The proofreading exonuclease subunit ε of *Escherichia coli* DNA polymerase III is tethered to the polymerase subunit α via a flexible linker

Kiyoshi Ozawa[1], Slobodan Jergic[1,2], Ah Young Park[1], Nicholas E. Dixon[1,2] and Gottfried Otting[1,*]

[1]Research School of Chemistry, Australian National University, Canberra ACT 0200 and [2]School of Chemistry, University of Wollongong, NSW 2522, Australia

ABSTRACT

Escherichia coli DNA polymerase III holoenzyme is composed of 10 different subunits linked by noncovalent interactions. The polymerase activity resides in the α-subunit. The ε-subunit, which contains the proofreading exonuclease site within its N-terminal 185 residues, binds to α via a segment of 57 additional C-terminal residues, and also to θ, whose function is less well defined. The present study shows that θ greatly enhances the solubility of ε during cell-free synthesis. In addition, synthesis of ε in the presence of θ and α resulted in a soluble ternary complex that could readily be purified and analyzed by NMR spectroscopy. Cell-free synthesis of ε from PCR-amplified DNA coupled with site-directed mutagenesis and selective ^{15}N-labeling provided site-specific assignments of NMR resonances of ε that were confirmed by lanthanide-induced pseudocontact shifts. The data show that the proofreading domain of ε is connected to α via a flexible linker peptide comprising over 20 residues. This distinguishes the α:ε complex from other proofreading polymerases, which have a more rigid multidomain structure.

INTRODUCTION

The DNA polymerase III (Pol III) holoenzyme is the major chromosomal replicase in *Escherichia coli* (1). This enzyme complex is composed of 10 different polypeptide subunits. The catalytic core contains one each of the α (130 kDa), ε (27 kDa) and θ (8.8 kDa) subunits encoded by the *dnaE*, *dnaQ* and *holE* genes, respectively. The α-subunit contains the 5′–3′ DNA polymerase active site (2,3), the ε-subunit is responsible for 3′–5′ proofreading exonuclease activity (4) and the θ-subunit has no identified enzymatic activity (5). The α:ε:θ core complex is active alone as a proofreading DNA polymerase, and copurification of these three subunits demonstrates their tight physical association (6,7). Direct interactions between ε and α (8) and ε and θ (5) have been demonstrated using purified subunits, but no interaction has been detected between α and θ.

The θ-subunit is not essential, as a Δ*holE* mutant is normally viable (9) and θ has only a modest stimulatory effect on the exonuclease activity of ε on a mispaired primer terminus (5). Genetic studies with the temperature-sensitive *dnaQ49* mutant allele indicated that θ stabilizes the structure of ε (10), and this effect may also be achieved by the θ homolog HOT (homolog of theta) encoded by bacteriophage P1 (11). Remarkably, the phage genome (94 kb) relies on *E. coli* Pol III for its replication but, with the exception of the Ban (DnaB analog) DNA helicase and SSB (single-stranded DNA-binding protein), does not encode any other replication proteins (12).

Two crystal structures of α have been determined: of a C-terminally truncated version of *E. coli* α (13), and of full-length *Thermus aquaticus* α (14). Structures are also known of the N-terminal globular domain of ε (ε186), both alone (15) and in complex with HOT (16,17). The structure of the ε186:θ complex determined by nuclear magnetic resonance (NMR) spectroscopy (18,19) is in agreement with the ε186:HOT structure. No structure has been determined of full-length ε but residues following ε186 within its C-terminal region (in the following referred to as C-terminal segment (CTS) of ε or εCTS; Figure 1) are known to be responsible for binding of α (20,21). This 57-residue segment contains a Q-linker sequence proposed to provide a flexible tether between domains (22), followed by the C-terminal 40-residue segment that has been shown to interact tightly with α when fused to maltose-binding protein (21).

The present study was carried out to obtain structural information about the α:ε:θ complex. By studying the ε:θ,

```
  1 MSTAITRQIV LDTETTGMNQ IGAHYEGHKI IEIGAVEVVN RRLTGNNFHV
 51 YLKPDRLVDP EAFGVHGIAD EFLLDKPTFA EVADEFMDYI RGAELVIHNA
101 AFDIGFMDYE FSLLKRDIPK TNTFCKVTDS LAVARKMFPG KRNSLDALCA
151 RYEIDNSKRT LHGALLDAQI LAEVYLAMTG GQTSMAFAME GETQQQQGEA
201 TIQRIVRQAS KLRVVFATDE EIAAHEARLD LVQKKGGSCL WRA
```

Figure 1. Amino acid sequence of full-length ε. Residues 7–180 have a defined conformation in the crystal structure of ε186 (15). The 57 residues of the εCTS are highlighted in bold. Secondary structure prediction suggests an α-helical segment in the εCTS with high propensity; the corresponding residues are underlined. Alanine and threonine residues in the CTS of ε are highlighted in red and yellow, respectively.

α:ε and α:ε:θ complexes and probing potential weak interactions between α and ε186 by a novel two-pronged NMR assignment strategy, we show that residues in the εCTS are flexible and that those in the Q-linker region remain so even when assembled in the Pol III core complex. This provides the first experimental evidence that ε is indeed tethered to α by a flexible peptide linker, suggesting a mechanism for transition between polymerization and proofreading modes in Pol III that is fundamentally different from those in other polymerases whose structures in both modes are known. In addition, experiments to produce ε by cell-free protein synthesis shed further light on the function of θ.

