Biochemistry

 Volume 42, Number 13 April 8, 2003

Articles

Elucidation of the $\epsilon-\theta$ Subunit Interface of *Escherichia coli* DNA Polymerase III by NMR Spectroscopy†

Eugene F. DeRose,‡ Thomas Darden,‡ Scott Harvey,§ Scott Gabel,‡ Fred W. Perrino,§ Roel M. Schaaper,|| and Robert E. London*,‡

Laboratories of Structural Biology and Molecular Genetics, National Institute of Environmental Health Sciences, Box 12233, Research Triangle Park, North Carolina 27709, and Department of Biochemistry, Wake Forest University Medical Center, Winston-Salem, North Carolina 27157

Received August 21, 2002

ABSTRACT: The DNA polymerase III holoenzyme (HE) is the primary replicative polymerase of *Escherichia coli*. The epsilon (ϵ) subunit of HE provides the 3′→5′ exonucleolytic proofreading activity for this complex. ϵ consists of two domains: an N-terminal domain containing the proofreading exonuclease activity (residues 1−186) and a C-terminal domain required for binding to the polymerase (α) subunit (residues 187−243). In addition to α, ϵ also binds the small (8 kDa) theta (θ) subunit. The function of θ is unknown, although it has been hypothesized to enhance the 3′→5′ exonucleolytic proofreading activity of ϵ. Using NMR analysis and molecular modeling, we have previously reported a structural model of ϵ186, the N-terminal catalytic domain of ϵ [DeRose et al. (2002) *Biochemistry 41*, 94]. Here, we have performed 3D triple resonance NMR experiments to assign the backbone and C^{β} resonances of [U-^{2}H,^{13}C,^{15}N] methyl protonated ϵ186 in complex with unlabeled θ. A structural comparison of the ϵ186$-\theta$ complex with free ϵ186 revealed no major changes in secondary structure, implying that the overall structure is not significantly perturbed in the complex. Amide chemical shift comparisons between bound and unbound ϵ186 revealed a potential binding surface on ϵ for interaction with θ involving structural elements near the ϵ catalytic site. The most significant shifts observed for the ϵ186 amide resonances are localized to helix α1 and β-strands 2 and 3 and to the region near the beginning of α-helix 7. Additionally, a small stretch of residues (K158−L161), which previously had not been assigned in uncomplexed ϵ186, is predicted to adopt β-strand secondary structure in the ϵ186$-\theta$ complex and may be significant for interaction with θ. The amide shift pattern was confirmed by the shifts of aliphatic methyl protons, for which the larger shifts generally were concentrated in the same regions of the protein. These chemical shift mapping results also suggest an explanation for how the unstable *dnaQ49* mutator phenotype of ϵ may be stabilized by binding θ.

The mechanisms by which organisms accurately replicate their DNA are of considerable interest. The best characterized replisome is DNA polymerase III holoenzyme (HE) of the bacterium *Escherichia coli*, which has been studied in significant detail both genetically and biochemically. The holoenzyme, consisting of 17 proteins (10 distinct), contains two core polymerases, responsible for simultaneously replicating the leading and lagging strands (*1*, *2*). Each core consists of three tightly bound subunits, the 130 kDa α subunit (the polymerase), the 27.5 kDa ϵ subunit (a proofreading 3′→5′ exonuclease), and the small (8 kDa) θ subunit, whose function is as yet undetermined. These three subunits

† This work was supported in part by the intramural NIEHS research program (EFD, TD, SAG, RMS, REL) and by NIH Grant CA75350 (FWP).

* To whom correspondence should be addressed. E-mail: london@niehs.nih.gov.

‡ Laboratory of Structural Biology, NIEHS.

§ Department of Biochemistry, Wake Forest University Medical Center.

|| Laboratory of Molecular Genetics, NIEHS.

10.1021/bi0205451 CCC: $25.00
Published on Web 03/13/2003

are bound together in the linear order α-ϵ-θ. Although θ-deficient mutants appear to be phenotypically similar to the wild-type *E. coli* (*3*), effects of θ on the mutation rate can be unmasked under extreme conditions or in organisms in which the stability of ϵ has been reduced by mutations (*4*). Under such conditions, it appears that θ provides stabilization to ϵ, which can significantly improve the accuracy of replicated DNA, presumably via enhancement of the function of the proofreading subunit.

The fidelity of the replication process is determined primarily by (i) the polymerase, which preferentially selects correct nucleotides over incorrect nucleotides during the base insertion step, and (ii) the ϵ proofreading subunit, which can remove polymerase misinsertion errors through its 3′→5′ exonuclease activity. The estimated contributions to overall fidelity are 10^4-10^5 for the base insertion step (misinsertion rate 10^{-4} to 10^{-5}), and 10^2-10^3 for the proofreading step, producing a final misincorporation rate of about 10^{-7} (*5*). Mechanistically, the precise interaction between the polymerase and the exonuclease is of particular interest, as these two activities are located at distinct sites, requiring shuttling of the 3′-terminal base from one site to the other, which may be a rate-limiting factor in fidelity.

The ϵ proofreading subunit has been shown to consist of two domains: an N-terminal domain containing the exonuclease activity (residues 1−186), and a C-terminal domain required for binding to the α (polymerase) subunit (*6*, *7*). We have previously studied the exonuclease domain of ϵ (ϵ186) using a combination of NMR[1] spectroscopy and molecular modeling in order to determine the solution structure of the enzyme (*8*). Using this approach, it was demonstrated that ϵ186 is structurally homologous to the proofreading domain of the Klenow fragment, bacteriophage RB69 (or T4) DNA polymerase, and *E. coli* exonuclease I, despite very limited sequence homology. Recently, an X-ray structure of ϵ186 has been reported which appears to closely resemble the NMR model structure (*9*). In addition to retaining catalytic activity, ϵ186 retains the ability to bind to θ (*6*, *7*). We have previously reported that the binding of ϵ186 to θ leads to significant changes in the $^1H-^{15}N$ HSQC spectrum of θ (*10*). Keniry et al. also noted that preliminary NMR data on θ in the presence of ϵ186 indicated that a flexible loop located between Leu20 and Asn32 of θ is immobilized by the interaction with ϵ186 (*11*). In the present study, we report on the NMR of the ϵ186−θ complex using triply labeled [U-^{2}H,^{13}C,^{15}N] methyl protonated ϵ186 with unlabeled θ. The studies indicate that the overall structure of ϵ186 does not change dramatically upon formation of the complex. However, chemical shift mapping clearly defines a region of ϵ that is involved in complex formation. These results provide insight into the structural basis of the stabilization of ϵ by θ and the basis for the possible modulation of the proofreading function by θ.

EXPERIMENTAL PROCEDURES

The ϵ186−θ Protein Complex. [U-^{2}H,^{13}C,^{15}N] methyl protonated (Val, Leu, and Ile-δ1 methyl) ϵ186 was prepared as described previously (*6*, *8*) with one exception. Unlabeled L-tyrosine (60 mg/L) and L-phenylalanine (60 mg/L) were added in addition to the [3,3-$^2H_2$,U-$^{13}C_4$]-α-ketobutyrate (80 mg/L, a metabolic precursor to isoleucine) and [3-^{2}H,U-$^{13}C_5$]-α-ketoisovalerate (80 mg/L, a metabolic precursor to leucine and valine) when the cultures reached $A_{595} = 0.4$. This labeling approach introduces the possibility of obtaining methyl−aromatic NOE information, although no aromatic resonances in the ^{13}C dimension are available (*12*). Unlabeled θ was prepared as described by Li et al. (*10*). To prepare the ϵ−θ complex, the labeled ϵ186 and unlabeled θ were concentrated to 0.125−0.25 mM in 0.9 M Tris, pH 7.0, buffer using a Centricon concentrator with a 10 kDa cutoff (ϵ186) or a 3 kDa cutoff (θ) (Millipore, Inc., Bedford, MA). The buffer also contained 100 μM 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF) to retard proteolysis resulting from any contaminating proteases which might be present. Protein concentrations were then determined spectrophotometrically [$\epsilon_{280}(\epsilon) = 7920$; $\epsilon_{280}(\theta) = 8250$]. The solution containing θ was slowly added to the solution containing ϵ so that a 1:1.05 ϵ:θ ratio was achieved. The samples were then further concentrated to a 600 μL volume for NMR studies, using the NMR buffer 0.2 M Tris-d_{11} (pH 7.0) in 90%/10% H_2O/D_2O, 2.0 mM $MgSO_4$, 1 mM EDTA, 2 mM D,L-1,4-dithiothreitol-d_{10} (Isotec, Inc., Miamisburg, OH), 2 mM sodium azide, and 100 μM AEBSF. The lower Tris concentration was found to be sufficient for studies of the more stable ϵ−θ complex, while ϵ186 by itself required ∼0.9 M Tris in order to exhibit sufficient stability at high concentrations for long-term NMR studies.

