JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

How a Low-Fidelity DNA Polymerase Chooses Non-Watson–Crick from Watson–Crick Incorporation

Wen-Jin Wu,[†,▽] Mei-I Su,[†,‡,▽] Jian-Li Wu,[†] Sandeep Kumar,[§,○] Liang-hin Lim,[†,∥] Chun-Wei Eric Wang,[†,∥] Frank H. T. Nelissen,[⊥] Ming-Chuan Chad Chen,[†] Jurgen F. Doreleijers,[#] Sybren S. Wijmenga,[⊥] and Ming-Daw Tsai[†,‡,§,∥,*]

[†]Institute of Biological Chemistry, and [‡]Genomics Research Center, Academia Sinica, 128 Academia Road Sec. 2, Nankang, Taipei 115, Taiwan

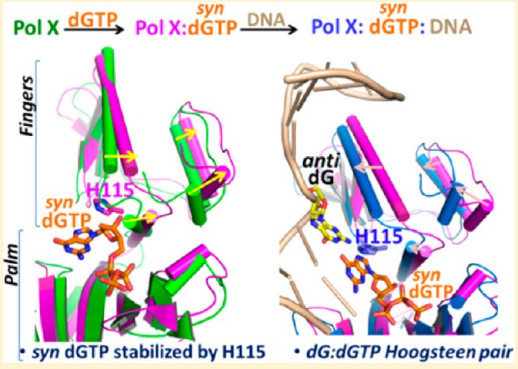
[§]Department of Chemistry, The Ohio State University, 100 West 18th Avenue, Columbus, Ohio 43210, United States

[∥]Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan

[⊥]Department of Biophysical Chemistry, and [#]Centre for Molecular & Biomolecular Informatics, Radboud University, Nijmegen 6525, The Netherlands

Supporting Information

ABSTRACT: A dogma for DNA polymerase catalysis is that the enzyme binds DNA first, followed by MgdNTP. This mechanism contributes to the selection of correct dNTP by Watson–Crick base pairing, but it cannot explain how low-fidelity DNA polymerases overcome Watson–Crick base pairing to catalyze non-Watson–Crick dNTP incorporation. DNA polymerase X from the deadly African swine fever virus (Pol X) is a half-sized repair polymerase that catalyzes efficient dG:dGTP incorporation in addition to correct repair. Here we report the use of solution structures of Pol X in the free, binary (Pol X:MgdGTP), and ternary (Pol X:DNA:MgdGTP with dG:dGTP non-Watson–Crick pairing) forms, along with functional analyses, to show that Pol X uses multiple unprecedented strategies to achieve the mutagenic dG:dGTP incorporation. Unlike high fidelity polymerases, Pol X can prebind purine MgdNTP tightly and undergo a specific conformational change in the absence of DNA. The prebound MgdGTP assumes an unusual *syn* conformation stabilized by partial ring stacking with His115. Upon binding of a gapped DNA, also with a unique mechanism involving primarily helix αE, the prebound *syn*-dGTP forms a Hoogsteen base pair with the template *anti*-dG. Interestingly, while Pol X prebinds MgdCTP weakly, the correct dG:dCTP ternary complex is readily formed in the presence of DNA. H115A mutation disrupted MgdGTP binding and dG:dGTP ternary complex formation but not dG:dCTP ternary complex formation. The results demonstrate the first solution structural view of DNA polymerase catalysis, a unique DNA binding mode, and a novel mechanism for non-Watson–Crick incorporation by a low-fidelity DNA polymerase.



INTRODUCTION

DNA polymerases play central roles in genomic replication and maintenance. For the better part of the last century, researchers strived to elucidate the structural and mechanistic basis for high fidelity of DNA polymerases.[1−9] A key factor contributing to the high fidelity of the catalysis is the Watson–Crick base pairing between the template base and the incoming dNTP. In the past decade, many translesion synthesis (TLS) or mutagenic DNA polymerases with important biological functions have been discovered,[10−13] and the question has been reversed to how these enzymes catalyze their specialized reactions or non-Watson–Crick incorporations. Extensive structural studies have revealed a wealth of unique features for individual enzymes.[7,8,14−20] However, the fundamental question of how the enzymes overcome Watson–Crick base pairing to catalyze mismatch incorporation remains poorly understood.

The deadly African swine fever virus (ASFV) is a double stranded DNA virus that has caused major economic losses, and in recent years it is turning into a global threat.[21−23] The DNA polymerase X encoded by ASFV[24] (referred to as "Pol X" hereafter) is highly prone to a dG:dGTP mismatch incorporation with an efficiency approaching that of a dG:dCTP pair on the basis of the studies from our laboratory,[25] although subsequent reports suggested that the dG:dGTP preference (relative to dG:dCTP) varies with the condition used.[26−28] Another unique feature of Pol X is that it is a half-sized DNA polymerase (only 174 residues) lacking the lyase domain and the duplex DNA binding subdomain of its mammalian homolog DNA polymerase β (Pol β). Pol X is therefore considered as the minimal version of an evolutionarily

Received: October 5, 2013
Published: March 11, 2014

dx.doi.org/10.1021/ja4102375 | J. Am. Chem. Soc. 2014, 136, 4927−4937

conserved Pol β-like DNA polymerase core.[8,24] The structure of free Pol X, solved by NMR in 2001,[29,30] consists of only two subdomains: an N-terminal subdomain (residues 1−105) and a C-terminal subdomain (residues 106−174). Several unique questions are pertinent to the structure and mechanism of Pol X: how it catalyzes the non-Watson−Crick dG:dGTP incorporation, whether its overall low fidelity reflects poor enzyme-dNTP interactions, how it binds DNA without the DNA-binding subdomains, and whether the "half-sized polymerase" Pol X undergoes MgdNTP-induced subdomain-closing conformational change characteristic of other DNA polymerases. We and others have actively pursued these questions by various approaches in the past decade,[26,31−36] but the progress has been limited because of the lack of structure of the enzyme:DNA:MgdNTP ternary complex.

Although NMR has never been used previously to solve the structure of the ternary complex of a DNA polymerase because of the large size of the complex and the chemical exchange of multiple components, we have achieved the challenging task of determining the solution structure of the ternary complex (Pol X:DNA:MgdGTP) with a dG:dGTP mismatch. In addition, since Pol X seems to be able to bind MgdNTP in the absence of DNA[29,30] and its kinetics seems to deviate from the ordered sequential mechanism (DNA binding first followed by dNTP),[37] we also solved the solution structure of the elusive Pol X:MgdGTP binary complex. For comparison with the new structures, the previously reported structure of the free Pol X was also refined with residual dipolar coupling (RDC) data. The results indicate several novel structural features for Pol X, which were further supported by functional analyses. The results indicate that Pol X uses multiple unprecedented strategies to achieve the mutagenic dG:dGTP incorporation.

MATERIALS AND METHODS

This section describes only a summary of key materials and samples used and the major experimental methods. Detailed experimental procedures are described in Supporting Information (SI) Materials and Methods.

Protein Samples for Structural Determination by NMR. Various uniformly and specifically labeled Pol X samples are prepared as described in SI Materials and Methods. The NMR samples of free Pol X (^{15}N-Pol X, ^{13}C,^{15}N-Pol X, ^{2}H,^{13}C,^{15}N-Pol X) used for refining the previously published reduced form-Pol X structure contain 50 mM KCl, 50 mM potassium borate, 10 mM DTT-$_{d10}$ at pH 6.5. The Pol X:MgdGTP binary complex was prepared by mixing 1−8 mol equiv of MgdGTP (indicated in figure legends) with Pol X in buffer containing 50 mM KCl, 50 mM potassium borate, 10 mM $MgCl_2$, 10 mM DTT-$_{d10}$ at pH 6.5. The Pol X:dhDNA:MgdGTP ternary complex for structural determination was prepared by first mixing 8 mol equiv of MgdGTP to Pol X, followed by addition of 1.2 mol equiv of DNA to the solution containing 10 mM DTT-$_{d10}$, 200 mM KCl, 50 mM potassium borate, 10 mM $MgCl_2$ at pH 6.5. Detailed conditions for experiments used to demonstrate specific properties may differ from these main conditions for structural determination, as described in specific figure legends.

Isotope-Labeled Double Hairpin DNA Samples for NMR. Two selectively isotope-labeled DNA samples were prepared via enzymatic synthesis (Figure S8 (SI)): DNA sample #1, 5′PO_3-$G^1G^2C^3G^4A^5A^6G^7C^8C^9\underline{G^{10}G^{11}G^{12}T^{13}G^{14}}C^{15}G^{16}A^{17}A^{18}G^{19}C^{20}A^{21}$-$C^{22}C^{23}$-3′dd, where 3′dd stands for dideoxy for the C3′ atom at C^{23}, and DNA sample #2, 5′PO_3-$G^1G^2C^3G^4A^5A^6G^7C^8C^9G^{10}G^{11}$-$G^{12}T^{13}G^{14}C^{15}\underline{G^{16}A^{17}A^{18}G^{19}C^{20}A^{21}C^{22}}C^{23}$-3′dd. In DNA #1 the underlined segment is labeled with ^{13}C and ^{15}N, while in DNA #2 the underlined segment was labeled with ^{13}C, ^{15}N, and deuterons at selective sites in order to reduce sensitivity loss due to fast transverse relaxation of the Pol X:DNA:MgdGTP ternary complex.

NMR Spectroscopy. NMR experiments were performed on Bruker AV600, AV800, and AVIII850 MHz spectrometers equipped with cryogenic probes. All NMR samples contained 5−10% D_2O for spectrometer lock. All spectra were processed using Topspin (Bruker) and analyzed with the program CARA[38] (www.nmr.ch). Resonance assignments were based on triple resonance experiments. Residual dipolar couplings (RDCs) were measured for partially aligned samples in Pf1-phage and/or C12E5/hexanol.