MATERIALS AND METHODS

Materials

L-[15N]Alanine and L-[15N]threonine were obtained from Spectra Stable Isotopes. Synthetic oligonucleotides were purchased from GeneWorks (Hindmarsh, SA, Australia); sequences of oligonucleotides used are listed in Table S1.

Protein purification

For large-scale isolation of α, *E. coli* cells (BL21::λDE3 *recA*) harboring the plasmid pND517 encoding the *dnaE* gene under control of a *tac* promoter (23) were grown aerobically at 30°C and in 20 l of Z medium of pH 7.3 (24) supplemented with 100 mg/l ampicillin in a fermenter with pH control (BIOSTAT C, B. Braun Biotech International, Melsungen, Germany). Overexpression of α was induced by addition of 0.5 mM isopropyl-β,D-thiogalactoside at $A_{595} = 1.5$ and the induced culture was grown to an A_{595} of ∼10 (4 h), yielding about 250 g of cells. The α-subunit was purified from 54 g of cells using a modified version of the procedure of Wijffels *et al.* (23), with two steps of chromatography on columns of DEAE-Toyopearl 650 M (2.5 × 45 cm) and phosphocellulose (5.5 × 22.5 cm). This was followed by chromatography using a heparin–Sepharose column (2.5 × 17.5 cm) to concentrate the protein. About 96 mg of purified α-subunit were obtained. The preparation of a C-terminal truncation mutant of α, α917 (13) is described in the Supplementary Data.

The Pol III θ-subunit was purified as described (25), except that cells were grown for 2 days at room temperature in an autoinduction medium (26), and the French press lysate was passed through a column of DEAE-Toyopearl in 40 mM Tris–HCl buffer, pH 7.6, prior to chromatography on a phosphocellulose column. 15N-labeled θ was produced using minimal medium supplemented with

15NH₄Cl (Cambridge Isotope Laboratories, Andover, MA, USA) as described (27). T7 RNA polymerase (28) and ε186 (25) were as described. Concentrations of samples of soluble pure proteins (α, ε186 and θ) were determined spectrophotometrically at 280 nm, using calculated values of ε_{280} of 95440, 7680 and 8250/M/cm, respectively (29).

Cell-free synthesis of the ε-subunit

S30 cell extracts from either *E. coli* strain Rosetta::λDE3/pRARE (from Novagen, Gibbtown, NJ, USA) or BL21Star::λDE3 (from Invitrogen, Carlsbad, CA, USA) were prepared by the procedure of Pratt (24,30,31), followed by concentration with polyethylene glycol 8000 as described by Kigawa *et al.* (31,32), and were used interchangeably. Cell-free protein synthesis was carried out for 6–7 h either using an autoinduction system that uses plasmid pKO1166 to direct production of T7 RNA polymerase in S30 extracts (33) at 37°C, or a standard coupled transcription/translation system with purified T7 RNA polymerase at 30 or 37°C as described previously (28,31). The plasmid template pSH1017 (25) was used at a concentration of 16 μg/ml for production of ε. All reaction mixtures containing expressed proteins were clarified by centrifugation (30 000*g*, 1 h) at 4°C.

Cell-free synthesis of selectively 15N-labeled ε in complex with θ

Samples of the ε:θ complex containing 15N-Ala or 15N-Thr labeled ε were prepared by synthesis of ε in the cell-free system in the presence of separately purified unlabeled θ (0.5 mg/ml), with 0.6 ml reaction mixtures at 30°C for 7 h. Following centrifugation, the supernatant was dialyzed against 2 l of NMR buffer (20 mM Tris–HCl, pH 6.9, 1 mM EDTA, 1 mM dithiothreitol) and concentrated to a final volume of about 0.5 ml using Millipore Ultra-4 centrifugal filters (molecular weight cutoff 10 000). D₂O was added to a final concentration of 10% (v/v) prior to NMR measurements.

Cell-free synthesis and purification of the α:ε:θ complexes with selectively 15N-labeled ε

Samples of the α:ε:θ complex containing 15N-Ala or 15N-Thr labeled ε were prepared by cell-free synthesis of ε in the presence of separately purified unlabeled θ (0.5 mg/ml) and α (5.0 mg/ml). Reaction mixtures (0.9 ml) were incubated at 30°C for 7 h. Following centrifugation, the supernatant was dialyzed against 2 l of buffer A (50 mM Tris–HCl, pH 7.6, 1 mM EDTA, 1 mM dithiothreitol), then loaded onto a DEAE-Toyopearl 650 M column (2.5 × 2.5 cm) pre-equilibrated with buffer A and eluted with a linear gradient of 0–1 M NaCl in 150 ml of buffer A. The combined fractions containing the α:ε:θ complex (eluting at about 0.22 M NaCl) were dialyzed against 2 l of buffer A, loaded onto a heparin–Sepharose column (2.5 × 17.5 cm) in buffer A and eluted with a linear gradient of 0–1 M NaCl in 300 ml of buffer A. The α:ε:θ complex eluted at about 0.25 M NaCl and was subsequently dialyzed against NMR buffer, concentrated and prepared for NMR spectroscopy as described above for the ε:θ complex. The purity of the α:ε:θ complex was assessed by

15% SDS–PAGE followed by staining with Coomassie brilliant blue; its concentration was estimated by the method of Bradford (34).