NMR Studies. The sequential backbone and C^β resonance assignments of ϵ186 were reported previously by DeRose et al. (*8*). NMR experiments on ϵ186 and ϵ186−θ were performed using a Varian (Palo Alto, CA) 600 MHz UNITY INOVA spectrometer equipped with a Varian 5 mm ^{1}H-{^{13}C,^{15}N} triple resonance probe, with actively shielded *z*-axis gradients and variable temperature capability. The spectra of ϵ186 and ϵ186−θ were collected at 20 and 30 °C, respectively, since the greater stability of the ϵ186−θ complex made it possible to work at higher temperature and greatly improved the sensitivity of the 3D NMR experiments. To quantify changes in backbone amide and side chain methyl chemical shifts, of ϵ186, upon binding θ, ^{15}N and ^{13}C HSQC spectra of the complex were also obtained at 20 °C in 0.9 M Tris. The sequential backbone and C^β resonance assignments of ϵ186−θ were established by the combined analysis of CT-HNCA (*13*), CT-HN(CO)CA, HN(CA)CB, and HN(COCA)CB (*14*) spectra. Backbone carbonyl resonances were assigned from an HNCO experiment (*15*). TROSY-based versions of these experiments were used to improve sensitivity and resolution (*16*, *17*). An ^{15}N-edited NOESY experiment was run to assign aromatic proton chemical shifts of ϵ186 (*18*). A modified version of the 3D CT-^{13}C,^{13}C-edited NOESY experiment was conducted to assign methyl−aromatic NOEs of ϵ186 (*12*, *19*). Leucine, valine, and isoleucine δ-methyl resonances, of ϵ186−θ and free ϵ186, were assigned using the H(CCO)NH-TOCSY and the (H)C(CO)NH-TOCSY experiments as described previously (*8*). The NMR data were processed using NMRPipe (*20*), and resonances were assigned using NMRView (*21*). TROSY-based HNCA, HN(CO)CA, HN(CA)CB, and HN(COCA)CB experiments were used as described by Yamazaki et al. (*13*, *14*) and Yang and Kay (*17*), with pulse

[1] Abbreviations: CSI, chemical shift indexing; NMR, nuclear magnetic resonance; TROSY, transverse relaxation-optimized spectroscopy; Tris-d_{11}, perdeuterated tris(hydroxymethyl)aminomethane.

sequences obtained from the Kay group. In the HNCA and HN(CO)CA experiments 64 × 32 × 512 complex points were acquired, with acquisitions times of 16.0, 17.8, and 64.0 ms in t_1, t_2, and t_3, respectively. In the HN(CA)CB and HN(COCA)CB experiments 64 × 32 × 512 complex points were acquired, with acquisition times of 7.0, 17.8, and 64.0 ms in t_1, t_2, and t_3, respectively. These experiments were acquired with 32 scans per increment, and a 1.0 s delay between scans, resulting in total acquisition time of approximately 82 h. The HNCO experiment was acquired using Varian's ghn_co pulse sequence, with 64 × 32 × 512 complex points and acquisition times of 21.2, 17.8, and 64.0 ms in t_1, t_2, and t_3, respectively. Sixteen scans per increment were acquired with a 1.0 s delay between scans, resulting in a total acquisition time of approximately 41 h. The ^{15}N-edited NOESY experiment on free ϵ186 was carried out using Varian's gnoesyNhsqc experiment, with 128 × 32 × 512 complex points and acquisition times of 16.0, 17.8, and 64.0 ms in t_1, t_2, and t_3, respectively. Sixteen scans per increment were acquired with a 1.0 s delay between scans and a mixing time of 150 ms, resulting in a total acquisition time of approximately 92 h. A 2D methyl−aromatic NOESY spectrum of free ϵ186 was obtained using a modified version of the 3D ^{13}C,^{13}C-edited NOESY experiment (*12*, *19*), with 74 × 512 complex points and acquisition times of 24.7 and 64 ms in t_1 and t_2, respectively. Five hundred twelve scans per increment were acquired, with a 1.0 s delay between scans and a mixing time of 175 ms, resulting in a total acquisition time of approximately 27 h. The triple resonance experiments were processed using NMRPipe as described in ref *14*, with apodization functions of 90° shifted sine-bell squared in t_3 and 90° shifted sine-bell in t_1 and t_2. The ^{15}N-edited NOESY spectrum of ϵ186 was processed using a polynomial filter to suppress the solvent signal in the acquisition dimension, followed by apodization with an 90° shifted sine-bell squared function, and zero filled to 1024 points before Fourier transformation. The downfield half of the spectrum was extracted in the acquisition dimension. The data were apodized with 90° shifted sine-bell functions in t_1 and t_2. The data were zero filled to 256 points and extended via mirror image linear prediction before Fourier transformation in t_1 and t_2, respectively. Secondary structure predictions were made from an analysis of the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^{13}C'$ chemical shifts using the CSI program of Wishart and Sykes [*22*; Protein Engineering Network Centres of Excellence (PENCE)/Medical Research Council of Canada (MRC) Joint Software Centre].

RESULTS

The objective of this study was to compare the secondary structure of $\epsilon186-\theta$ with free ϵ186 determined previously (*8*) and to map chemical shift changes upon binding θ. For this purpose, we reassigned the backbone and $^{13}C^{\beta}$ resonances of ϵ186 in the complex. One hundred fifty-seven amide resonances (87%), 170 $^{13}C^{\alpha}$ resonances (91%), 157 $^{13}C^{\beta}$ resonances (91%), and 139 $^{13}C'$ resonances (75%) were assigned from the HNCA, HN(CO)CA, HN(CA)CB, HN(COCA)CB, and HNCO spectra of the complex. Secondary structure predictions for $\epsilon186-\theta$ derived from a CSI analysis of the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^{13}C'$ chemical shifts (*22*) or derived using the program TALOS (*23*) are given in Figure 1. For comparison, the CSI and TALOS secondary structure predictions obtained previously in our study of uncomplexed ϵ186 and the secondary structures obtained from a crystallographic analysis of ϵ186 in the presence of Mn^{2+} ions and thymidine 5′-monophosphate (*9*) are also included in Figure 1. In general, the secondary structure predictions derived from the NMR studies of ϵ186 and the $\epsilon186-\theta$ complex are very similar to each other and similar to the recently presented crystal structure. We note that a few of the structural elements present in the crystal structure are not predicted by the CSI analysis but do show up in the TALOS and NOE analyses (data not shown) (Figure 1; also, see Figure 2 of ref *8*). In particular, helix α2 identified in the crystal structure was not identified by the CSI analysis but was identified by TALOS and NOE analyses for ϵ186 and TALOS analysis for $\epsilon186-\theta$ in the present study. For consistency with the nomenclature used in the crystallographic study (*9*), we have now renumbered the helices to include α2. As in the previous study of ϵ186, the CSI analysis also failed to identify helix α1 and strand β5, while these elements were identified by the TALOS analysis (Figure 1). All of these elements were present in the modeled structure. Additionally, the 3_{10} helix structure for residues Lys120−Phe124 identified in the crystal was absent in the modeled structure, although TALOS (*23*) assigned Thr121 to a helical geometry. Finally, we note the greater extension of the β3 and α3 elements in the crystal structure as compared with the NMR/modeled structure.