Structure Calculation. Briefly, the structure of free Pol X was refined from the previously reported reduced form[29] by applying the residual dipolar coupling (RDC) data of backbone N−H. The structure of Pol X:MgdGTP binary complex was calculated from restraints of intramolecular NOEs and backbone dihedral angles and 13 intermolecular NOEs between Pol X and dGTP (Table S1 (SI)), and refined with backbone N−H RDC data. The structure of Pol X in the ternary complex was calculated from the MgdGTP-bound binary Pol X structure using backbone N−H RDC data of the ternary complex. The CYANA calculated ternary Pol X structure was then subjected to HADDOCK[39,40] modeling for the complex with gapped DNA and the incoming nucleotide MgdGTP, utilizing the restraints obtained from NMR experiments including chemical shift perturbation data and intermolecular NOEs between Pol X, DNA, and dGTP (Table S1 (SI)). All structural calculations using experimentally determined distances and dihedral and dipolar coupling restraints (Table S1−S5 (SI)) were performed with CYANA 3.0,[41] and the final ensembles of NMR structures were subjected to explicit water refinement under HADDOCK-CNS.[42] The crystal structure of the Pol β:DNA:MgdNTP ternary complex consists of two Mg^{2+} ions at the active site,[8,43,44] based on the studies described in SI Materials and Methods (including ITC results of D49S, D51S, and D100S mutants shown in Figure S3F−I (SI)), we assumed that both Mg^{2+} ions are present in the binary complex and used the distance restraints from the crystal structure of Pol β for calculation.

The chemical shift assignments for free Pol X, Pol X:MgdGTP binary complex, and Pol X in the Pol X:MgdGTP:DNA ternary complex have been deposited in the BioMagResBank under access numbers 18933, 18934, and 18935, respectively. The coordinates of free Pol X, Pol X:MgdGTP binary complex, Pol X in the Pol X:MgdGTP:DNA ternary complex, and HADDOCK-calculated Pol X:MgdGTP:DNA ternary complex have been deposited in the RCSB Protein Data Bank with access codes 2m2t, 2m2u, 2m2v, and 2m2w, respectively.

RESULTS

Solution Structures of Free, Binary, and Ternary Complexes of Pol X by NMR. We have used NMR to determine the solution structures of Pol X in the free (refined with backbone N−H RDC), binary (Pol X:MgdGTP), and ternary (Pol X:DNA:MgdGTP) forms. Detailed procedures are described in the Materials and Methods and the SI Materials and Methods, which involve specific isotope labeling of the protein, DNA, and dNTP.

Ribbon diagrams and ensembles of these structures are shown in Figure 1A (free Pol X), Figure 1B (Pol X:MgdGTP binary complex), and Figure 1C (Pol X:DNA:MgdGTP ternary complex). The stereo views of the ensemble structures are shown in Figure S1A (SI) (free), Figure S1B (SI) (binary), Figure S1C (SI) (ternary, protein only, before HADDOCK), Figure S1D (SI) (ternary, protein only, after HADDOCK), and Figure S1E (SI) (Pol X:DNA:MgdGTP). The structural statistics are listed in Tables S2−S5 (SI). As shown by the superimposition of the 2D ^{1}H,^{15}N-TROSY-HSQC spectra of the three forms (Figure S1F (SI)), there are notable differences. The unique structural features of the binary and the ternary complexes will be addressed in later sections. The sequence of Pol X and its secondary structural elements are shown in Figure S1G (SI). The DNA used in the ternary complex is a single-

dx.doi.org/10.1021/ja4102375 | J. Am. Chem. Soc. 2014, 136, 4927−4937

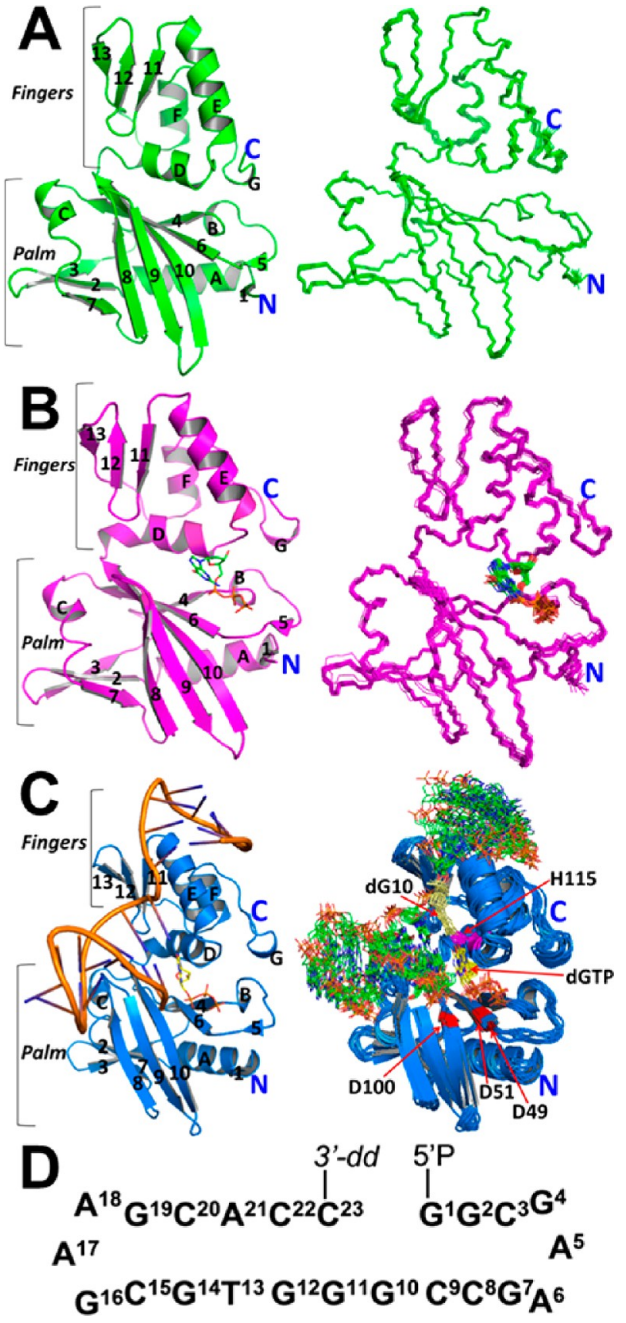


Figure 1. Solution structures of Pol X in different states. (A) Free Pol X (ribbon diagram and ensemble of 20 structures). The N-terminal subdomain (residue 1−105) is also referred to as "Palm", and the C-terminal subdomain (residue 106−174) as "Fingers" according to our previous reports[9,30] and the nomenclature by Steitz et al. for DNA polymerases based on functional alignment (left-handed analogy).[45] Note that the "Fingers" would correspond to the "Thumb" in the right-hand analogy.[4,46] (B) Pol X:MgdGTP binary complex (ribbon diagram and ensemble of 20 structures). (C) HADDOCK calculated structure of the Pol X:MgdGTP:DNA ternary complex. The right figure shows the ensemble of 10 structures. (D) Sequence of the double hairpin gapped DNA used in (C). The nucleotide C23 is dideoxy at its 2′ and 3′ carbons in order to prevent the reaction from taking place. The reducing condition was used for all structural studies (in the presence of 10 mM DTT).

stranded, double hairpin DNA (abbreviated as "dhDNA" or simply "DNA" hereafter) that forms a single-nucleotide gap steadily as shown in Figure 1D. Such a DNA has been shown previously to bind to Pol β almost identically to the corresponding gapped DNA formed by three pieces of oligonucleotides.[47] The structure of Pol X in the ternary complex was particularly challenging because of the chemical exchange between three components. This is the first case of a protein/gapped DNA/MgdNTP ternary complex solved by solution NMR.

dGTP Adopts a *syn* Conformation Stabilized by His115 in the Binary Complex. A major finding of this work is that, in the absence of DNA, MgdGTP binds to Pol X in an unusual *syn* conformation about the glycosidic bond, suggesting that the binding is specific. Mg^{2+} is required for the dGTP binding, as dGTP alone caused no detectable shifts in the HSQC spectrum of Pol X. As shown in Figure 2A (surface charge representation) and Figure 2B (stick representation), the MgdGTP binding cleft formed between the two subdomains is narrow, forcing dGTP to adopt the more compact *syn* conformation (compared to the extended *anti* conformation). Some of the evidence for the unusual *syn* conformation of the bound nucleotide is summarized in Figure 2C-a. Using ^{13}C,^{15}N-labeled MgdNTP, we observed stronger H8−H1′ NOE than the H8−H2′ or H8−H2″ NOE (Figure 2C-b) as the H8 proton is closer to H1′ than to H2′ or H2″ in the *syn* conformation.[48] In addition, ^{31}P NOE difference data of MgdGTP (Figure 2C-c) showed NOE between P_α and the NH_2 protons of dGTP, suggesting an intramolecular H-bond as indicated in Figure 2C-a, a unique stabilizing factor only possible in *syn* purine nucleotides. The dGTP-purine/His115 imidazole interaction through partial pi−pi stacking (Figure 2B) appears to be critical for the binding of *syn* dGTP. The partial pi−pi stacking is supported by the intermolecular NOEs between His115 Cδ2 proton and H1′ and H8 protons of dGTP (Figure 2C-d), and by the ring current shift of the H8 proton of the bound *syn* dGTP (from 8.1 ppm to 6.25 ppm) (Figure S2A (SI)). We then constructed the H115A mutant to test the importance of this partial pi−pi stacking in dGTP binding. The MgdGTP affinity of Pol X decreased by >20-fold in H115A (rows 1 and 5 of Table 1 and Figure S3 (SI)). In addition, the upfield shift of the H8 proton was completely abolished (Figure S2B (SI)). These results support that His115 is the primary residue responsible for stabilizing *syn* dGTP. In addition, the surrounding residues shown in Figure 2B appear to collectively provide a hydrophobic pocket, which is supported by intermolecular NOEs between dGTP and Pol X residues (Figure 2C-d,e and Table S1 (SI)), and dense inter-residue NOEs between the Pol X residues shown here (NOE table available from RCSB Protein Data Bank under access code 2m2u). Furthermore, the calculated binary structure shows a possible hydrogen bond between dGTP 3′-OH and Val120 backbone NH as indicated in Figure 2B. This interaction was supported by intermolecular NOEs between Val120 γ2 methyl protons and dGTP ribose protons H1′, H2′, H2″, and H3′ (Figure 2C-e), and the chemical shift change of Val120 N−H (from 8.87 ppm to 9.42 ppm).