Cell-free synthesis and purification of the ε:θ complex containing ^{15}N-labeled θ

Cell-free synthesis of unlabeled ε at 30°C for 7 h in the presence of separately purified uniformly ^{15}N-labeled θ (0.5 mg/ml) was done in a reaction volume of 0.9 ml. Following centrifugation, the supernatant was dialyzed against buffer A, loaded onto a DEAE-Toyopearl column and eluted as above for the α:ε:θ complex. Excess θ remained unbound, and the first protein peak eluting at about 60 mM NaCl contained the ε:θ complex. Following dialysis against 2 l of NMR buffer and concentration to a final volume of about 0.5 ml as above, D$_2$O was added to a final concentration of 10% (v/v) for NMR measurements. The purity was assessed by 15% SDS–PAGE followed by staining with Coomassie brilliant blue.

Cell-free synthesis of mutant ε:θ complexes from PCR-amplified DNA

Samples of the ε:θ complex containing site-specific mutants of ε were produced by cell-free protein synthesis from linear PCR-amplified DNA with 5′-phosphorylated primers to promote ligation to circular DNA in the reaction mixture (35). The DNA was prepared in two steps, using Vent DNA polymerase (New England Biolabs, Ipswich, MA, USA) for all PCR reactions. First, the site-specific mutation was introduced using two separate PCR reactions (50 μl each) with 20–30 ng of pSH1017 template, where the first reaction used primer 1133 and one of the reverse primers of Table S1 (containing the desired mutation) and the second reaction used one of the forward primers of Table S1 and primer 1134. The PCR products were mixed in an equimolar ratio and the residual primers removed using the Qiaquick PCR purification kit (Qiagen, Hilden, Germany). In the second step, the T7 promoter and terminator sequences were added in two separate PCR reactions (50 μl each) using the primer pairs 1131 and 1134, and 1132 and 1133, respectively (Table S1), with 20–30 ng of purified PCR product from the first step. The two PCR products were mixed in an approximately equimolar ratio and the residual primers removed as above, denatured at 95°C (5 min) and reannealed at room temperature (∼5 min). This generated DNA with complementary single-stranded 8-nt overhangs suitable for cyclization by the intrinsic ligase activity of the cell extract (35). The reannealed PCR solution was used as the template for subsequent cell-free protein synthesis at a concentration of about 10 μg/ml of reaction mixture. Four mutants of ε (A186G, A188G, A200G and A243G) labeled with ^{15}N-Ala and two (T193A, T201A) labeled with ^{15}N-Thr were thus prepared in six parallel reactions in the presence of unlabeled θ.

Genes encoding two additional ε mutants (the double mutant S2A/T3S and the C-terminal deletion mutant ε217) were prepared using PCR reactions with appropriate primer pairs (Table SI) and inserted between the NdeI and EcoRI sites of the T7 expression vector pRSET-5b (36).

The resulting plasmids were used as templates to prepare PCR-amplified DNA products by following the procedure of the second step described above. The resulting purified and reannealed mixtures of PCR products were used directly as templates for cell-free protein synthesis of the mutant ε-subunits in the presence of θ, as above.

NMR spectroscopy

All NMR spectra were recorded at 25°C using a Bruker 800 MHz NMR spectrometer equipped with a cryoprobe. ^{15}N-HSQC spectra were recorded using 5 mm sample tubes (except for samples of the α:ε:θ complex which were concentrated to 200 μl and measured in 3 mm sample tubes), t_{1max} = 32 ms, t_{2max} = 102 ms, and total recording times between 1 h and 13 h.

RESULTS

Cell-free synthesis of full-length ε in the presence of θ and/or α

Cell-free synthesis of full-length ε at 37°C produced similar yields in both the autoinduction system with coexpression of T7 RNA polymerase from plasmid pKO1166 (33) and in our standard system with purified T7 RNA polymerase (28,31). Yields were about 3–4 mg/ml at 37°C and 2 mg/ml at 30°C. At both temperatures, the protein was produced in insoluble form. Similarly, the protein was in inclusion bodies when overexpression was carried out *in vivo* at 30 or 37°C, using standard expression protocols in LB medium (25,37), or at room temperature using the autoinduction medium described by Studier (26).

In an attempt to improve the solubility of full-length ε, its cell-free synthesis was carried out in the presence of either purified α or θ. At 37°C, the presence of α alone did not improve the solubility of ε. Instead, a coprecipitate of the α:ε complex was obtained (Figure 2A), although α alone remained soluble when added to the cell-free reaction mixture. Expression at 30°C produced mainly soluble α:ε complex (data not shown), suggesting that the effect is caused by thermal instability of ε. In contrast, cell-free synthesis of ε in the presence of 0.5–1.0 mg/ml of θ led to a soluble ε:θ complex (Figure 2B) even at 37°C, and the simultaneous presence of α and θ also gave a soluble α:ε:θ core complex (Figure 2C). The yield of expressed ε was unaffected by the presence of θ and/or α.

A similar set of experiments performed with C-terminally truncated ε, ε217, and C-terminally truncated α, α917 (13), also resulted in coprecipitation of ε217 and α917, indicating that the interaction with α does not depend only on the C-terminal 26 residues of ε (Figure S1).