The NMR-based model reported here is based on that reported previously with some minor modifications. A few revisions to the model were made on the basis of 17 observed aliphatic−aromatic NOE interactions that are summarized in Table 1. These measurements were made possible by the inclusion of unlabeled phenylalanine and tyrosine in the specific labeling medium, so that the *de novo* synthesis of deuterated phenylalanine and tyrosine was suppressed (see Experimental Procedures). The aliphatic methyl resonances were assigned using the H(CCO)NH-TOCSY and the (H)C(CO)NH-TOCSY experiments as described previously (*8*). Aromatic proton resonances were assigned using the ^{15}N-edited NOESY experiment. In general, all of these assignments were consistent with the previously presented structure and allowed some further refinement of the model. Finally, the conformation of the segment from G17 to I21 was revised so that M18 rather than I21 interacts with the nucleotide substrate. This revision is based on the stronger analogy with the Exo I structure (*24*) noted by Hamdan et al. (*9*). The revised model with the enhanced set of NOE constraints was then refined using molecular dynamics protocols as described previously (*8*). The coordinates for the modeled structure now have been deposited in the Protein Data Bank, ID 1MGZ.

A comparison of the CSI data for the ϵ186 and $\epsilon186-\theta$ complex indicates that the secondary structure of ϵ186 is not significantly altered upon binding θ (Figure 1). A small stretch of previously unassigned residues in ϵ186 (K158−L161) (*6*) is predicted to adopt β-strand secondary structure in the complex, which may be significant for the interaction with θ. In addition, according to the CSI analysis, only residues I5−T15 of the complex are predicted to comprise β-strand 1, as compared to residues T3−T15 in the free protein. This is consistent with a complete analysis of NOE data and prediction of ϕ and ψ dihedral angles, via TALOS, made on ϵ186 previously (*8*). Residues M18−I21 also are not predicted to form another β-strand, again in agreement

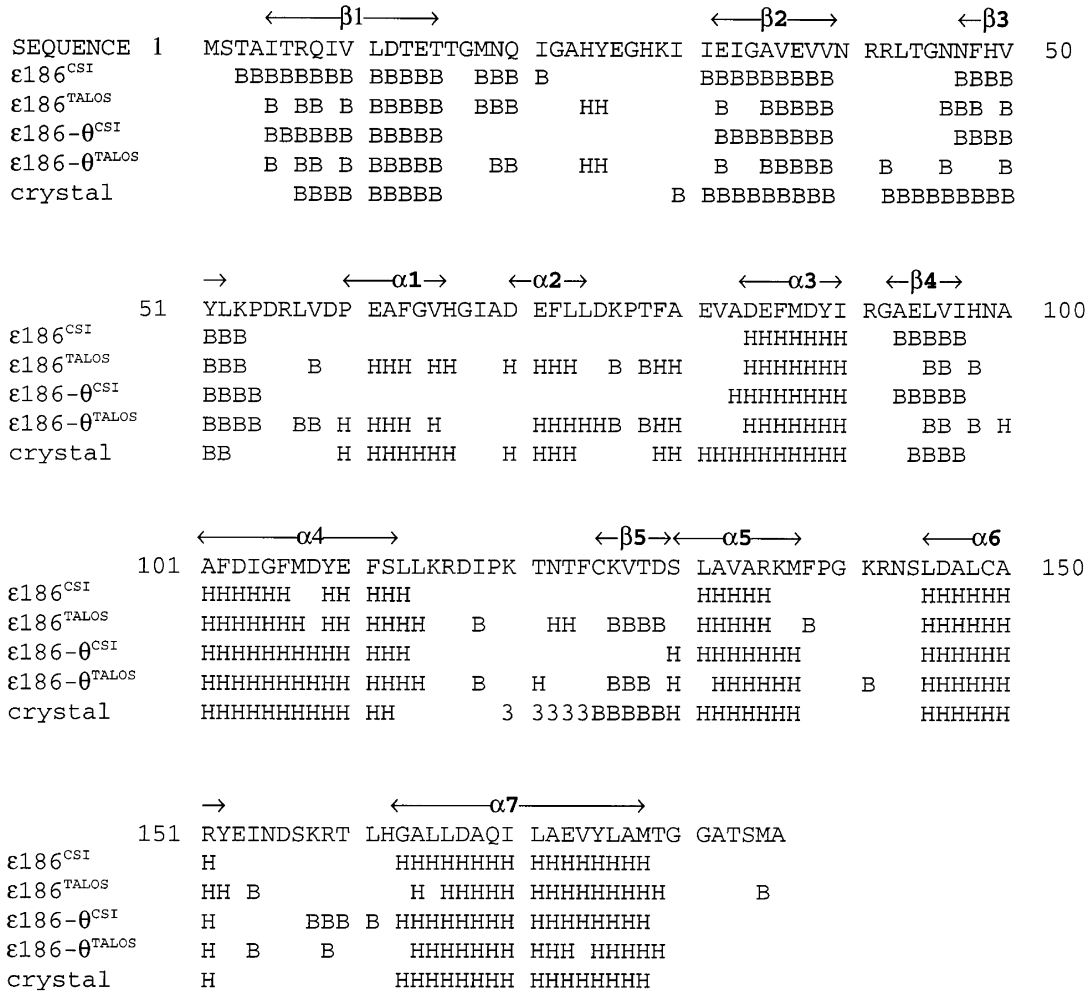


FIGURE 1: Comparison of the secondary structure predictions for ϵ186 and ϵ186−θ. Entries correspond to predictions based on CSI and TALOS calculations for ϵ186 uncomplexed (from Figure 2 of ref *8*) and complexed with θ (present work). Results of a recent crystal structure study of ϵ186 in the presence of Mn^{2+} and thymidine 5′-monophosphate by Hamdan et al. (*9*) are also shown. Symbols B, H, and 3 refer to an extended strand usually involved in a β-sheet, helix, or 3_{10} helix, respectively. The concensus secondary structure is based on CSI and TALOS data.

Table 1: Assigned Methyl−Aromatic NOEs in Selectively Protonated ϵ186[a]

aliphatic methyl	aromatic residue	aliphatic methyl	aromatic residue
I9	Y89	L113	Y109
I154	Y152	L11	F86
I68	Y51	L114	F111, F79
I30 CD	F79	V174	Y152
V96	Y175	V50	F48, Y51, F86
V133	Y175	V82	F48, F86, F111

[a] These NOEs connect the selectively protonated methyl groups of leucine, valine, or isoleucine with the aromatic protons of phenylalanine and tyrosine residues.

with a complete analysis of the NMR data from free ϵ186 (*8*) and with the recently published crystal structure (*9*).

A comparison of the amide shifts obtained for ϵ186 and for the ϵ186−θ complex immediately revealed that most of the perturbations were concentrated in several regions of the protein, particularly the loop preceding the final α7 helix and the beginning of α7, α-helix 1, β-strand 3, and to a lesser extent β-strand 2. However, this comparison was limited by the fact that the two systems were studied under different buffer and temperature conditions. We thus obtained a ^{1}H−^{15}N HSQC spectrum for the ϵ186−θ complex at 20 °C and in 0.9 M Tris buffer. Approximately 84% of the amide resonances of ϵ186−θ in this spectrum could be reassigned unambiguously by comparison with the higher temperature spectra. The remaining amide shift assignments were corrected using a chemical shift factor derived from the average value for the reassigned peaks. In general, the chemical shift mapping results obtained from the initial comparison of the data obtained at different temperatures were in good qualitative agreement with the data obtained for the comparison at the same sample temperature.