While the *syn* conformation of bound dGTP was well demonstrated by the NOE analysis, Nikolova et al. recently reported that the *anti*-to-*syn* transition of a dNTP is characterized by a ca. +2 ppm shift of C8 and a ca. +3 ppm shift of C1′.[49] The corresponding shifts observed for dGTP upon binding to Pol X were +1.88 and −1.25 ppm, respectively (Figure S2A (SI)). The differences can be reasonably explained by the different chemical environments in the double stranded DNA[49] and in the active site of Pol X.

dx.doi.org/10.1021/ja4102375 | J. Am. Chem. Soc. 2014, 136, 4927−4937

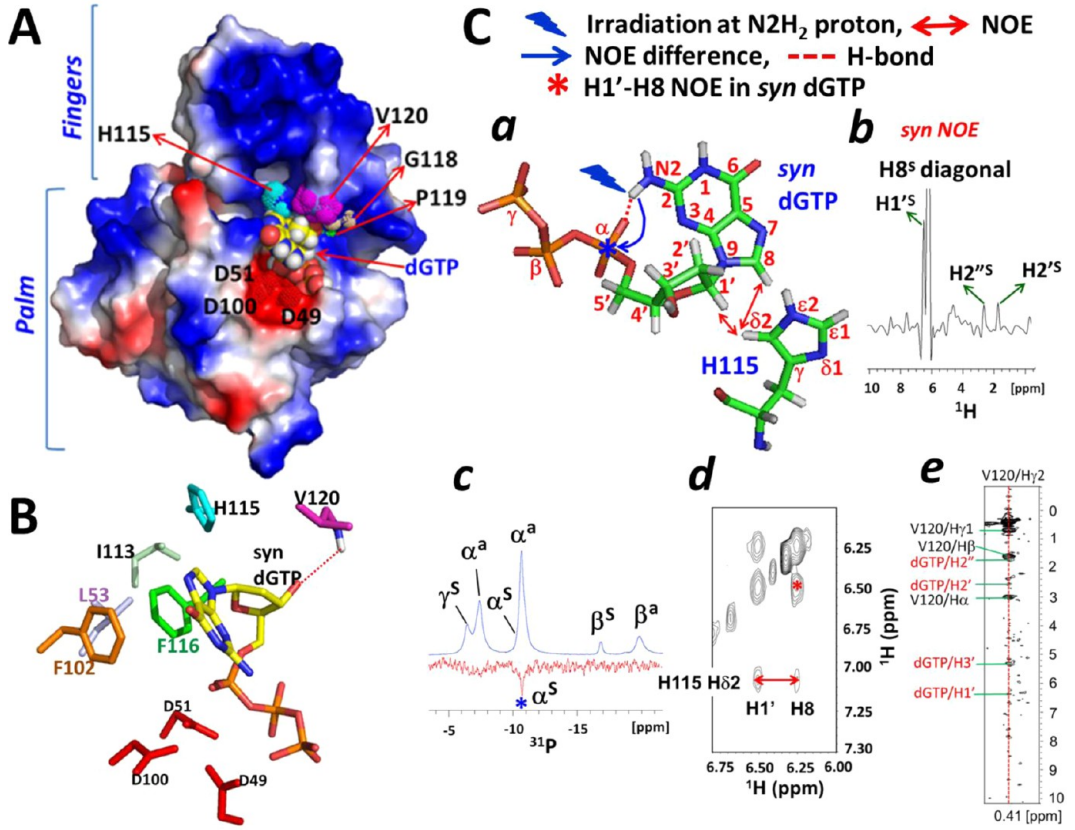


Figure 2. Detailed structure and NMR evidence for the Pol X:MgdGTP binary complex. (A) Surface charge representation of Pol X:MgdGTP binary complex with positive, negative, and neutral potentials shown in blue, red, and white color, respectively. (B) Stick representation of the active site residues important for MgdGTP binding. The red dashed line indicates an intermolecular H-bond between dGTP 3′-OH and V120 backbone N−H. (C) NMR evidence for *syn*-dGTP conformation and interaction between dGTP and His115 and Val120. (a) The stick structure of dGTP showing the *syn* conformation and the relationship with His115, as well as key experimental evidence. (b) Column vectors of H8 proton in the *syn* form from the first 2D $^{1}H-^{1}H$ plane of a 3D ^{13}C-NOESY-HSQC (100 ms mixing time) of [^{13}C,^{15}N-His]-U-[^{2}H]-Pol X:[^{13}C,^{15}N]-MgdGTP (1:1) binary complex, showing stronger H8−H1′ NOE than the H8−H2′ or H8−H2″ NOE. (c) [Proton saturation]-1D ^{31}P NOE difference experiment of Pol X:MgdGTP (1:8) binary complex. A NOE effect was observed for P_α (blue * in the lower spectrum) upon irradiation of NH_2 protons (blue lightning bolt symbol in (a)) at 6.2 ppm in the *syn* form. The upper spectrum shows the off-resonance irradiation spectrum (irradiation at 30 ppm). (d) A 2D plane of a ^{13}C-NOESY-HSQC experiment showing intermolecular NOE between the Cδ2 proton of His115 and H1′ and H8 of the *syn* dGTP. (e) NOE strip of V120 Hγ2 methyl protons from a 3D ^{13}C-NOESY-HSQC (120 ms mixing time) showing its intermolecular NOEs to dGTP ribose protons of H1′, H2′, H2″, and H3′.

Pol X binds dATP similarly to dGTP but binds pyrimidine nucleotides weakly. Isothermal titration calorimetry (ITC) analysis showed that Pol X binds purine MgdNTP more tightly than pyrimidine MgdNTP by a factor of 20−40 in the absence of DNA (Figure S3A−D (SI) and rows 1−4 and 8−11 of Table 1). This differs from the previous report that Pol X binds the four dNTPs with similar affinity.[29] The higher affinity for purine than pyrimidine nucleotides is also reflected in the 2D $^{1}H,^{15}N$-HSQC spectra of Pol X binary complexes (Figure S4A−C (SI)), which show that MgdATP and MgdGTP induce nearly identical shifts, while MgdCTP and MgdTTP induce multiple sets of relatively minor shifts. Because of the weak and nonspecific binding, the binary complexes with MgdCTP or MgdTTP were not further examined. Using $^{13}C,^{15}N$-MgdATP, we observed that the bound MgdATP also exists in the *syn* conformation in the WT but not in the H115A mutant (Figure S2C,D (SI)). In addition, binding of MgdATP to the H115A mutant is weakened by 7.4-fold relative to the WT (rows 9 and 12 of Table 1). The results together suggest that the structure of Pol X:MgdATP should be very similar to that of Pol X:MgdGTP. The above results suggest that binding of dGTP and dATP is a specific property of Pol X. This is a unique feature for Pol X since the full-length Pol β and its C-terminal 22 kDa domain corresponding to Pol X have previously been reported not to bind MgdNTP.[29] To ensure a fair comparison between Pol β and Pol X under the same condition, we also examined the binding property of full-length Pol β by NMR (Figure S5A (SI)) and ITC (Figure S5C (SI) and row 6 of Table 1), and confirmed that it does not bind MgdNTP tightly.

MgdGTP Binding in the Absence of DNA Induces Conformational Change Different from That of Pol β in the Presence of DNA. Upon MgdGTP binding, helix αD moves away from the Palm subdomain in order to create space for the bulky incoming MgdGTP, which also causes outward movements of helices αE, αF, and αG (Figure 3A). This conformational change of Pol X in the absence of DNA is of comparable scale to that of Pol β and many other polymerases in the presence of DNA (Figure 3B, showing only the subdomains corresponding to Pol X). However, as indicated by the yellow arrows, the direction of the conformational change is different between Figure 3A and B. The MgdNTP induced conformational change of Pol β is a subdomain "closure" mode, whereas in Pol X it is a subdomain "opening" mode in order to accommodate the *syn* dGTP. A closer view with ensemble

dx.doi.org/10.1021/ja4102375 | J. Am. Chem. Soc. 2014, 136, 4927−4937

Table 1. K_d Values for the Binding of dNTP (Rows 1−14) or dhDNA (Rows 15−18) to DNA Polymerases Determined by ITC[a]

row	enzyme	ligand	K_d (μM)
Buffer: 2 mM MgCl2, 400 mM KCl, pH 7.2			
1	Pol X	dGTP	1.1 ± 0.1
2	Pol X	dATP	2.1 ± 0.3
3	Pol X	dCTP	41 ± 6
4	Pol X	dTTP	42 ± 5
5	H115A	dGTP	25 ± 14
6	Pol β	dNTP[d]	n.d.[b]
7	Pol β, Y271H	dNTP[e]	n.d.[b]
Buffer: 10 mM MgCl2, 200 mM KCl, pH 8.0			
8	Pol X	dGTP	0.26 ± 0.07
9	Pol X	dATP	0.46 ± 0.04
10	Pol X	dCTP	9.4 ± 0.1
11	Pol X	dTTP	8.61 ± 1.55
12	H115A	dATP	3.4 ± 0.5
13	H115Y	dGTP	0.38 ± 0.07[c]
14	H115Y	dCTP	5.9 ± 0.9[c]
15	Pol β	dhDNA	13.8 ± 0.5
16	Pol X	dhDNA	4.0 ± 0.1
17	R125Q, R127Q	dhDNA	n.d.[b]
18	R125A, R127A	dhDNA	n.d.[b]

[a]All samples contain 50 mM Tris-borate or Tris/HCl and 2−10 mM β-mercaptoethanol. The ± signs stand for standard deviation. [b]n.d.: nondetectable by ITC measurement. [c]n (Hill coefficient) is close to 0.5 for the H115Y mutant. [d]N = A, G, C, T. [e]N = G, C.

structures indicates that the movement of helix αD enables partial ring stacking between His115 and the incoming dGTP (Figure 3C). In addition, the overall conformational change is in part achieved by disruption of extensive hydrophobic interactions at the interface between the two subdomains in the Pol X structure. As shown in Figure 3D, in the free form (light color) Phe116 phenyl ring packs against Phe102 phenyl ring with a pi−pi interaction, while His115 imidazole makes contact with Phe102, Ile113 and Phe116. Upon dGTP binding (heavy color), the phenyl ring of Phe116 rotates away from that of Phe102 (orthogonal instead of parallel now), and His115 moves further away to create a pocket large enough to accommodate the bulky purine ring of dGTP.