^{15}N-HSQC spectra of ^{15}N-labeled ε in complex with θ

Cell-free synthesized samples of the 36 kDa ε:θ complex, where ε was selectively labeled with ^{15}N-Ala or ^{15}N-Thr, were found to be adequate for the acquisition of ^{15}N-HSQC NMR spectra without chromatographic purification (Figure 3). The cross-peaks of the N-terminal domain of ε were readily assigned by comparison with the reported assignments of ε186 (38,39). Conservation
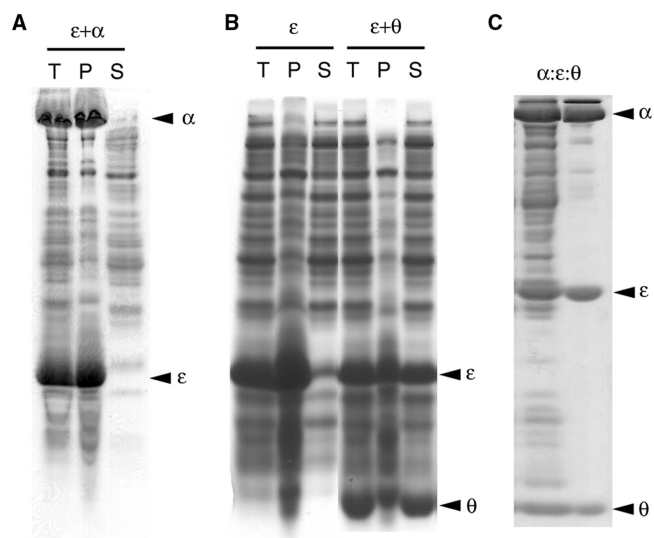
Figure 2. Subunit θ solubilizes nascent ε and α:ε:θ forms a soluble isolable complex when ε is made in the presence of α and θ. The ε-subunit was produced by cell-free synthesis. The gels were stained with Coomassie brilliant blue. T, P and S denote the total reaction mixtures, the insoluble fractions (pellet) and the supernatants, respectively. (**A**) The 10% SDS–PAGE of ε synthesized at 37°C in the presence of 5 mg/ml of α-subunit. Nascent ε coprecipitates α. (**B**) The 15% SDS–PAGE of ε synthesized at 37°C in the presence of 0.1 mg/ml (lanes 1–3, labeled ε) or 1 mg/ml (lanes 4–6, labeled ε + θ) of θ-subunit. (**C**) The 15% SDS–PAGE of ε synthesized at 30°C in the presence of 5 mg/ml of the α-subunit and 0.5 mg/ml of the θ-subunit. The left and right lanes show the soluble fraction of the cell-free reaction mixture and the purified α:ε:θ complex, respectively.

of their chemical shifts indicated that neither the structure of the N-terminal domain nor its interface with θ is significantly affected by the presence of the εCTS, and provide no indication that the εCTS interacts with portions of the folded core of ε186. The additional cross-peaks observed must arise from Ala and Thr residues in the εCTS (Figure 1). Their chemical shifts are characteristic of a random-coil polypeptide chain. Many of them are also more intense than the signals from the globular N-terminal domain but their line widths are not sufficiently narrow to indicate that the segment is completely free in solution. In addition, the ^{15}N-Ala labeled sample displays at least one more cross-peak than expected from the amino acid sequence and peak doubling is observed also in the ^{15}N-Thr labeled sample (see below). This may be explained by a heterogeneous chemical environment due to nonspecific interaction with other components of the cell extract, considering that the cell-free reaction mixture inhibited the binding of the ε:θ complex to α (see below).

^{15}N-HSQC spectra of ^{15}N-labeled θ in the ε:θ complex

The ^{15}N-HSQC spectrum of uniformly labeled θ in complex with full-length ε was indistinguishable from that of θ in complex with ε186 (data not shown). This indicates that θ does not interact with the εCTS in the ε:θ complex.

Analysis of the α:ε:θ complex containing ^{15}N-Thr or ^{15}N-Ala labeled ε

Initially, we attempted to form the α:ε:θ complex by the addition of purified α to the NMR sample containing the
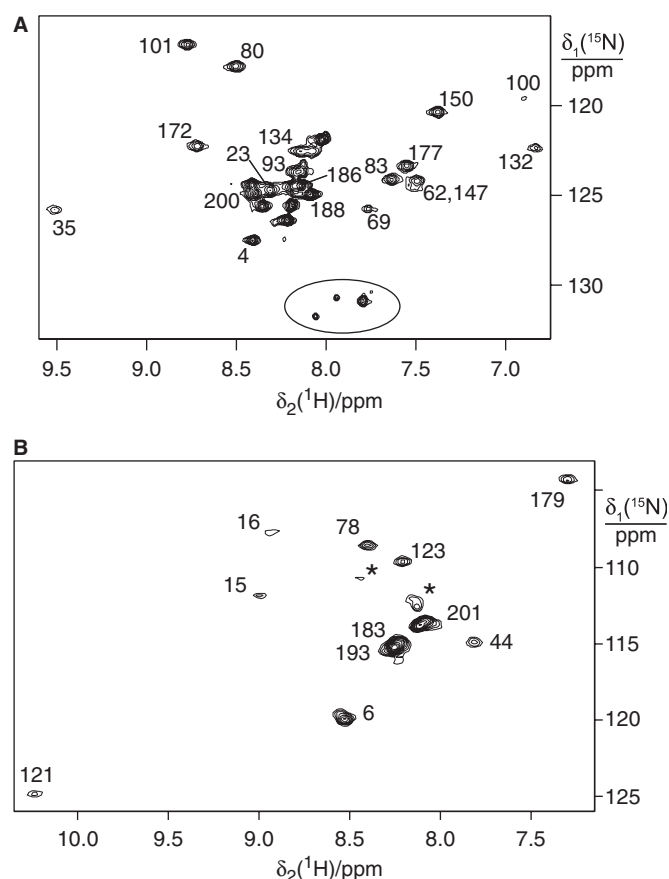