Figure 2 shows an overlay of the ^{1}H−^{15}N HSQC spectra of ϵ186−θ (red) and free ϵ186 (black). Residues that exhibited ^{15}N amide shifts greater than 1 ppm are also indicated. A bar graph summarizing the change in ^{15}N amide shift upon binding θ is shown in Figure 3A. The amide proton chemical shifts are summarized in Figure 3B. These plots demonstrate that the maximum chemical shift changes occur primarily in β-strand 3, the loop preceding α-helix 7 and the beginning of α7, α-helix 1, and to a lesser extent in β-strand 2. Residues D59 and A69, which are located before and after α1, and residues L161 and H162, which are located in a loop just before α7, exhibit some of the largest shifts.

Since the perdeuterated protein was grown on media containing the metabolic precursors for methyl-protonated

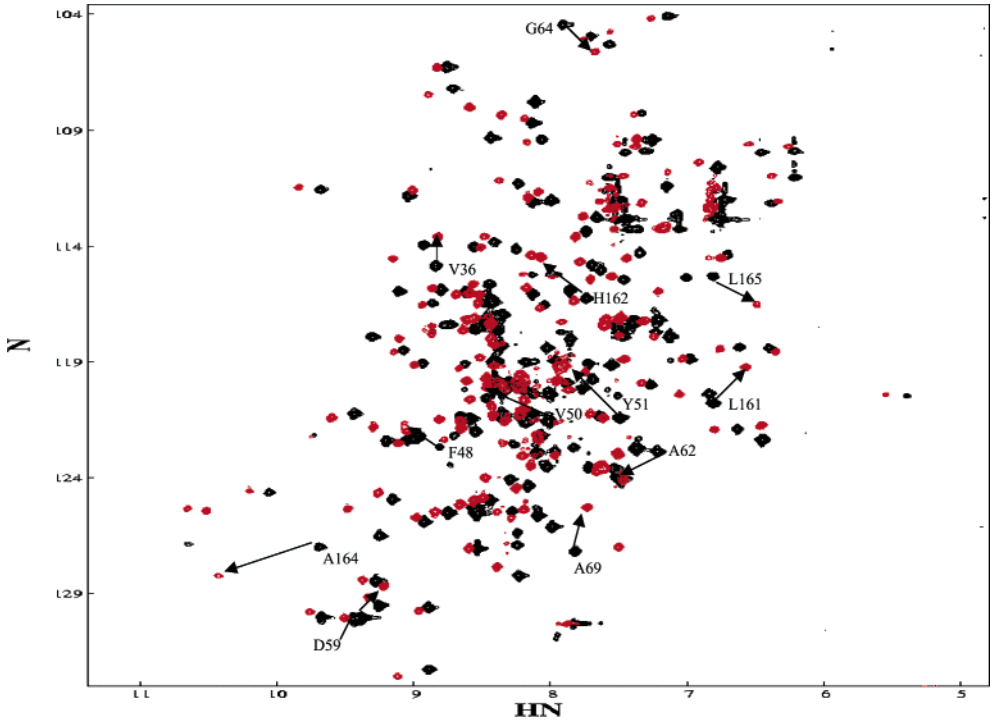
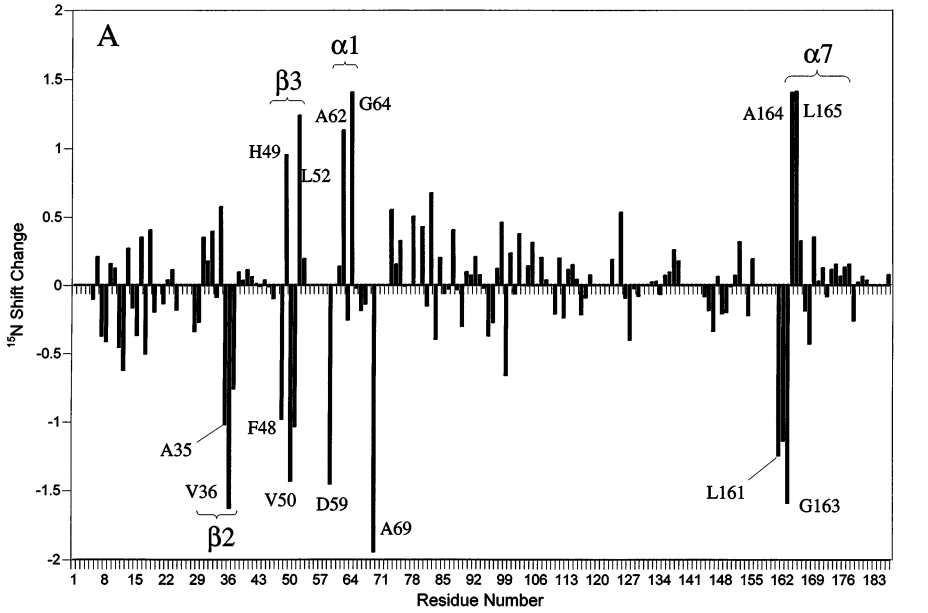
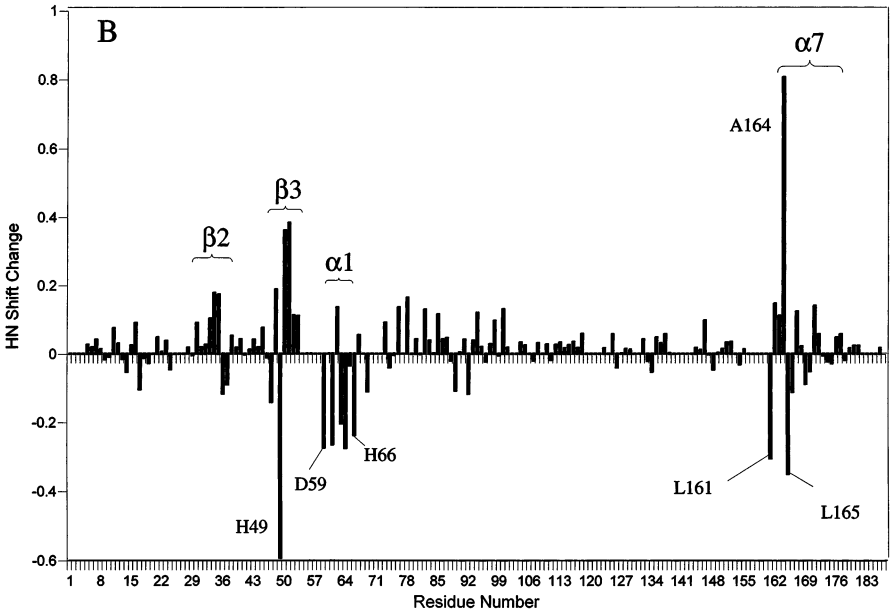


FIGURE 2: Overlay of the $^1H-^{15}N$ HSQC spectra of [U-^2H,^{13}C,^{15}N] methyl protonated ϵ186 in complex with unlabeled θ (ϵ186$-\theta$) shown in red and free [U-^2H,^{13}C,^{15}N] methyl protonated ϵ186 shown in black. Residues exhibiting ^{15}N chemical shift changes greater than 1 ppm are labeled. The spectra were obtained using Varian's gNhsqc pulse sequence on a UNITY INOVA 600 MHz spectrometer with 128 × 512 complex points and acquisition times of 71 and 64 ms in t_1 and t_2, respectively. Eight scans were acquired per increment with a 1.0 s delay between scans. The free enzyme was in 0.9 M Tris-d_{11}, pH 7.0, buffer, 1 mM DTT-d_{10}, and 1 mM $MgCl_2$. The complex was in 200 mM Tris-d_{11}, pH 7.0, buffer, 1 mM DTT-d_{10}, and 2 mM $MgCl_2$. The spectrum of the complex was obtained at 30 °C, and the spectrum of free ϵ186 was acquired at 20 °C.