Pol X Uses a Unique DNA Binding Mode with Helix αE to Bind DNA. Pol X lacks the two subdomains responsible for DNA binding in Pol β.[50] Nevertheless, Pol X binds DNA almost as tightly as Pol β does on the basis of a previous report,[30] and even more tightly than Pol β (by a factor of 3.5) toward the dhDNA used in this study when examined by ITC under identical conditions (rows 15, 16 of Table 1 and Figure S6B,E (SI)). This dichotomy raises interesting questions about how Pol X binds to DNA. The DNA binding surface of Pol X in the absence of dNTP has been mapped by chemical shift changes previously.[29,30] The structure of the Pol X:MgdGTP:DNA ternary complex in this work provides details of the actual DNA binding site. As shown by the electrostatic representation of the ternary complex (Figure 4A), the major positive patch of the helix αE in Pol X forms a major binding interface with the gapped DNA, while the major negative patch formed by the three active site aspartates accommodates the triphosphate moiety of MgdGTP. The gapped DNA adopts a sharp turn after the template G10, enabling G10 to mispair with the preformed *syn* dGTP. These structural features suggest that

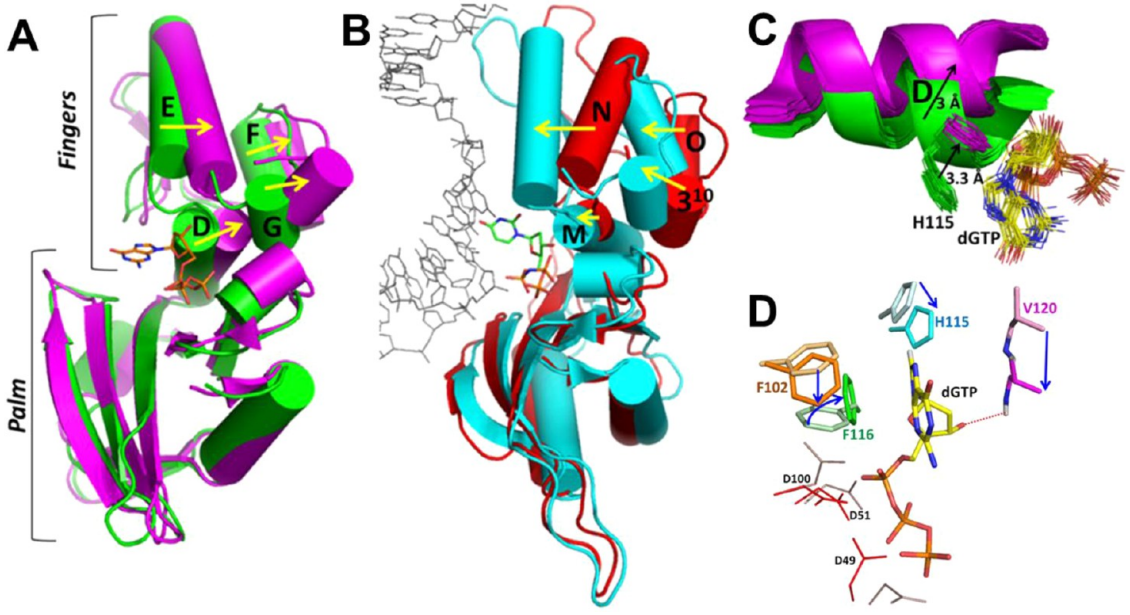


Figure 3. Conformational change induced by MgdGTP binding to free Pol X in the absence of DNA. (A) Free Pol X in green, binary Pol X in magenta, and dGTP in stick (both are the lowest energy NMR structures). The two structures were aligned according to their Palm-subdomains (a.a. 1−105). (B) Conformational change of the Pol X-corresponding subdomains between Pol β:DNA binary complex (red ribbon, PDB code 1BPX) and Pol β:DNA:dUMPNPP ternary complex (cyan ribbon, PDB code 2FMS). The two structures were also aligned on their Palm-subdomains (a.a. 149−261). (C) Overlaid 20 lowest-energy NMR structures showing the conformational transition of helix αD and His115 between free Pol X (green) and binary Pol X (magenta) with MgdGTP. The movement of D-helix was measured from Cα carbons of His115 (3 Å), and the movement of His115 imidazole rings was measured from their Cε1 carbons (3.3 Å). (D) Rearrangement of active site hydrophobic residues and His115 upon MgdGTP binding to Pol X. Residues in the Pol X:MgdGTP binary complex are shown in colored sticks, while the corresponding residues in the free Pol X are shown in the same but lighter color.

dx.doi.org/10.1021/ja4102375 | J. Am. Chem. Soc. 2014, 136, 4927−4937

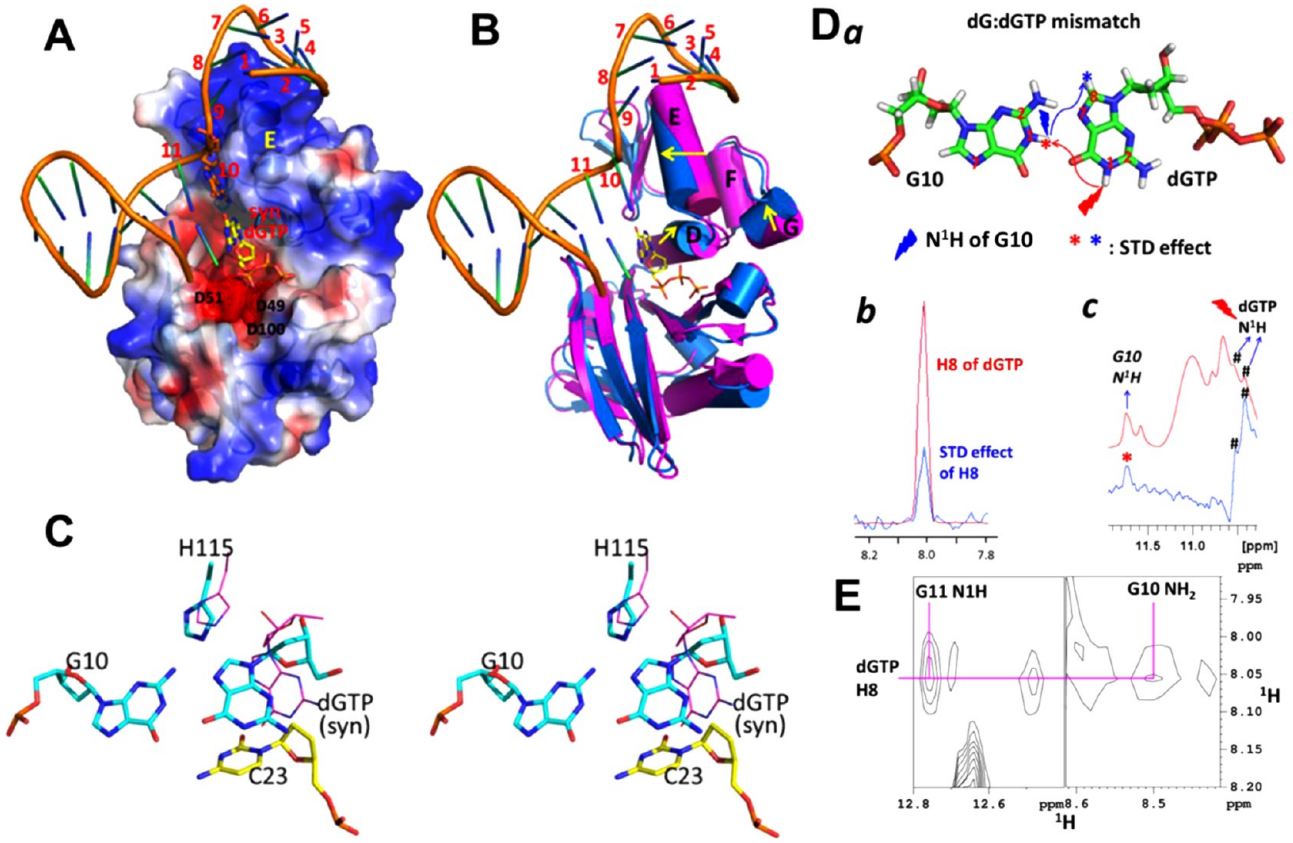


Figure 4. Detailed structure and NMR evidence for the Pol X:MgdGTP:DNA ternary complex. (A) Electrostatic representation for the ternary complex. (B) Structural transition from binary (magenta) to ternary (marine) states. (C) Stereo view for the binary-to-ternary transition at the dG:dGTP mismatch site. For simplicity, the triphosphate moiety of dGTP is omitted. (D) Saturation transfer difference (STD) for the dG:dGTP mismatch. (a) Schematic demonstration of the experiments. The lightning bolt symbol indicates the proton on which a train of saturation pulse was applied, and the asterisk symbol indicates the STD effect on that proton. (b) 1D H8 row vectors from STD 2D $^{1}H,^{13}C$-HSQC spectra of dGTP in the ternary complex. The irradiation was applied to N1H of the template G10 (blue lightning bolt symbol in (a)), while the blank irradiation was applied at 25 ppm. (c) STD spectra with P1331 water suppression scheme for the ternary complex. The bottom trace is the difference spectrum between the control (top spectrum, blank irradiation at 25 ppm) and the irradiation (at 10.45 ppm of $^{15}N,^{13}C$-dGTP N1H proton signal; the doublet signals, labeled by # signs, are due to N−H J-coupling). For irradiation, a train of 50 ms Gaussian shape pulse was applied for 0.8 s. (E) Cross-strand NOEs between H8 of the incoming dGTP and both G10 N1H imino proton and G11 NH_2 amino protons of the bound DNA in a 2D jump-return NOESY spectrum (120 ms mixing time) on a 0.5 mM ternary complex in 8% D_2O at 293 K at 600 MHz. The molar ratio was 1:1.5:10 for $[^{2}H,^{13}C,^{15}N]$-Pol X:dhDNA:dGTP. The buffer contained 10 mM $MgCl_2$, 500 mM KCl, and 50 mM potassium borate, pH 7.5.

the concentrated positive charge on the surface of helix αE contributes significantly to binding of the gapped DNA and helps to direct the formation of the *anti*:*syn* dG:dGTP Hoogsteen base pair (see next section). In addition, this interaction of helix αE with DNA leads to additional conformational change of Pol X as shown in Figure 4B, which indicates that αE moves toward the phosphate backbone of the G1-C9 DNA segment, possibly forming extensive Coulombic interactions between several positively charged residues on the helix (Arg125, 127 and Lys131, 132, 133) and the phosphate backbone of DNA.