Figure 3. ^{15}N-HSQC NMR spectra of ε in complex with θ. The cross-peaks are assigned with the residue numbers. (**A**) ^{15}N-Ala labeled ε in complex with unlabeled θ. The circle identifies a set of unassigned peaks indicative of sample heterogeneity. They were not observed for the α:ε:θ complex (Figure 5A, spectral region not shown). (**B**) ^{15}N-Thr labeled ε in complex with unlabeled θ. The asterisks label peaks which were not observed in the spectra of ε186 or the α:ε:θ complex, indicating sample heterogeneity.

unpurified ε:θ complex. ^{15}N-HSQC spectra of samples containing ^{15}N-Ala or ^{15}N-Thr labeled ε were, however, indistinguishable from spectra in the absence of α even after addition of α in 2-fold excess. This seemed to be due to nonspecific association of ε with other proteins (or nucleic acids) in the cell-free reaction mixture since after partial purification via a DEAE column, the ε:θ complex readily bound to α (data not shown). Therefore, we prepared the α:ε:θ complex by cell-free synthesis of ε in the presence of purified α and θ. Subsequent purification of the complex was straightforward, requiring only two steps of chromatography (Figure 2C).

The ^{15}N-HSQC spectra of the ternary complex contained only a few cross-peaks (Figure 4A). Since nuclear magnetization relaxes faster with decreasing molecular tumbling rates, the high molecular weight of the complex (about 165 kDa) leads to excessively fast nuclear relaxation for most residues and cross-peaks can be observed only for residues located in polypeptide segments of increased mobility.

We next set out to assign the cross-peaks observed for the α:ε:θ complex containing ^{15}N-Thr labeled ε. The peaks

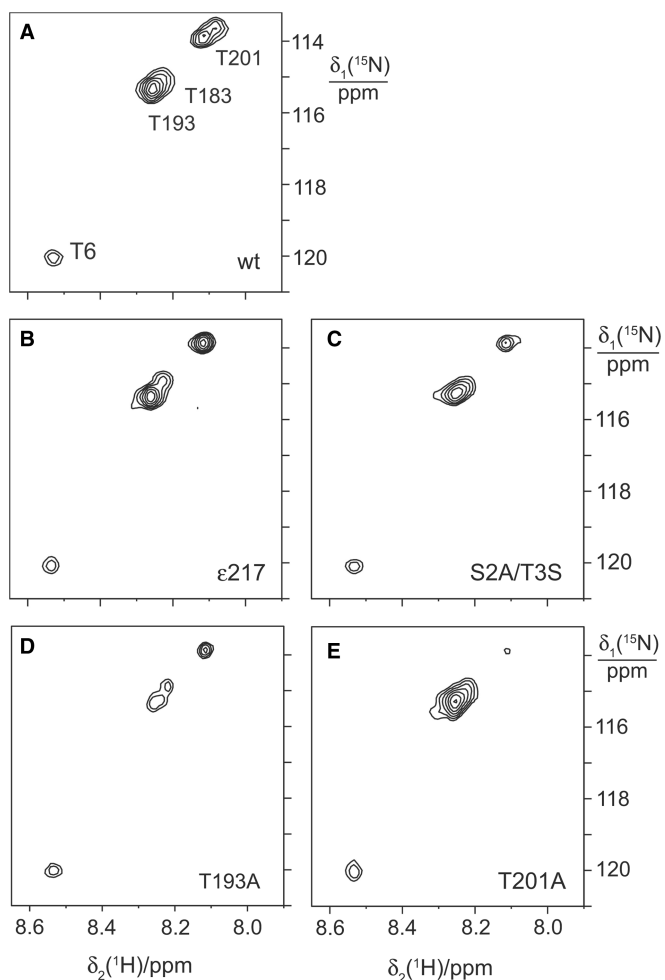**Figure 4.** ^{15}N-HSQC NMR spectra of α:ε:θ and mutant ε:θ complexes containing ^{15}N-threonine labeled ε. For best comparison, the spectra were Fourier transformed with the same parameters and scaled for similar intensity of the cross-peak of Thr6 in all spectra. (**A**) Spectrum of the α:ε:θ complex. (**B–E**) Spectra of samples of the ε:θ complex prepared with mutant ^{15}N-threonine labeled ε. The mutations are indicated.

Figure 5. ^{15}N-HSQC NMR spectra of the α:ε:θ complex and PCS induced by Dy^{3+}, Er^{3+} or Tb^{3+} ions. Resonance assignments are indicated for the cross-peaks observed in the absence of lanthanide ions. Arrows identify PCS induced by paramagnetic lanthanide ions bound to the active site of ε. (**A**) ^{15}N-Alanine labeled ε in complex with α and θ. The spectra in the absence of lanthanide ion (black), with Er^{3+} (blue) and with Dy^{3+} (brown) are superimposed. (**B**) ^{15}N-Threonine labeled ε in complex with α and θ. The spectra in the absence of lanthanide ion (black), with Er^{3+} (blue) and with Tb^{3+} (red) are superimposed.

observed for α:ε:θ coincide with the most intense peaks observed for the ε:θ complex (Figure 3B). Therefore, the assignment could be obtained in the absence of α by analyzing appropriate ε mutants in complex with θ.