aliphatic residues, it was also possible to monitor the methyl shift sensitivity of these residues to the presence of θ. Changes in the ^{13}C chemical shifts of protonated leucine, isoleucine, and valine methyl groups, upon binding θ, are shown in Figure 3C. In generating this figure, we selected the valine or leucine methyl resonance which exhibited the larger of the two shift values. We note that this type of comparison is biased toward regions that contain these aliphatic residues, but this complements the information from the amide ^1H and ^{15}N shifts which are not subject to this type of bias. An overlay of the $^1H-^{13}C$ HSQC spectra of ϵ186 and ϵ186$-\theta$, shown in Figure 4, covering the region containing the leucine, valine, and isoleucine δ-methyl protons, indicates that most of the methyl groups do not show significant shifts, consistent with the overall preservation of secondary and tertiary structure of ϵ. Both spectra were acquired at 20 °C, and the samples were dissolved in 0.9 M Tris buffer. These data were used to generate Figure 3C. Among the more significantly shifted methyl resonances are I31, V50, V58, I68, L74, I154, L161, L165, and L166. A comparison of these results with the amide shifts summarized in Figure 3A,B indicates that most of these residues are situated in regions of the protein showing larger amide shifts. V31 is located at the edge of β1 and V50 in β2 which, based on amide shifts, are located in the vicinity of the ϵ186$-\theta$ interface. V58 and I68, which bracket helix α1, are located adjacent to D59 and A69, which, as discussed above, exhibit large amide shifts. L161 is in the loop preceding α7, and L165 and L166 are located in α7. Alternatively, L161 and L165 are fully exposed and hence probably undergo direct hydrophobic interactions with θ. The remaining residues are partially exposed and hence may be involved in some hydrophobic contact with θ. The significant methyl shifts for L161 and L165 parallel the large amide shifts for these residues noted above and are consistent with a hydrophobic interaction with θ. The ^{13}C δ-methyl resonance of I154 is significantly shifted (Figure 3C), yet I154 is not in a region of large amide shifts. This apparent discrepancy may arise due to ring current effects, caused by a change in the position of the nearby Y152 residue upon binding θ.

Panels A and B of Figure 5 show the modeled structure of ϵ186 determined previously (*8*) with the residues color-coded to show the amide ^{15}N chemical shift changes upon binding θ. The color scale varies from white (no chemical shift change) to red (maximum chemical shift change). A "front" view looking into the active site is shown in Figure 5A. The highly conserved acidic residues D12, E14, D103, and D167, involved in divalent metal ion binding at the active site (*8*), are also shown in this figure. A "top" view of the protein, looking down into the active site, is shown in Figure 5B. This view is tilted forward with respect to the view in Figure 5A and contains a modeled trinucleotide substrate A-C-G. It is apparent from this figure that many of the residues exhibiting the largest amide ^{15}N chemical shift changes upon binding θ are located above and behind the active site where the substrate binds. However, the amide resonances of A62 and H162, which also exhibit large amide chemical shift changes upon binding θ, are positioned in close proximity to the trinucleotide substrate (Figure 6). A62 is positioned near the guanine base of the substrate, and H162 is positioned near the guanosine 5′-phosphate. The side chain of H162 may play a catalytic role, by positioning a water





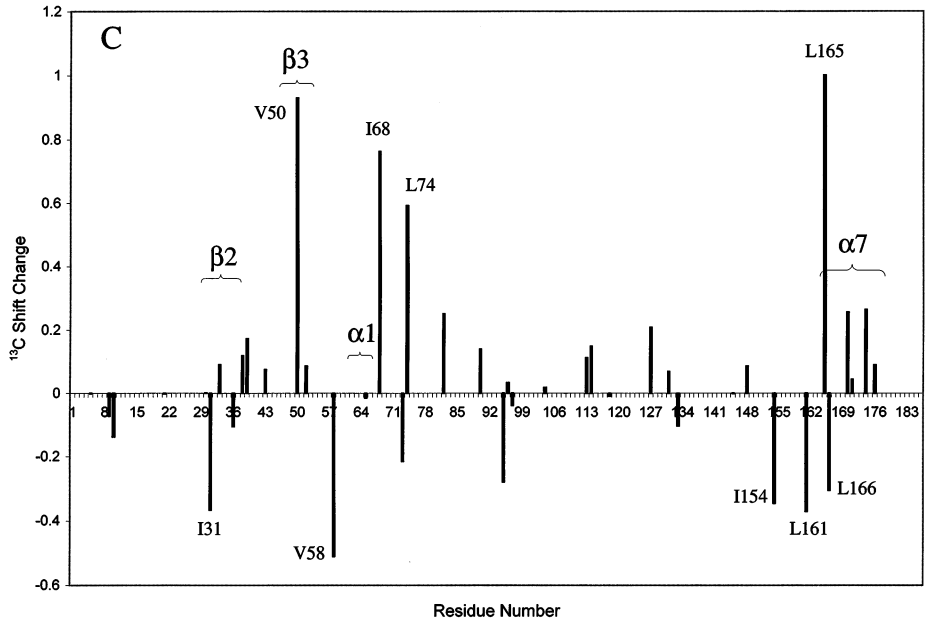


FIGURE 3: (A) Bar graph summary of the ^{15}N amide chemical shift changes of ϵ186 upon binding θ. Several of the residues and structural elements showing the most significant shifts are indicated. To compare the amide shifts of ϵ186$-\theta$ and ϵ186, 84% of the ϵ186$-\theta$ amide shifts, assigned at 30 °C, were reassigned in the ^{1}H$-^{15}$N HSQC spectrum at 20 °C and 0.9 M Tris buffer conditions. The remaining shift changes were corrected by subtracting the difference between the average amide shift changes recorded for the complex at 30 and 20 °C. (B) Bar graph summary of the ^{1}H amide chemical shift changes of ϵ186 upon binding θ. (C) Bar graph summary of the leucine, valine, and isoleucine δ ^{13}C methyl chemical shift changes of ϵ186 upon binding θ.