The extensive DNA-interacting role of helix αE in Pol X is unique, as fewer such interactions are observed for other members of the X-family DNA polymerases. As shown in Figure S6A (SI), in the corresponding helix of Pol β, Pol λ, and Pol μ, there are only two positively charged residues that interact with the phosphate backbone of DNA. Elimination of two of the five DNA-interacting positively charged residues in the helix αE in R125AR127A or R125QR127Q significantly reduced the DNA affinity of Pol X (rows 17, 18 of Table 1 and Figure S6B−D (SI)). This result supports the importance of helix αE in the unique high DNA affinity of Pol X.

***Syn* dGTP Forms Hoogsteen Base Pair with Template dG in the Ternary Complex.** As shown in Figure 4C, the partial pi−pi stacking in the binary complex is disrupted in the ternary complex, freeing the guanine ring of dGTP to pair with the template dG. Importantly, the motion of the guanine is small and the *syn* conformation is still retained, leading to the Hoogsteen base pairing. As a result of these small motions, the imidazole ring of His115 becomes more (but not fully) coplanar and less ring-stacked with the guanine ring of dGTP. Here we describe some specific spectroscopic evidence for the structure of the ternary complex and the *anti*:*syn* dG:dGTP Hoogsteen base pairing. First, we prepared isotope-labeled DNA with the segment of G10G11G12T13G14 labeled with ^{13}C and ^{15}N (Figure S7 (SI) and SI Materials and Methods). The sequential imino−imino connectivity for the G10−G14 segment of the double hairpin DNA free in solution and in the dG:dGTP and dG:dCTP ternary complexes were assigned (Figure S8 (SI)). Detailed NMR analyses clearly showed no looping out of any nucleotide base, as explained in SI Materials

dx.doi.org/10.1021/ja4102375 | J. Am. Chem. Soc. 2014, 136, 4927−4937

and Methods. The labeled DNA allowed us to assign the NOE between the Nε2-H proton of His115 and the imino N−H proton of G11 as shown in Figure S8D (SI). Next, we prepared specifically labeled Pol X and assigned the methyl groups as shown in Figure S9A (SI). These specific assignments enabled identification of intermolecular NOEs between the methyl groups of Val120 and Ile124 of Pol X and the imino proton of G10 as shown in Figure S9B (SI).

Furthermore, saturation transfer difference (STD) effect was observed for the incoming dGTP C8−H8 cross peak in a 2D $^{1}H,^{13}C$-HSQC when the N1H proton signal of G10 was saturated (Figure 4D-a,b). Similarly, STD effect was observed for the N1H proton of the template G10 when the N1H signal of dGTP was saturated by selective irradiation (Figure 4D-c, and see Figure S9C (SI) for the assignment of dGTP N1H). In addition, a previous report showed that *anti*:*syn* GG mismatch in a DNA displays observable interstrand NOEs with a neighboring base.[51] We indeed observed two cross-strand NOEs between the H8 proton of MgdGTP and $N2H_2$ of G10 and N1H of G11 (Figure 4E).

Consecutive Changes from Free to Binary and Ternary Complexes. Taken together, the results indicate that helices αD/αE in the free form first move outward to accommodate the nucleotide in the binary complex (Figure 3A) and then move back inward to interact with DNA in the ternary complex (Figure 4B). These consecutive changes are illustrated in Figure 5A, which also shows the movement of His115 and its role in the binary and the ternary complexes as described above. Another important point is that the loop between helices αD and αE consists of the Gly118-Pro119 peptide bond (conserved in the X-family polymerase as *cis* Gly-Ser), which was shown to be *cis* in the free Pol X[29,30] and remains *cis* in the MgdGTP binary complex as indicated by the characteristic NOE between the H_α proton of Gly118 and H_α proton of Pro119, and in the ternary complex as evidenced by characteristic ^{13}C chemical shifts.[52] It could play a unique role in providing a sharp turn between helix αD and helix αE, while also providing necessary space for dGTP binding and for conformational transitions between the three forms. Note that while the conformational change in the binary complex enables formation of the H-bond between Val120 backbone N−H and dGTP 3′-OH group as mentioned above (Figure 2C), this H-bond is disrupted by the further conformational change induced by DNA binding as shown in Figure 5A. The chemical shift of Val120 N−H, which moves from 8.87 ppm (free) to 9.42 ppm (binary) as mentioned above, moves back to 8.41 ppm in the ternary complex.

His115 Is Not Needed in the dG:dCTP Ternary Complex with Watson−Crick Pairing. The results above suggest that Pol X catalyzes dG:dGTP incorporation by prebinding MgdGTP in the *syn* form, which then forms Hoogsteen base pairing with the template dG in the ternary complex (*anti*:*syn* for dG:dGTP). While this is a unique and effective mechanism to overcome Watson−Crick base pairing, it prompted us to ask how Pol X catalyzes Watson−Crick dG:dCTP incorporations when the prebinding of dCTP is much weaker as described above. Here we use 1D proton NMR spectra (Figure 5B) to show that the dG:dCTP ternary complex is also readily formed, but it does not depend critically on His115. As shown in the dhDNA assignments in Figure S8 (SI), the imino protons of G10 and T13 resonate at 10.72 and 13.71 ppm, respectively in the free DNA.

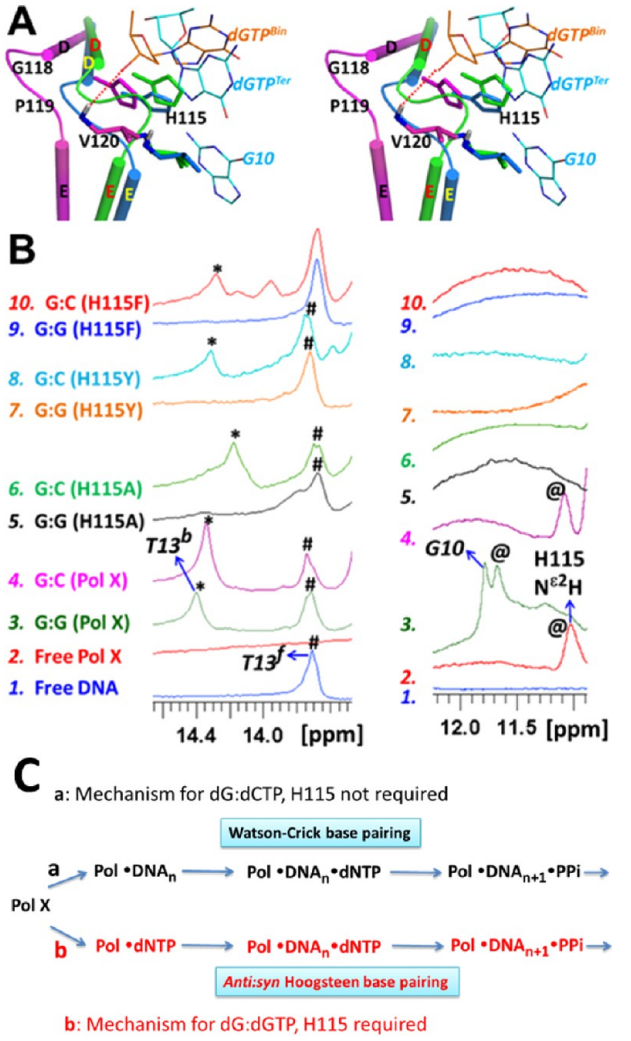


Figure 5. Comparison between dG:dGTP and dG:dCTP. (A) Conformational transitions of helices αD and αE from the free Pol X (green) to the binary complex (magenta) and the ternary complex (marine) as explained in the text. For simplicity, the ribose of template G10 and the triphosphate of dGTP are omitted. "$dGTP^{Bin}$" and "$dGTP^{Ter}$" stand for dGTP in the binary form and the ternary form, respectively. (B) 1D NMR spectra showing chemical shift evidence for dG:dGTP (G:G) or dG:dCTP (G:C) ternary complexes. $T13^f$ and $T13^b$ (marked with "#" and "*", respectively) stand for the N1H resonance of T13 in the free and bound states, respectively. His115 Nε2H proton is marked with "@". All spectra were acquired at 293 K. The molar ratio was 1:1.1:15 for each of the Pol X:DNA:dNTP ternary complex. The buffer contained 10 mM $MgCl_2$, 500 mM KCl, and 50 mM potassium borate, 10 mM DTT-d10, pH 7.5. Note that the distortion of baselines and some peak shapes is the nature of the P1331 pulse sequence (phase modulation as a function of resonance frequency), which has the advantage of reducing saturation transfer effect on labile protons. (C) Proposed model showing the different catalytic pathways of Pol X for dG:dCTP (a) and dG:dGTP (b).