The crystal structure of ε186 is missing electron density for the first six and last six residues of the protein (15) and NMR cross-peaks of these residues are narrower than average, indicating increased mobility. (The ^{15}N-HSQC cross-peaks of the three N-terminal residues cannot be observed in ε186 (38,39), presumably due to line broadening arising from hydroxide-catalyzed amide-proton exchange rates near the amino terminus.) Assuming that these segments are also highly mobile in the α:ε:θ complex, the cross-peaks of Ala4, Thr6 and Thr183 would be expected to appear at conserved positions in the ^{15}N-HSQC spectrum of the α:ε:θ complex. This is indeed the case (Figures 4A and 5). The conserved appearance of the spectrum obtained for the ε:θ complex containing the ^{15}N-Thr labeled S2A/T3S double-mutant of ε (Figure 4C)
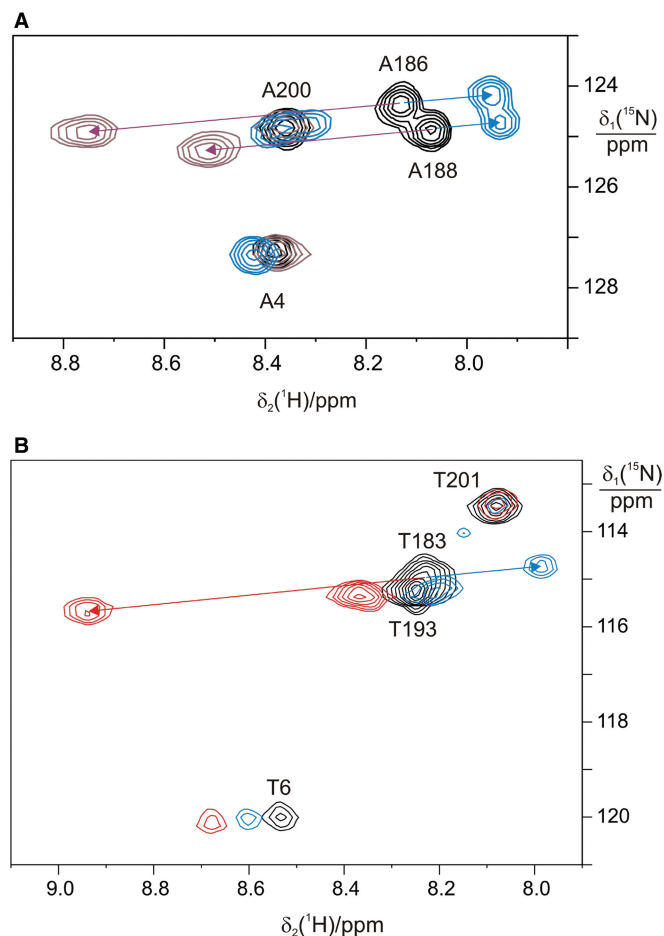
confirms that Thr3 remains unobservable in full-length ε and that the unassigned cross-peaks must arise from Thr residues located in the εCTS.

In order to assign the cross-peaks of these additional Thr residues, we synthesized ε in the presence of θ from DNA that had been mutated and amplified by PCR. Mutagenesis of Thr201 to Ala led to a much weaker cross-peak at the ^{15}N-chemical shift of 114 p.p.m. (Figure 4E). Similarly, mutation of Thr193 led to a much weaker cross-peak at the ^{15}N-chemical shift of 115.3 p.p.m. (Figure 4D). The residual cross-peak intensities arise from the presence of small amounts of wild-type protein due to residual amounts of wild-type plasmid template in the cell-free reactions (35).

The cross-peak of Thr201 appears as two peaks in the spectrum of the wild-type ε:θ complex (Figure 3B) and of the α:ε:θ complex (Figure 4A), but not in the spectra of the mutants. The origin of this peak doubling could be conformational heterogeneity or the inconsistent appearance

of a cross-peak of the only unassigned threonine residue in the εCTS, Thr218. This was examined by truncation of ε following Ala217 in the deletion mutant ε217. The result was inconclusive: the ε217 sample did not show doubling of the cross-peak assigned to Thr201 (Figure 4B), but neither did the T193A mutant (Figure 4D). As expected for a more mobile peptide segment, however, the truncation led to a more intense signal for Thr201.

Four ^{15}N-HSQC cross-peaks could be observed in the α:ε:θ complex prepared with ^{15}N-Ala labeled ε (Figure 5). They were also present at conserved positions in the spectrum of the ε:θ complex (Figure 3A). One of the cross-peaks was assigned to Ala4 by virtue of its similar position in the NMR spectrum of ε186. For additional assignments of the most intense peaks, mutant samples of the ε:θ complex were prepared. Mutation of the C-terminal residue of ε (Ala243) to glycine did not remove any of the intense Ala cross-peaks observed for the wild-type protein (data not shown). In contrast, mutations of Ala200 (Figure 6B) and Ala188 (Figure 6C) enabled straightforward assignment of two of the cross-peaks to these residues. Mutation of Ala186 resulted in disappearance of one cross-peak (assigned to Ala186) and significant change in chemical shift of the cross-peak assigned to Ala188, as might be expected for a next-neighbor residue (Figure 6D). Notably, most of the chemical shifts were insensitive toward mutations in other residues, in agreement with the random-coil nature of the εCTS until at least Thr201.