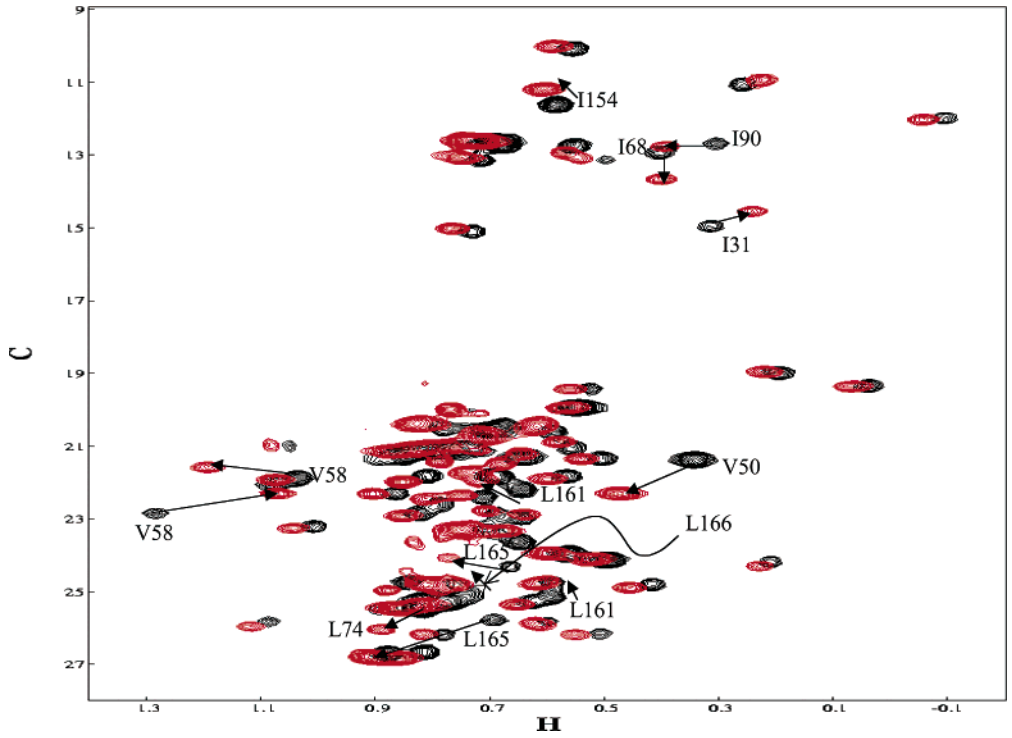


FIGURE 4: Constant time $^1H-^{13}C$ HSQC spectra of the region of the spectrum containing the aliphatic methyl groups arising from leucine, valine, and isoleucine δ methyl groups. The spectrum of ϵ186 is shown in black, while the spectrum of ϵ186, in the presence of unlabeled θ, is shown in red. Several of the more significantly shifted methyl resonances are indicated. The spectra were obtained using Varian's gChsqc pulse sequence on a UNITY INOVA 600 MHz spectrometer with 310 × 512 complex points and acquisition times of 26 and 64 ms in t_1 and t_2, respectively. Thirty-two scans were acquired per increment with a 1.0 s delay between scans. The free enzyme was in 0.9 M Tris-d_{11}, pH 7.0, buffer, 1 mM DTT-d_{10}, and 1 mM $MgCl_2$. The complex was in 0.9 M Tris-d_{11}, pH 7.0, buffer, 1 mM DTT-d_{10}, and 2 mM $MgCl_2$. Both spectra were acquired at 20 °C.

molecule or stabilizing a hydroxyl anion for nucleophilic attack of the 5′-phosphate of the terminal base (*8*), in a fashion similar to Y497 of the Klenow fragment of Pol I (*26*).

Figure 7 shows a view from the opposite side of the protein, with respect to the view in Figure 5A. This view shows all of the residues exhibiting amide ^{15}N chemical shift changes larger than ±1 ppm (side chain carbon atoms in green) and the residues exhibiting methyl carbon chemical shifts larger than ±0.3 ppm (side chain carbon atoms in cyan). This figure shows that most of the residues exhibiting significant chemical shift changes upon binding θ are located in β-strands 2 and 3, α-helix 1, and residues D59 and A69 positioned on either side of α1, the turn preceding α-helix 7, and the beginning of α-helix 7. According to the secondary structure analysis, α-helix 7 begins with residue G163. In addition, the previously unassigned residues K158, R159, and T160 (*8*) are predicted to adopt β-strand secondary structure in the complex and may also be participating in the interaction with θ. Thus, the interaction interface is comprised primarily of β-strands 2 and 3, α-helix 1, the loop preceding α-helix 7, and the N-terminal residues of α-helix 7, as shown in Figures 5B and 6.

DISCUSSION

The ϵ and θ subunits of the Pol III holoenzyme are bound tightly together as components of the Pol III core, but the function of this interaction is unknown. ϵ carries the 3′-exonuclease (proofreading) activity that is of critical importance to chromosomal replication fidelity, but no function has been assigned to θ. In vitro experiments with purified proteins revealed a small (~2.5-fold) enhancement by θ of the ϵ 3′-exonuclease activity, suggesting that θ might have some stimulatory effect on ϵ (*27*). On the other hand, a mutant strain lacking θ proved normally viable and did not show any noticeable phenotype (*3*). The *holE* gene, encoding θ (*28*), is well conserved among the Gram-negative bacteria. The Gram-negative organisms are characterized by a replicative polymerase that contains separate polypeptides for the polymerase and the proofreader. This is in contrast to the Gram-positive bacteria and higher organisms, in which the two activities of the replicating polymerase reside on the same polypeptide. In those other organisms, no protein homologous to θ is found, nor do the polymerases contain any (sub)domain with homology to θ. Thus, the possibility arises that the need for θ is related to the separate nature of the proofreading subunit. Hamdan et al. (*9*) have recently proposed that the structural analogy between θ and the C-terminal domain of exonuclease I (*24*), a nuclease related to ϵ, might imply a similar location and function. As discussed below, the present study provides some support for this suggestion.

In the present study, we have extended our previous analysis of the solution structure of ϵ (*8*) by examining the changes in NMR spectra of ϵ186 upon complexing with θ. All studies used ϵ186, the N-terminal domain of ϵ containing both the exonuclease activity and the θ binding site (*6*, *7*, *29*). This study reveals a defined area on ϵ where chemical shift changes were concentrated. While some of those changes could result from indirect, conformational adjustments to binding at more remote locations of the protein, the concentration of shifts in one spatially contiguous region

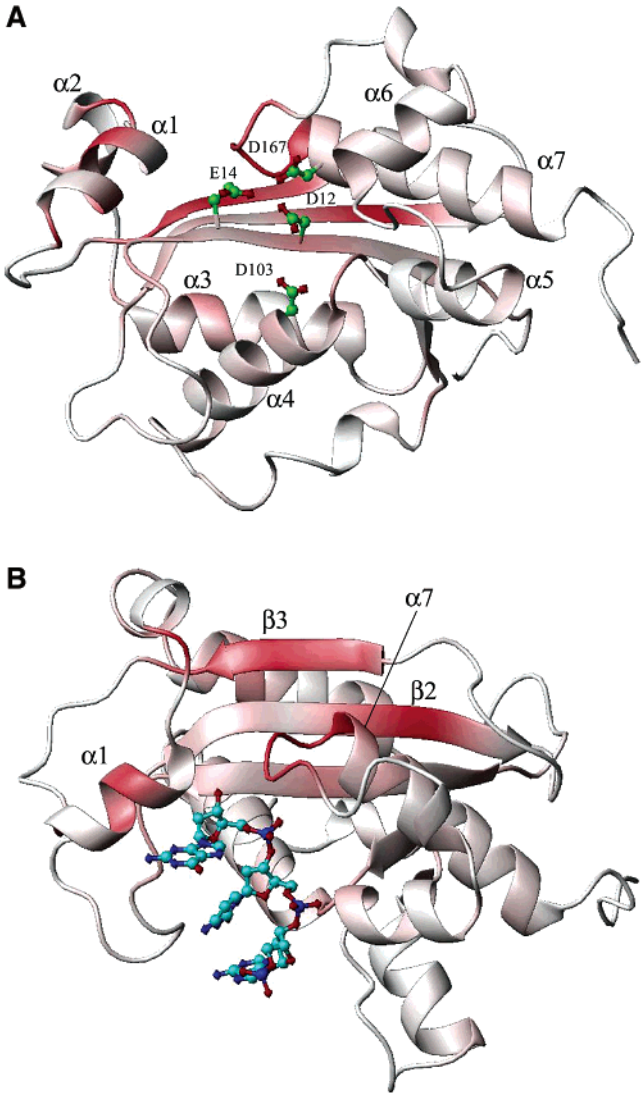


FIGURE 5: (A) Model of ϵ186 color-coded to show the amide ^{15}N chemical shift change upon binding θ. The color scale varies from white (no chemical shift change) to red (maximum chemical shift change). The catalytic residues D12, E14, D103, and D167 are also shown. (B) Model of ϵ186 complexed with the 5′ A-C-G trinucleotide color-coded to show the amide ^{15}N chemical shift change upon binding θ. Secondary structure elements exhibiting the greatest changes in amide ^{15}N chemical shift are also labeled. The model has been rotated approximately 90° forward to better illustrate the relationship of the ϵ186−θ interface and the active site. This figure was generated using the program MOLMOL (*25*).