Spectrum 1 in Figure 5B is for the free dhDNA where only the T13 peak is shown for simplicity. Spectrum 2 is for free WT Pol X where the peak at 11.03 ppm was assigned to the His115 Nε2H proton as explained in SI Materials and Methods and Figure S8E (SI). In the dG:dGTP ternary complex (spectrum 3), three distinct downfield shifts occur: the His115 Nε2H peak shifted to 11.67 ppm, the G10 imino proton peak shifted to 11.8 ppm (forming an apparent doublet with the His115 peak),

dx.doi.org/10.1021/ja4102375 | J. Am. Chem. Soc. 2014, 136, 4927−4937

Table 2. Catalytic Efficiencies for Incorporation of dCTP, dGTP, and 7-Deaza-dGTP Opposite Template G by Pol X and H115A Mutant[a]

row	enzyme	reaction	k_{pol} (s^{-1})	mutant/WT	$K_{d,app}$ (μM)	mutant/WT	$k_{pol}/K_{d,app}$ (s^{-1} M^{-1})	fidelity[c]
1	Pol X	dG:dCTP[b]	0.074 ± 0.005		200 ± 30		370	
2	Pol X H115A	dG:dCTP	0.16 ± 0.01	2.2	73 ± 15	0.36	2200	
3	Pol X	dG:dGTP[b]	0.0028 ± 0.0001		21 ± 1		110	3.8
4	Pol X H115A	dG:dGTP	0.0016 ± 0.0001	0.57	690 ± 90	33	2.3	960
5	Pol X	dG:7-deaza-dGTP	0.00019 ± 0.00002		510 ± 130		0.37	
6	Pol X H115A	dG:7-deaza-dGTP	0.00013 ± 0.00002	0.68	930 ± 300	1.8	0.14	

[a]The ± signs stand for standard deviation. [b]Data from Lamarche et al.[27] [c]The fidelity is defined by $[(k_{pol}/K_{d,app})_{dG:dCTP} + (k_{pol}/K_{d,app})_{dG:dGTP}]/(k_{pol}/K_{d,app})_{dG:dGTP}$.

and the T13 imino proton peak shifted to 14.4 ppm (also refer to Figure S9B–D (SI) for these assignments). Importantly, in the dG:dCTP ternary complex (spectrum 4), the apparent doublet at 11.67/11.8 ppm does not show up (His115 is not shifted whereas G10 is shifted further to 12.9 ppm, not shown here because of overlapping with other peaks), but the downfield shift of the T13 imino proton peak still happens. These characteristic shifts allowed us to monitor the formation of a dG:dGTP Hoogsteen base pairing or a dG:dCTP Watson–Crick base pairing in site-specific mutants. As shown in spectra 5–6, mutation of His115 to alanine abolished the formation of the dG:dGTP mismatch but had little effect on the dG:dCTP correct match. This result indicates that His115 is required for the formation of a dG:dGTP mismatch, but to the contrary it plays little role in the dG:dCTP Watson–Crick base pairing.

In addition, the fact that the relative intensities of the two T13 resonances are similar between GG and GC complexes (spectra 3, 4) suggests that both ternary complexes are formed readily under the condition used (enzyme:DNA:MgdNTP = 1:1.5:10), even though MgdCTP binds much more weakly than MgdGTP in the absence of DNA. Importantly, unlike the multiple forms in the MgdCTP binary complex (Figure S4B (SI)), the dG:dCTP ternary complex exists in one single form (Figure S4D (SI)). This result suggests that even though MgdCTP binds to Pol X more weakly than MgdGTP in the absence of DNA, the Watson–Crick base pairing can compensate for the difference. Taken together, the results suggest that Pol X uses different mechanisms for dG:dGTP (His115 and Hoogsteen base pairing) and dG:dCTP (His115 side chain not required) incorporations.

Comparison of Pol X His115 with Pol β Tyr271. It is interesting that His115 of Pol X corresponds to Tyr271 of Pol β. We previously reported that Y271F, Y271A, and Y271S mutations of Pol β led to modest perturbations in its kinetic constants.[53] In light of the important roles of His115 in the functions of Pol X, we examined whether Y271H Pol β will acquire the MgdNTP binding property of Pol X, and whether H115Y Pol X will reduce its MgdNTP binding property. Our NMR (Figure S5B (SI)) and ITC (Figure S5D (SI) and row 7 of Table 1) analyses indicate that Y271H Pol β remains incapable of binding MgdNTP tightly in the absence of DNA. Likewise, for Pol X, H115Y retains the ability to bind MgdGTP tightly and MgdCTP weakly (15-fold difference in K_d) in the absence of DNA (rows 13, 14 of Table 1). These results suggest that the difference between Pol X and Pol β should not be attributed primarily to the difference in the single residue.

H115Y and H115F Mutations Disrupt dG:dGTP but Not dG:dCTP Ternary Complex. Although the H115Y mutation did not perturb the binding affinity of MgdGTP and MgdCTP, it affected the ternary complex. In contrast to WT Pol X, which forms both dG:dGTP and dG:dCTP ternary complexes readily as described above, H115Y is unable to form the dG:dGTP ternary complex (Figure 5B, spectrum 7), while its ability to form the dG:dCTP ternary complex is unperturbed (spectrum 8). Similar results were obtained for the H115F mutant (spectra 9, 10). Although the structural basis for the properties of His115 mutants requires further investigation, the results provide additional support for the conclusion that His115 is the key residue for dG:dGTP non-Watson–Crick incorporation but is little relevant to the mechanism of the Watson–Crick incorporation catalyzed by the same enzyme.

Formation of dG:dGTP Non-Watson–Crick Pairing Is Not Sensitive to DNA Sequence. Previous reports suggested the DNA-sequence dependence of dNTP incorporation by other members of the X family DNA polymerases such as Pol β,[54] Pol λ,[55] and Pol μ.[56] In contrast, our recent kinetic analyses showed that the dG:dGTP incorporation by Pol X is insensitive to the DNA sequence.[27] This difference can be explained by two factors. One is that for dG:dGTP non-Watson–Crick catalysis, binding of dGTP precedes that of DNA. The other reason is that, unlike other polymerases that bind DNA extensively in a hand-holding like mode, Pol X mainly uses E-helix to bind DNA, and the main interaction between Pol X and DNA involves the positively charged arginine/lysine side chains and the backbone phosphates of DNA bases 1–9 (Figure 4A,B and S6A (SI)). To provide further support to this explanation, we examined the effect of changing the bases in this region. As shown in Figure S11 (SI), the apparent doublet for the templating base *G10* imino proton and the His115 Nε2H proton remain intact in the three spectra with different DNA sequences.

Support of the Results of Structural Analyses by Kinetics and 7-Deaza-dGTP. While structural analyses can provide insightful information on the intermediate complexes, the new information from structural analyses should be tested by kinetic analysis, since kinetic analysis can provide information through the transition state. This is particularly true since it has been established previously that the chemical step is the rate-limiting step in Pol X catalysis.[57] For the same reason, the order of substrate binding will not affect the rate k_{pol}. We thus performed single turnover kinetics as described previously.[27,37] As shown in Table 2, mutation of WT to H115A caused relatively minor effects (within a factor of 2) on the k_{pol} and $K_{d,app}$ values for dG:dCTP incorporation (rows 1, 2), or the k_{pol} for dG:dGTP incorporation (rows 3, 4), but it led to 33-fold increase in $K_{d,app}$ for dG:dGTP incorporation (rows 3, 4). These data fully support the conclusions reached from structural analyses. Further support was provided by the use of 7-deaza-dGTP, a dGTP analogue in which the N-atom at the 7-position is replaced by carbon, which thus is incapable of

forming one of the two H-bonds when the Hoogsteen edge of the base is used.[58] As expected, when dGTP was replaced by 7-deaza-dGTP, the H115A/WT ratios returned to within a factor of 2 (rows 5, 6). Furthermore, analyses in the last two columns of Table 2 indicate that the fidelity (dG:dCTP vs dG:dGTP) increases from 3.8 for WT to 960 for H115A, demonstrating the functional role of H115 in dG:dGTP incorporation.

DISCUSSION

Pol X Uses Multiple Unique Strategies for Catalysis. In this work we use NMR to show that Pol X uses active site residue His115 to stabilize and preform a *syn* dGTP conformation in the absence of gapped DNA, as a key mechanism to overcome the Watson−Crick base pairing. This binding of MgdGTP (in the absence of DNA) also induces a conformational change that is different from the subdomain-closing conformational change of enzyme−DNA binary complexes induced by MgdNTP for other DNA polymerases. In the presence of gapped DNA, the *syn* dGTP forms a mismatch with the template dG and forms an *anti*:*syn* dG:dGTP Hoogsteen base pair in the Pol X:DNA:MgdGTP ternary complex, explaining the structural basis of the mismatch incorporation. Although missing the DNA binding subdomain, Pol X uses its helix αE to bind the DNA. While in the absence of DNA Pol X binds MgdCTP weakly and in multiple forms, the Watson−Crick dG:dCTP base pairing occurs readily in the ternary complex. H115A mutation abolished dG:dGTP but not dG:dCTP. The results suggest that Pol X catalyzes dG:dGTP and dG:dCTP incorporations with distinct mechanisms. The dG:dGTP mismatch formation via the prebinding mechanism is likely to occur under physiological conditions because the K_d value for MgdGTP (0.26 ± 0.07 μM) is lower than the average physiological concentration of dGTP in different mammals (5.2 ± 4.5 μM).[59] Some of the points relevant to the mechanism of DNA polymerase catalysis are elaborated below, whereas the implication of the results on the biological functions of Pol X is addressed in SI Discussion.

Prebinding of *syn*-MgdNTP Followed by Mismatch Formation Is Unprecedented in DNA Polymerases. Prebinding of MgdNTP in the absence of DNA has also been reported in a few different DNA polymerases including the template-independent terminal deoxynucleotidyl transferase (TdT),[60] but Pol X is the only one that binds MgdNTP in a *syn* conformation, and also the only one that goes on to form an *anti*:*syn* dG:dGTP mispair for subsequent mismatch incorporation. It is well accepted that high-fidelity replicative DNA polymerases or medium-fidelity gap-filling DNA polymerases cannot bind MgdNTP in the absence of DNA. In the two binary complex structures reported previously for Pol I[61] and Pol β,[62] the bound nucleotide was not located at the correct nucleotide binding site. A notable exception is a recent report that the bacterial DNA polymerase *Thermus thermophilus* HB8 Pol X (tt-Pol X) binds MgdNTP tightly in the absence of gapped DNA. The crystal structure of tt-Pol X:CadGTP binary complex indicates that the bound dGTP exists in both *syn* and *anti* forms in equilibrium, but only the *anti* form exists in the ttPolX:DNA:CaddNTP ternary complex, resulting in an *anti*:*anti* dC:dGTP correct pairing.[63] Thus the mechanistic and functional properties of tt-Pol X and ASFV Pol X are substantially different, even though they share the ability to bind dNTP in the absence of DNA.