To verify that these ^{15}N-alanine and ^{15}N-threonine resonance assignments also apply to the α:ε:θ complex, we added paramagnetic lanthanides to Pol III core complexes containing appropriately labeled ε (Figure 5). The ε-subunit has a natural metal-binding site that can bind a single lanthanide ion, resulting in substantial pseudocontact shifts (PCS) (40), which displace each ^{15}N-HSQC cross-peak by a chemical shift increment that is similar in the ^{1}H and ^{15}N dimensions. PCS could indeed be measured in the α:ε:θ complex (Figure 5). The sign and magnitudes of the PCS observed for Ala4, Thr6, Thr183 and Ala186 with Dy^{3+}, Tb^{3+} and Er^{3+} were in agreement with those measured previously for ε186 (39). Along the sequence Ala186, Ala188, Thr193, Ala200 and Thr201, the PCS steadily decreased toward zero. This would be expected for a flexible polypeptide chain for which positive and negative PCS are averaged by extensive sampling of the space around the metal ion.

In summary, we have demonstrated that the Pol III α:ε:θ core complex is characterized by substantially increased mobility of all Ala and Thr residues in the polypeptide segment between Thr183 and Thr201 of ε, whereas residues farther toward the C-terminus are immobilized by binding to α.

Lack of interaction between ε186 and α

The N-terminal domain of ε, ε186, has been shown previously to be incapable of forming an isolable complex with α (21,41). If the ε:θ complex is tethered to α via a long flexible linker involving the C-terminal residues of ε, weak binding of the globular N-terminal domain of ε to
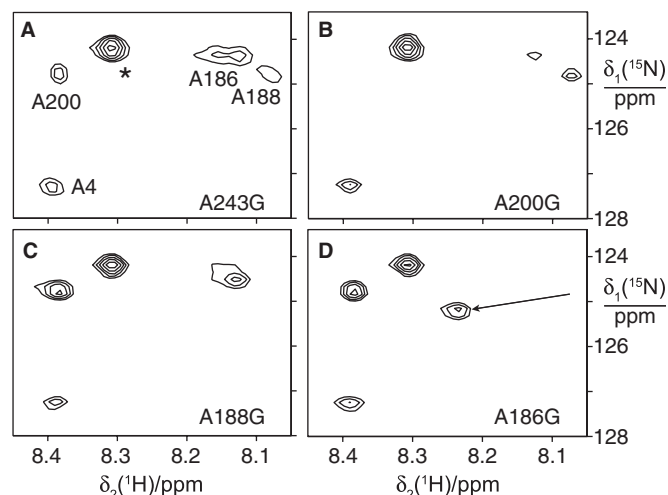

Figure 6. Assignment of ^{15}N-HSQC spectra of ^{15}N-alanine labeled ε:θ complexes by site-directed mutagenesis. ε was labeled with ^{15}N-alanine and θ was unlabeled. The four spectra are of samples prepared with ε where different alanine residues were mutated to glycine. (**A**) A243G mutant. The asterisk identifies a cross-peak that was not observed in the wild-type protein (Figure 3A). It may arise from a low-molecular weight metabolite. (**B**) A200G mutant. (**C**) A188G mutant. (**D**) A186G mutant. The arrow identifies the putative shift of the cross-peak of Ala188 from its position in the wild-type protein.

α may still be sufficient to put the exonuclease domain in a defined location with respect to the polymerase subunit. ^{15}N-HSQC spectra of the ε186:θ complex with uniformly ^{15}N-labeled ε186, however, were not affected by the presence of α (data not shown), indicating that any association between ε186 and α would involve a dissociation constant $K_d > 0.1$ mM. This lower limit for K_d is consistent with a reported estimate from gel filtration data (21).

DISCUSSION

In contrast to the soluble N-terminal domain of the ε-subunit, ε186, overexpression of full-length *dnaQ* leads to insoluble protein which in the past could only be solubilized by a denaturation and refolding protocol (37). Cell-free synthesis circumvents this problem elegantly by allowing the formation of protein–protein complexes from nascently produced ε in the presence of its natural binding partners θ and α. In addition, nascently synthesized ε readily formed the ternary α:ε:θ complex, whereas the ε:θ complex was unable to bind to α in the reaction mixture of cell-free synthesis. This effect appears to be due to an engagement of the εCTS in nonspecific interactions with other components of the cell-free reaction mixture, as partial purification of the ε:θ complex by anion exchange chromatography restored the binding capacity to α. The nonspecific interactions also explain the apparent heterogeneity observed for the NMR signals from the CTS (Figure 3).

Our experiments confirm the role of θ as an important factor for stabilizing the structure of ε. The ε186 domain is prone to irreversible denaturation at 25°C (42), whereas the complex with θ can be studied at 30°C without

degradation of NMR signals over time (38). The temperature sensitivity of the N-terminal domain of ε explains why the presence of θ is required during cell-free synthesis of ε to obtain the full-length ε:θ complex in soluble form. In the absence of θ, cell-free synthesis of ε leads to precipitation even in the presence of α. Precipitation is more pronounced at higher temperatures, in agreement with the sensitivity of the N-terminal domain to denaturation. The fact that α coprecipitates with ε highlights the independent roles of the N- and C-terminal domains of ε, where binding to α via the εCTS is still possible if the globular N-terminal domain unfolds and aggregates.