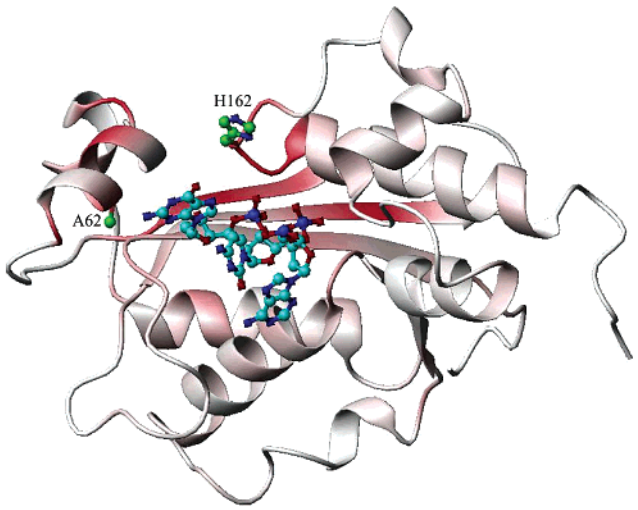


FIGURE 6: Model of ϵ186/A-C-G trinucleotide complex color-coded to show the amide ^{15}N chemical shift change upon binding θ. The color scale varies from white (no chemical shift change) to red (maximum chemical shift change). Also shown are the side chains of amino acids A62 and H162 of ϵ186 in close proximity to the trinucleotide substrate. The amide ^{15}N chemical shifts of A62 and H162 changed by more than 1 ppm upon binding θ. This figure, showing the same view of the enzyme as in Figure 5A, was generated using the program MOLMOL (*25*).

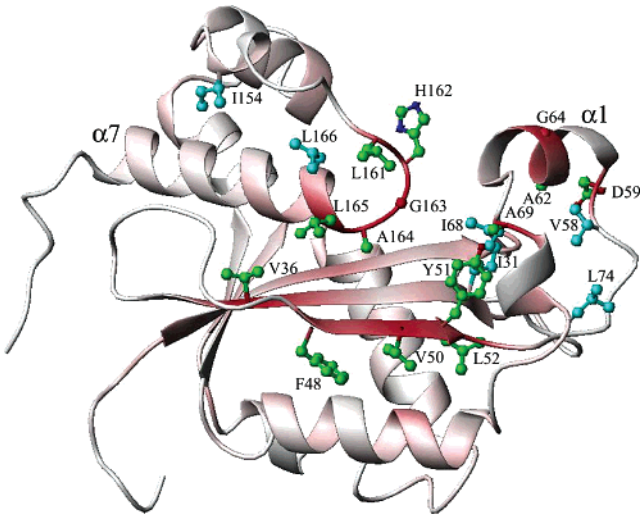


FIGURE 7: Model of ϵ186 showing the side-chains of residues exhibiting the largest ^{15}N amide and ^{13}C methyl chemical shifts. The ribbon is color coded to show the amide ^{15}N chemical shift change upon binding θ. The color scale varies from white (no chemical shift change) to red (maximum chemical shift change). The side-chains of residues exhibiting ^{15}N chemical shifts of more than ±1.0 ppm are shown with carbon atoms colored green. The side-chains of residues exhibiting ^{13}C chemical shifts of more than ±0.3 ppm are shown with carbon atoms colored cyan. V50, L161, and L165 also exhibited ^{13}C methyl shifts larger than ±0.3 ppm. The figure was generated using the program MOLMOL (*25*).

most probably indicates the binding surface of the two subunits. The binding surface is comprised primarily of β-strands 2 and 3, α-helix 1, the loop preceding α-helix 7, and the N-terminal residues of α-helix 7, as shown in Figures 5B and 7. In the views shown in Figures 5 and 7, in which the substrate binding site is located above and toward the front side of the central β-sheet formed by strands β1, β2, and β3, the ϵ−θ interaction surface is located above and behind the DNA binding site. Of particular interest are the interactions with the N-terminal residues of helix α7 and the loop immediately preceding α7. These regions contain the catalytically important residues His162 and Asp167 (*30*−*32*), which are part of the exonuclease Exo III motif (HxxxxD) (*30*). This interaction, resulting possibly in stabilization of components of the catalytic site, would be consistent with the observed enhancement, albeit moderate, of the exonuclease activity by θ (*27*).

A stabilization of ϵ by θ has also been inferred from recent genetic experiments utilizing unstable *dnaQ* alleles (the *dnaQ* gene encodes ϵ). Notably, the *dnaQ49* mutation, characterized by a V96G substitution in ϵ (*30*, *33*), yields an unstable protein and a temperature-dependent mutator phenotype. The poor stability of the V96G variant of ϵ results in a strong mutator phenotype even at low temperatures (*4*). Our recent

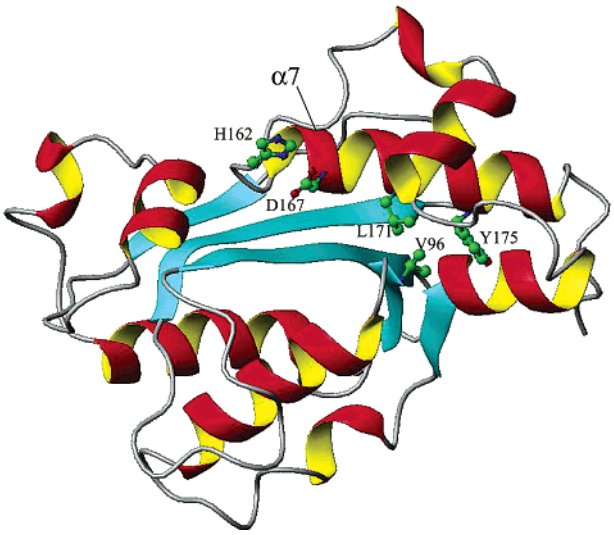


FIGURE 8: Model of ϵ186 showing the close proximity of V96 to L171 and Y175 of α-helix 7. The catalytic residues D12, E14, D103, H162, and D167 are also shown. This figure was generated using the program MOLMOL (*25*).

studies indicate that deletion of the holE gene to eliminate θ from the *dnaQ49* mutant leads to a very strong mutator phenotype even at low temperatures (*4*), suggesting that the $\epsilon-\theta$ interaction is of even more critical importance for stabilizing the V96G mutant. These data suggest that θ stabilizes mutant ϵ and, by implication, the wild-type ϵ. Our model for ϵ186 along with the proposed $\epsilon-\theta$ binding surface can provide a rationale for these observations. Specifically, V96 of ϵ is located in a hydrophobic pocket that includes L171 and Y175 in helix α7 (Figure 8). The observation of a V96 methyl−Y175 aromatic NOE in the present study (Table 1) further supports the existence of this structural motif. The hydrophobic interactions of V96 with these residues are likely important for holding α-helix 7 in place for the proper positioning of D167 and H162 in the active site, and loss of the V96 interactions with the α7 residues (in addition to a disruption of strand β4) would readily account for the temperature-dependent proofreading inactivation of the *dnaQ49* mutant. The interaction of θ with the N-terminal portion of α7 could provide some additional stabilization to overcome the effect resulting from loss of the hydrophobic interactions noted above. As described above, the $\epsilon-\theta$ interface defined on the basis of amide chemical shift data involves strong interactions with the N-terminus of α-helix 7 and the loop preceding it, including H162. This suggests that θ also stabilizes α7 for proper positioning with respect to the central β-sheet, in effect acting like a clamp for the β-sheet and the α-helix. Lack of both stabilizing factors in the V96G mutant lacking θ may make it impossible to maintain the critical H162 and D167 residues in close proximity to the catalytic site, accounting for the complete loss of proofreading in this mutant (*4*).

In the recent study of the crystal structure of ϵ186 (*9*), it was noted that the structure of θ is reminiscent of the small C-terminal domain, residues 420−466, of exonuclease I (*24*). This domain interacts with the β3−α1 loop and α1 and α2 helices of the N-terminal exonuclease domain. The interaction of θ with the corresponding region of ϵ186 would place it in a position where it could influence DNA binding. The proposed interaction with the β3 strand and α1 helix is consistent with the amide shift perturbations observed in the present study and, hence, provides further support for a functional role of θ which may be similar to the C-terminal domain of exonuclease I. However, as noted above, θ also significantly perturbs the shifts of the N-terminal residues of helix α7, and as noted above, we believe that this interaction may be significant for stabilizing the structure of the V96G mutant. There is also a significant shift change observed for the catalytically important residue H162. Possibly, the disordered C-terminal region of θ (*11*) could bridge the active site, analogous to residues 355−358 of exonuclease I, encircling the substrate and leading to higher processivity. It is anticipated that a determination of the structure of the ϵ186−θ complex will shed additional light on the nature of these interactions.

