (eds). *Proceedings of the Second International Symposium on Information Theory*. Akademiai Kiado, Budapest, pp. 267–281.
30. Beard,W.A., Shock,D.D., Batra,V.K., Pedersen,L.C. and Wilson,S.H. (2009) DNA polymerase beta substrate specificity side chain modulation of the 'A-Rule'. *J. Biol. Chem.*, **284**, 31680–31689.
31. London,R.E., Wingad,B.D. and Mueller,G.A. (2008) Dependence of amino acid side chain (13)C shifts on dihedral angle: Application to conformational analysis. *J. Am. Chem. Soc.*, **130**, 11097–11105.
32. Hansen,D.F., Neudecker,P. and Kay,L.E. (2010) Determination of isoleucine Side-Chain conformations in ground and excited states of proteins from chemical shifts. *J. Am. Chem. Soc.*, **132**, 7589–7591.
33. Ahn,J., Werneburg,B.G. and Tsai,M.D. (1997) DNA polymerase beta: Structure-fidelity relationship from pre-steady-state kinetic analyses of all possible correct and incorrect base pairs for wild type and R283A mutant. *Biochemistry*, **36**, 1100–1107.
34. Ahn,J.W., Kraynov,V.S., Zhong,X.J., Werneburg,B.G. and Tsai,M.D. (1998) DNA polymerase beta: effects of gapped DNA substrates on dNTP specificity, fidelity, processivity and conformational changes. *Biochem. J.*, **331**, 79–87.
35. Butterfoss,G.L., DeRose,E.F., Gabel,S.A., Perera,L., Krahn,J.M., Mueller,G.A., Zheng,X.H. and London,R.E. (2010) Conformational dependence of C-13 shielding and coupling constants for methionine methyl groups. *J. Biomol. NMR*, **48**, 31–47.
36. Kim,S.J., Lewis,M.S., Knutson,J.R., Porter,D.K., Kumar,A. and Wilson,S.H. (1994) Characterization of the tryptophan fluorescence and hydrodynamic properties of rat DNA-polymerase-beta. *J. Mol. Biol.*, **244**, 224–235.
37. Mauldin,R.V., Carroll,M.J. and Lee,A.L. (2009) Dynamic dysfunction in dihydrofolate reductase results from antifolate drug binding: modulation of dynamics within a structural state. *Structure*, **17**, 386–394.
38. Boehr,D.D., McElheny,D., Dyson,H.J. and Wright,P.E. (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science*, **313**, 1638–1642.
39. Peng,Y., Hansen,A.L., Bruschweiler-Li,L., Davulcu,O., Skalicky,J.J., Chapman,M.S. and Bruschweiler,R. (2017) The michaelis complex of arginine kinase samples the transition state at a frequency that matches the catalytic rate. *J. Am. Chem. Soc.*, **139**, 4846–4853.
40. Rosenzweig,R., Moradi. S,, Zarrine-Afsar,A., Glover,J.R. and Kay,L.E. (2013) Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. *Science*, **339**, 1080–1083.