The dG:dGTP Hoogsteen Base Pairing for Pol X Is Unique. Hoogsteen base pairing has been shown to occur in the translesion synthesis (TLS) such as Pol ι with 8-oxo-G and incoming dCTP[64] and with ^{6}O-methylG and incoming dTTP[65] (both are *syn*:*anti*), Dpo4 with 5′ thymine of the cyclobutane pyrimidine dimer and incoming ddATP[66] and with 8-oxo-G and incoming CadATP (*syn*:*anti*),[67] and Pol β with 8-oxoG and incoming dAMPCPP (*syn*:*anti*).[68] For correct base pairing, Hoogsteen base pairing has been suggested to offer a basis for the preferred dA:dTTP (*syn*:*anti*) incorporation by Pol ι.[69] For mismatch incorporation, Hoogsteen base pairing has only been reported for Pol β with template dA and incoming 8-oxo-dGTP (*anti*:*syn*),[70] and for dG:dGMP mismatch between the template dG and the dG at the primer terminus in *Bacillus stearothermophilus* DNA polymerase I fragment (BF)[71] and human Pol λ.[72] In comparison with these examples, the dG:dGTP Hoogsteen pairing in this work is unique in that the incoming nucleotide is a natural dGTP and is likely prebound by Pol X in the *syn* conformation before encountering the template dG. It is interesting to note that the *anti*:*syn* geometry for the dG:dGTP pairing determined in this study was also predicted previously by molecular dynamic simulation.[31] On the other hand, even though computational studies predicted that high salt concentration can lead to a conformational change in Pol X,[32] our NMR spectra (Figure S12 (SI)) suggested that the conformational change induced by MgdGTP is very different from the salt effect.

Comparison with the Mechanism of Mismatch Incorporation by Other DNA Polymerases. To understand the molecular basis of fidelity, the structure of a mismatched ternary complex of a polymerase is of paramount importance, since the fidelity is based on the comparison between correct and mismatched incorporations, but the vast majority of research deals with only correct incorporations. Crystallization attempts for mismatched ternary complexes, however, have not been straightforward because of the low affinity of the mismatched dNTP. Except for a few cases described above, earlier reports on the mismatched complex structures have the mismatch occurring between the template and the primer terminus,[71] or between the dNTP and the next template base.[73] For Dpo4, it was shown that dT:dGTP mismatch led to destabilized base stacking.[74] Recently Bebenek et al.[75] used crystal structures of a lower fidelity Pol λ variant to demonstrate a dT:dGMPCPP mismatch with a Watson−Crick geometry consistent with a minor tautomeric or ionized base pair, confirming the existence of such a mechanism that was proposed by Watson and Crick.[76] Like in their study, our NMR structure also demonstrates nascent base pairing of the mismatched complex, but the mechanism used by Pol X to achieve the mismatched complex is clearly different from that of Pol λ.

Role of Enzyme-dNTP Interaction in the DNA Polymerases with Specialized Functions. An interesting feature of the Y family polymerase Dbh[77,78] and Dpo4 is that there are relatively few interactions between the enzyme and DNA, and between the enzyme and the incoming dNTP. This appears not to be the case for the ASFV Pol X. As described above, Pol X binds dGTP with extensive interactions. On the other hand, there are increasing examples for the involvement of specific enzymatic residues in binding the incoming dNTP and/or the template base in a specific conformation in order to achieve the specialized catalytic function of the polymerase. The most dramatic case is Rev1, which achieves its high specificity toward dG:dCTP incorporation by using protein groups to direct both the incoming dCTP and the template G

dx.doi.org/10.1021/ja4102375 | J. Am. Chem. Soc. 2014, 136, 4927−4937

evicted from DNA.[20] While Rev1 uses Arg324 side chain as a template to pair up with the incoming dCTP for its dCTP preference, Pol X selects for *syn*-dGTP via His115 in the absence of DNA.

Although only Rev1 and Pol X are known to use specific active site residues to select for preferred dNTP, several other polymerases are known to use active site residues to orient the template base in a specific conformation. For example, Pol ι binds the template 8-oxo-G in a *syn* conformation to form a correct 8-oxoG:dCTP Hoogsteen pair for an error-free replication,[64] and Pol η accommodates the *cis-syn* cyclobutane pyrimidine dimer with the reinforced B-form DNA conformation via complementary DNA binding surface as well as hydrophobic interaction.[79] Also, Dpo4 bypasses the 2-aminofluorene lesion via error-free and error-prone mechanisms; the latter was promoted by interactions between the enzyme and the bulky lesion.[80]

Conformational Change Occurs upon MgdNTP Binding in the Absence of DNA. Most high-fidelity polymerases undergo the MgdNTP-induced subdomain-closing conformational change upon binding of MgdNTP to the polymerase:DNA binary complex. However, low fidelity Y-family DNA polymerase DinB has been shown to pre-exist in a closed conformation in the free form,[78] while Pol λ, an X family DNA polymerase, shows no significant conformational change on going from the E:DNA binary form to the E:DNA:MgdNTP ternary complex for a match[81] or a mismatch[75] formation. These studies raised the possibility that the open-to-closed conformational change may not be required for error-prone polymerases, or that the lack of it could be a cause of their low fidelity. For the Pol X:MgdGTP binary complex, our results indicate a novel mechanistic and structural feature in that MgdGTP binding in the absence of DNA induces a substantial conformational change that is of a subdomain-opening mode.

Overview of the Catalytic Mechanisms of Pol X. Here we propose an overall picture for the catalytic mechanism of Pol X, based on the results of this work and previous studies.[25,27,30,37] Pol X appears to use two different mechanisms to catalyze dNTP incorporations: the canonical mechanism for dG:dCTP, which does not involve His115, and the dNTP prebinding mechanism for dG:dGTP mediated by His115, as illustrated by the model in Figure 5C. The incorporation of the other base pairs (matched and mismatched) likely can use one or both mechanisms, depending on the nature of dNTP (different affinity for prebinding), and the physiological concentrations of both dNTP and DNA.

CONCLUSION

We present the first solution structural view of DNA polymerase catalysis, a unique DNA binding mode, and a novel mechanism for non-Watson−Crick incorporation by a DNA polymerase. Since the half-sized DNA polymerase X likely represents the minimal version of an evolutionarily conserved DNA polymerase core, the novel mechanistic and structural properties of Pol X discovered in this study shall provide valuable information for researchers in the field of DNA polymerases. In particular, it will be very interesting to examine whether other low-fidelity DNA polymerases also use the strategy of prebinding specific dNTP to choose non-Watson−Crick from Watson−Crick base pairing. This question is particularly interesting in light of the newly published review[82] suggesting a common "nucleotide gateway" for some families of DNA polymerases in the presence of DNA. Furthermore, the structures of the binary and ternary complexes of Pol X may also provide a basis for drug discovery against the deadly Africa swine fever virus that has caused devastating global economic loss in the recent years.[21−23] In particular, our results suggest that Pol X can potentially be inhibited by nucleotide analogues before binding DNA. Most significantly, this work has addressed important questions fundamental to enzyme catalysis.

ASSOCIATED CONTENT

Supporting Information

SI Materials and Methods; SI Discussion; Tables S1−S5; Figures S1−S10. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

mdtsai@gate.sinica.edu.tw

Present Address

[○]C-PCS Chemistry Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States.

Author Contributions

[▽]W.-J. Wu and M.-I. Su contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Drs. A. Bax, N. Fitzkee, and J. Lorieau for useful suggestions on RDC measurements, Dr. M. Zweckstetter for useful suggestions on PALES, Dr. M. Summers for useful discussion, Dr. W. Bermel for helpful suggestions on pulse sequence programming, Dr. P. Güntert for help in using CYANA 3.0, Dr. J. Feigon for suggestion on DNA synthesis, Dr. P. Balbo for critical reading of the manuscript, Drs. L. E. Kay and V. Tugarinov for useful suggestions on the HMCMCGCACB experiment, and Drs. A. Bonvin and M. van Dijk for useful suggestions on HADDOCK. The WeNMR project (European FP7 e-Infrastructure Grant, Contract No. 261572, www.wenmr.eu) is acknowledged for the Haddock computation. We thank Dr. M. L. Salas for histag-Pol X construct. This work was supported by funding from Academia Sinica and National Science Council. NMR data were collected at the High-Field NMR Center supported by the National Research Program for Genomic Medicine, NRPB and NCFPB of Taiwan.

REFERENCES

(1) Steitz, T. A. *J. Biol. Chem.* **1999**, *274*, 17395.
(2) Hubscher, U.; Maga, G.; Spadari, S. *Annu. Rev. Biochem.* **2002**, *71*, 133.
(3) Joyce, C. M.; Benkovic, S. J. *Biochemistry* **2004**, *43*, 14317.
(4) Beard, W. A.; Wilson, S. H. *Chem. Rev.* **2006**, *106*, 361.
(5) Johnson, K. A. *Biochim. Biophys. Acta* **2010**, *1804*, 1041.
(6) Batra, V. K.; Perera, L.; Lin, P.; Shock, D. D.; Beard, W. A.; Pedersen, L. C.; Pedersen, L. G.; Wilson, S. H. *J. Am. Chem. Soc.* **2013**, *135*, 8078.
(7) Freudenthal, B. D.; Beard, W. A.; Shock, D. D.; Wilson, S. H. *Cell* **2013**, *154*, 157.
(8) Moon, A. F.; Garcia-Diaz, M.; Batra, V. K.; Beard, W. A.; Bebenek, K.; Kunkel, T. A.; Wilson, S. H.; Pedersen, L. C. *DNA Repair* **2007**, *6*, 1709.
(9) Showalter, A. K.; Lamarche, B. J.; Bakhtina, M.; Su, M.-I.; Tang, K.-H.; Tsai, M.-D. *Chem. Rev.* **2006**, *106*, 340.