NOTE ADDED IN REVISION

After the initial submission of this paper the coordinates of the X-ray structure from Hamdan et al. were released (PDB ID 1J53 and 1J54). Superposition of the modeled structure (PDB ID 1MGZ) onto the X-ray structure yields an RMS deviation of 2.4 Å when α carbons of all residues resolved in the X-ray structure are compared, with a smaller RMS deviation of 1.6 Å when comparison is restricted to residues within regions of secondary structure. The active site is also quite similar in the two structures.

ACKNOWLEDGMENT

The authors dedicate this paper to our late friend and colleague, Dr. Greg Mullen, who made major contributions to the structural characterization of polymerases and related enzymes. The authors gratefully acknowledge the research group of Dr. Lewis Kay, University of Toronto, for providing the NMR pulse sequences and Dr. Joseph Krahn, NIEHS, for contributions to the molecular modeling.

REFERENCES

1. Kelman, Z., and O'Donnell, M. (1995) DNA polymerase III holoenzyme: structure and functiona of a chromosomal replicating machine, *Annu. Rev. Biochem. 64*, 171−200.
2. Glover, B. P., and McHenry, C. S. (2001) The DNA polymerase III holoenzyme: an asymmetrical dimeric replicative polymerase complex with leading and lagging strand polymerases, *Cell 105*, 925−934.
3. Slater, S. C., Lifsics, M. R., O'Donnell, M., and Maurer, R. (1994) *holE*, the gene coding for the θ subunit of DNA polymerase III of *Escherichia coli*: characterization of a *holE* mutant and comparison with a *dnaQ* (ϵ-subunit) mutant, *J. Bacteriol. 176*, 815−821.
4. Taft-Benz, S. A., and Schaaper, R. M. (2002) (submitted for publication).
5. Schaaper, R. M. (1999) Base selection, proofreading, and mismatch repair during DNA-replication in *Escherichia coli*, *J. Biol. Chem. 268*, 23762−23765.
6. Perrino, F. W., Harvey, S., and McNeill, S. M. (1999) Two functional domains of the ϵ subunit of DNA polymerase III, *Biochemistry 38*, 16001−16009.
7. Taft-Benz, S. A., and Schaaper, R. M. (1999) The C-terminal domain of DnaQ contains the polymerase binding site, *J. Bacteriol. 181*, 2963−2965.
8. DeRose, E. F., Li, D., Darden, T., Harvey, S., Perrino, F. W., Schaaper, R. M., and London, R. E. (2002) Model for the catalytic domain of the proofreading ϵ subunit of *Escherichia coli* DNA polymerase III based on NMR structural data, *Biochemistry 41*, 94−110.
9. Hamdan, S., Carr, P. D., Brown, S. E., Ollis, D. L., and Dixon, N. E. (2002) Structural basis of proofreading during replication of the *Escherichia coli* chromosome, *Structure 10*, 535−546.

10. Li, D., Allen, D. L., Harvey, S., Perrino, F. W., Schaaper, R. M., and London, R. E. (1999) A preliminary CD and NMR study of the *Escherichia coli* DNA polymerase III θ subunit, *Proteins 36*, 111−116.
11. Keniry, M. A., Berthon, H. A., Yang, J. Y., Miles, C. S., and Dixon, N. E. (2000) NMR Solution structure of the θ subunit of DNA polymerase III from *Escherichia coli*, *Protein Sci. 9*, 721−733.
12. Aghazadeh, B., Zhu, K., Kubiseski, T. J., Liu, G. A., Pawson, T., Zheng, Y., and Rosen, M. K. (1998) Structure and mutagenesis of the Dbl homology domain, *Nat. Struct. Biol. 5*, 1098−1107.
13. Yamazaki, T., Lee, Revington, M., Mattiello, D. L., Dahlquist, F. W., Arrowsmith, C. H., Kay, L. E. (1994) An HNCA pulse scheme for the backbone assignment of ^{15}N, ^{13}C, ^{2}H-labeled proteins: Application to a 37-kDa Trp Repressor-DNA complex, *J. Am. Chem. Soc. 116*, 6464.
14. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., Kay, L. E. (1994) A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins with high sensitivity, *J. Am. Chem. Soc. 116*, 11655−11666.
15. Kay, L. E., Guang, Y. X., and Yamazaki, T. (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation, *J. Magn. Reson., Ser. A 109*, 129−133.
16. Pervushin, K., Riek, R., Wider, G., and Wüthrich (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 12366−12371.
17. Yang, D., and Kay, L. E. (1999) Improved 1HN-detected triple resonance TROSY-based experiments, *J. Biomol. NMR 13*, 3−10.
18. Zhang, O., Kay, L. E., Olivier, J. P., Forman-Kay, J. D. (1994) Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulse field gradient NMR techniques, *J. Biomol. NMR 4*, 845−858.
19. Zwahlen, C., Gardner, K. H., Sarma, S. P., Horita, D. A., Byrd, R. A., and Kay, L. E. (1998) An NMR experiment for measuring methyl−methyl NOEs in ^{13}C-labeled proteins with high resolution, *J. Am. Chem. Soc. 120*, 7617−7625.
20. Delaglio, F., Grzeseik, 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.
21. Johnson, B. A., and Blevins, R. A. (1994) NMRView: a computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−614.
22. Wishart, D. S., and Sykes, B. D. (1994) The ^{13}C Chemical-Shift Index: a simple method for the indentification of protein secondary structure using ^{13}C chemical-shift data, *J. Biomol. NMR 4*, 171−180.
23. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR 13*, 289−302.
24. Breyer, W. A., and Matthews, B. W. (2000) Structure of *Escherichia coli* Exonuclease I suggests how processivity is achieved, *Nat. Struct. Biol. 7*, 1175−1128.
25. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.
26. Beese, L. S., and Steitz, T. A. (1991) Structural basis for the 3′−5′ exonuclease activity of *Escherichia coli* DNA polymerase I.; a two metal ion mechanism, *EMBO J. 10*, 25−33.
27. Studwell-Vaughan, P. S., and O'Donnell, M. (1993) DNA polymerase III accessory proteins. V. θ encoded by *holE*, *J. Biol. Chem. 268*, 11785−11791.
28. Carter, J. R., Franden, M. A., Aebersold, R., Kim, D. R., and McHenry, C. S. (1993) Isolation, sequencing and overexpression of the gene encoding the θ subunit of DNA polymerase III holoenzyme, *Nucleic Acids Res. 21*, 3281−3286.
29. Jonczyk, P., Nowicka, A., Fijalkowska, I. J., Schaaper, R. M., and Ciesla, Z. (1998) *In vivo* protein interactions within the *Escherichia coli* DNA polymerase III core, *J. Bacteriol. 180*, 1563−1566.
30. Taft-Benz, S. A., and Schaaper, R. M. (1998) Mutational analysis of the 3′→5′ proofreading exonuclease of *Escherichia coli* DNA polymerase III, *Nucleic Acids Res. 26*, 4005−4011.
31. Barnes, M. H., Spacciapoli, P., Li, D. H., and Brown, N. C. (1995) The 3′−5′ exonuclease site of DNA polymerase III from gram-positive bacteria: definition of a novel motif structure, *Gene 165*, 45−50.
32. Slupska, M. M., King, A. G., Lu, L. I., Lin, R. H., Mao, E. F., Lackey, C. A., Chiang, J.-H., Baikalov, C., and Miller, J. H. (1998) Examination of the role of DNA polymerase proofreading in the mutator effect of miscoding tRNAs, *J. Bacteriol. 180*, 5712−5717.
33. Takano, K., Nakabeppu, Y., Maki, H., Horiuchi, T., and Sekiguchi, M. (1986) Structure and function of *dnaQ* and *mutD* mutators of *Escherichia coli*, *Mol. Gen. Genet. 205*, 9−13.

BI0205451