(10) Ohmori, H.; Friedberg, E. C.; Fuchs, R. P. P.; Goodman, M. F.; Hanaoka, F.; Hinkle, D.; Kunkel, T. A.; Lawrence, C. W.; Livneh, Z.; Nohmi, T.; Prakash, L.; Prakash, S.; Todo, T.; Walker, G. C.; Wang, Z.; Woodgate, R. *Mol. Cell* **2001**, *8*, 7.
(11) Yang, W. *FEBS Lett.* **2005**, *579*, 868.
(12) Yamtich, J.; Sweasy, J. B. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 1136.
(13) Broyde, S.; Patel, D. J. *Nature* **2010**, *465*, 1023.
(14) Rothwell, P. J.; Waksman, G. *Adv. Protein Chem.* **2005**, *71*, 401.
(15) McCulloch, S. D.; Kunkel, T. A. *Cell Res.* **2008**, *18*, 148.
(16) Schneider, S.; Schorr, S.; Carell, T. *Curr. Opin. Struct. Biol.* **2009**, *19*, 87.
(17) Hubscher, U.; Maga, G. *Curr. Opin. Chem. Biol.* **2011**, *15*, 627.
(18) Xia, S.; Wang, J.; Konigsberg, W. H. *J. Am. Chem. Soc.* **2012**, *135*, 193.
(19) Banerjee, S.; Brown, K. L.; Egli, M.; Stone, M. P. *J. Am. Chem. Soc.* **2011**, *133*, 12556.
(20) Nair, D. T.; Johnson, R. E.; Prakash, L.; Prakash, S.; Aggarwal, A. K. *Science* **2005**, *309*, 2219.
(21) Costard, S.; Wieland, B.; de Glanville, W.; Jori, F.; Rowlands, R.; Vosloo, W.; Roger, F.; Pfeiffer, D. U.; Dixon, L. K. *Philos. Trans. R. Soc., B* **2009**, *364*, 2683.
(22) Callaway, E. *Nature* **2012**, *488*, 565.
(23) Alonso, C. *Virus Res.* **2013**, *173*, 1.
(24) Oliveros, M.; Yáñez, R. J.; Salas, M. a. L.; Salas, J.; Viñuela, E.; Blanco, L. *J. Biol. Chem.* **1997**, *272*, 30899.
(25) Showalter, A. K.; Tsai, M.-D. *J. Am. Chem. Soc.* **2001**, *123*, 1776.
(26) García-Escudero, R.; García-Díaz, M.; Salas, M. a. L.; Blanco, L.; Salas, J. *J. Mol. Biol.* **2003**, *326*, 1403.
(27) Lamarche, B. J.; Kumar, S.; Tsai, M.-D. *Biochemistry* **2006**, *45*, 14826.
(28) Maga, G.; Villani, G.; Crespan, E.; Wimmer, U.; Ferrari, E.; Bertocci, B.; Hubscher, U. *Nature* **2007**, *447*, 606.
(29) Maciejewski, M. W.; Shin, R.; Pan, B.; Marintchev, A.; Denninger, A.; Mullen, M. A.; Chen, K.; Gryk, M. R.; Mullen, G. P. *Nat. Struct. Mol. Biol.* **2001**, *8*, 936.
(30) Showalter, A. K.; Byeon, I. J.; Su, M. I.; Tsai, M. D. *Nat. Struct. Biol.* **2001**, *8*, 942.
(31) Sampoli Benítez, B. A.; Arora, K.; Balistreri, L.; Schlick, T. *J. Mol. Biol.* **2008**, *384*, 1086.
(32) Sampoli Benítez, B. A.; Arora, K.; Schlick, T. *Biophys. J.* **2006**, *90*, 42.
(33) Voehler, M. W.; Eoff, R. L.; McDonald, W. H.; Guengerich, F. P.; Stone, M. P. *J. Biol. Chem.* **2009**, *284*, 18434.
(34) Jezewska, M. J.; Marcinowicz, A.; Lucius, A. L.; Bujalowski, W. *J. Mol. Biol.* **2006**, *356*, 121.
(35) Jezewska, M. J.; Bujalowski, P. J.; Bujalowski, W. *Biochemistry* **2007**, *46*, 12909.
(36) Tang, K.-H.; Niebuhr, M.; Aulabaugh, A.; Tsai, M.-D. *Nucleic Acids Res.* **2008**, *36*, 849.
(37) Kumar, S.; Bakhtina, M.; Tsai, M.-D. *Biochemistry* **2008**, *47*, 7875.
(38) Keller, R. *The Computer Aided Resonance Assignment Tutorial*; CANTINA Verlag: Goldau, Switzerland, 2004.
(39) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. *J. Am. Chem. Soc.* **2003**, *125*, 1731.
(40) van Dijk, M.; van Dijk, A. D. J.; Hsu, V.; Boelens, R.; Bonvin, A. M. J. J. *Nucleic Acids Res.* **2006**, *34*, 3317.
(41) Güntert, P.; Mumenthaler, C.; Wüthrich, K. *J. Mol. Biol.* **1997**, *273*, 283.
(42) van Dijk, A. D.; Bonvin, A. M. J. J. *Bioinformatics* **2006**, *22*, 2340.
(43) Arndt, J. W.; Gong, W.; Zhong, X.; Showalter, A. K.; Liu, J.; Dunlap, C. A.; Lin, Z.; Paxson, C.; Tsai, M.-D.; Chan, M. K. *Biochemistry* **2001**, *40*, 5368.
(44) Batra, V. K.; Beard, W. A.; Shock, D. D.; Krahn, J. M.; Pedersen, L. C.; Wilson, S. H. *Structure* **2006**, *14*, 757.
(45) Steitz, T. A.; Smerdon, S. J.; Jäger, J.; Joyce, C. M. *Science* **1994**, *266*, 2022.
(46) Beard, W. A.; Wilson, S. H. *Mutat. Res.* **2000**, *460*, 231.
(47) Kirby, T. W.; DeRose, E. F.; Beard, W. A.; Wilson, S. H.; London, R. E. *Biochemistry* **2005**, *44*, 15230.
(48) Patel, D. J.; Kozlowski, S. A.; Nordheim, A.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 1413.
(49) Nikolova, E. N.; Kim, E.; Wise, A. A.; O'Brien, P. J.; Andricioaei, I.; Al-Hashimi, H. M. *Nature* **2011**, *470*, 498.
(50) Casas-Finet, J. R.; Kumar, A.; Karpel, R. L.; Wilson, S. H. *Biochemistry* **1992**, *31*, 10272.
(51) Skelly, J. V.; Edwards, K. J.; Jenkins, T. C.; Neidle, S. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 804.
(52) Shen, Y.; Bax, A. *J. Biomol. NMR* **2010**, *46*, 199.
(53) Kraynov, V. S.; Werneburg, B. G.; Zhong, X.; Lee, H.; Ahn, J.; Tsai, M. D. *Biochem. J.* **1997**, *323*, 103.
(54) Beard, W. A.; Shock, D. D.; Wilson, S. H. *J. Biol. Chem.* **2004**, *279*, 31921.
(55) Duym, W. W.; Fiala, K. A.; Bhatt, N.; Suo, Z. *J. Biol. Chem.* **2006**, *281*, 35649.
(56) Zhang, Y.; Wu, X.; Yuan, F.; Xie, Z.; Wang, Z. *Mol. Cell. Biol.* **2001**, *21*, 7995.
(57) Bakhtina, M.; Roettger, M. P.; Kumar, S.; Tsai, M.-D. *Biochemistry* **2007**, *46*, 5463.
(58) Kretulskie, A. M.; Spratt, T. E. *Biochemistry* **2006**, *45*, 3740.
(59) Traut, T. W. *Mol. Cell. Biochem.* **1994**, *140*, 1.
(60) Delarue, M.; Boule, J. B.; Lescar, J.; Expert-Bezancon, N.; Jourdan, N.; Sukumar, N.; Rougeon, F.; Papanicolaou, C. *EMBO J.* **2002**, *21*, 427.
(61) Sloan, D. L.; Loeb, L. A.; Mildvan, A. S. *J. Biol. Chem.* **1975**, *250*, 8913.
(62) Sawaya, M.; Pelletier, H.; Kumar, A.; Wilson, S.; Kraut, J. *Science* **1994**, *264*, 1930.
(63) Nakane, S.; Ishikawa, H.; Nakagawa, N.; Kuramitsu, S.; Masui, R. *J. Mol. Biol.* **2012**, *417*, 179.
(64) Kirouac, K. N.; Ling, H. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3210.
(65) Pence, M. G.; Choi, J.-Y.; Egli, M.; Guengerich, F. P. *J. Biol. Chem.* **2010**, *285*, 40666.
(66) Ling, H.; Boudsocq, F.; Plosky, B. S.; Woodgate, R.; Yang, W. *Nature* **2003**, *424*, 1083.
(67) Di Noia, J. M.; Neuberger, M. S. *Annu. Rev. Biochem.* **2007**, *76*, 1.
(68) Batra, V. K.; Shock, D. D.; Beard, W. A.; McKenna, C. E.; Wilson, S. H. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 113.
(69) Nair, D. T.; Johnson, R. E.; Prakash, S.; Prakash, L.; Aggarwal, A. K. *Nature* **2004**, *430*, 377.
(70) Batra, V. K.; Beard, W. A.; Hou, E. W.; Pedersen, L. C.; Prasad, R.; Wilson, S. H. *Nat. Struct. Mol. Biol.* **2010**, *17*, 889.
(71) Johnson, S. J.; Beese, L. S. *Cell* **2004**, *116*, 803.
(72) Picher, A. J.; García-Díaz, M.; Bebenek, K.; Pedersen, L. C.; Kunkel, T. A.; Blanco, L. *Nucleic Acids Res.* **2006**, *34*, 3259.
(73) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. *Cell* **2001**, *107*, 91.
(74) Vaisman, A.; Ling, H.; Woodgate, R.; Yang, W. *EMBO J.* **2005**, *24*, 2957.
(75) Bebenek, K.; Pedersen, L. C.; Kunkel, T. A. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 1862.
(76) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964.
(77) Silvian, L. F.; Toth, E. A.; Pham, P.; Goodman, M. F.; Ellenberger, T. *Nat. Struct. Biol.* **2001**, *8*, 984.
(78) Zhou, B. L.; Pata, J. D.; Steitz, T. A. *Mol. Cell* **2001**, *8*, 427.
(79) Biertümpfel, C.; Zhao, Y.; Kondo, Y.; Ramón-Maiques, S.; Gregory, M.; Lee, J. Y.; Masutani, C.; Lehmann, A. R.; Hanaoka, F.; Yang, W. *Nature* **2010**, *465*, 1044.
(80) Rechkoblit, O.; Kolbanovskiy, A.; Malinina, L.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Nat. Struct. Mol. Biol.* **2010**, *17*, 379.
(81) Garcia-Diaz, M.; Bebenek, K.; Krahn, J. M.; Kunkel, T. A.; Pedersen, L. C. *Nat. Struct. Mol. Biol.* **2005**, *12*, 97.
(82) Wu, S.; Beard, W. A.; Pedersen, L. G.; Wilson, S. H. *Chem. Rev.* **2014**, *114*, 2759.