Whether θ acts as a chaperone for ε *in vivo* is not clear, in view of the low cellular concentration of Pol III subunits (43). In any case, neither θ nor the N-terminal domain of ε seem to act as a specific receptor of the εCTS in the absence of α, as NMR spectra of ε:θ samples containing selectively labeled ε and uniformly labeled θ did not reveal significant differences from corresponding spectra of the ε186:θ complex.

Among the C-terminal 57 residues of ε, the segment between Ala186 and Thr201 is much more mobile than the residues closer to the C-terminus. Judging by the signal intensities of the ^{15}N-Ala residues (Figures 3A and 5A), this holds both for the ε:θ and α:ε:θ complexes. Therefore, the 'sticky' part of the εCTS is limited to residues following Thr201. In the complex with α, the contacts must involve residues close to or even including the C-terminal alanine residue of ε, as no NMR signal could be observed for this residue. This is consistent with properties of the *dnaQ932* mutant allele, which encodes a version of ε lacking only the last three residues. Its recessive phenotype suggests it is incapable of competing with wild-type ε for binding to α (44). Considering that ε217 can coprecipitate a C-terminally truncated form of α, α917 (Supplementary Data), at least some of the residues between Thr201 and Thr218 seem to be involved in binding to α. This indicates a large binding epitope, which would be difficult to explore by site-directed mutagenesis.

Residues 183–201 of ε thus form the core of a highly flexible tether in the α:ε:θ complex. Combined with the absence of detectable interactions between the ε186:θ complex and α, this means that the proofreading exonuclease domain enjoys an extraordinary degree of positional and orientational freedom with respect to the polymerase subunit to which it is tethered.

Similar to ε, the polymerase active site of the α-subunit binds a divalent metal ion (Mg^{2+}) in a similar coordination environment (13,14). It is likely that this ion can also be replaced by a lanthanide ion, but there was no evidence for PCS induced in ε by a paramagnetic lanthanide bound to α in any of our experiments, even though PCS induced by Dy^{3+} can extend over distances >40 Å (45). As all PCS were readily explained as intramolecular effects from a lanthanide bound to the active site of ε (Figure 5), it is unlikely that the proofreading domain is positioned close to a lanthanide ion in the polymerase active site of α in the α:ε:θ complex. We do not know yet whether the presence of primer-template DNA or of any of the other subunits of the polymerase III holoenzyme complex would change this; there has certainly been no indication in published

work of an interaction between ε and any Pol III subunit other than α and θ.

Wieczorek and McHenry (46) reported that a 320-residue N-terminal segment of α binds ε with the same affinity (K_d about 5 nM) as the 1160-residue full-length α-subunit. In the crystal structures of α (13,14), the N-terminal 270 residues fold into a globular (php) domain of about 40 Å diameter. Considering that a 20-residue peptide in extended conformation can span well over 60 Å, the flexible tether by which ε is attached to α could readily bridge the distance between almost any attachment site on the php domain and the polymerase active site.

The α:ε:θ Pol III core complex is distinctly different from proofreading polymerases whose structures have been determined in that the polymerase and exonuclease active sites reside in different subunits. In examples of structures of single-chain proofreading polymerases like DNA polymerase I (47,48) and RB69 DNA polymerase (49), the proofreading domain shares a large and well-defined interface with the polymerase domain, suggesting that both domains are rigidly associated. Repair of a mismatched template-primer DNA end thus involves translocation of the DNA from the polymerase to the exonuclease active sites which are located about 30 Å apart.

Our observations that neither ε186 nor θ interact to any appreciable degree with the εCTS or with α, and that the long linker between the catalytic domain of ε and the C-terminal region that is firmly bound to α is still flexible in the Pol III core complex, suggests a fundamentally different mechanism for transition of Pol III from the polymerization to the editing mode. The transition might involve conformational changes in α that expose a cryptic binding site for the ε186 domain, such that the two active sites are brought closer together, rather than protein-mediated transfer of the DNA substrate as occurs in other enzymes. Since we have shown the exonuclease domain to be relatively freely mobile in the α:ε:θ complex, it might thus be able to swing in closer to the polymerase active site to repair a mismatched terminus when this is required.

CONCLUDING REMARKS

In this work, we have exploited several powerful new techniques for structural biology in solution to study the structure of the Pol III core complex. The assignment of NMR resonances by site-directed mutagenesis has long been used in situations where resonance assignments by other means are difficult (50). As mutations can change the chemical shifts by perturbing the 3D protein structure, this strategy is not always successful. While this problem hardly occurs in random-coil polypeptide segments as in the C-terminal region of ε, chemical shift changes were indeed induced by the mutation of Ala186 (Figure 6D). In this situation, PCS induced by paramagnetic lanthanide ions provided a conclusive second line of evidence. We anticipate that the combination of mutation by PCR, cell-free protein synthesis from PCR-amplified DNA (35) and PCS induced by lanthanide tags (51) will become

an attractive tool set for many hitherto difficult cases like those examined here.

The present study also illustrates the power of cell-free protein synthesis in the presence of separately purified cofactors for the production of soluble protein–protein complexes. In a similar manner, we were able to produce the complex between the χ- and ψ-subunits of *E. coli* Pol III in soluble form by cell-free synthesis of ψ in the presence of separately purified χ (33). In the χ:ψ complex, the N-terminal segment of 26 residues of ψ is unstructured in the complex with χ, but required for binding to the γ-subunit (33). Similarly, the Pol III τ-subunit binds to α via a C-terminal polypeptide segment which is unstructured in pure τ (52,53). The emerging picture of the polymerase III holoenzyme is that of a machine composed of globular building blocks that are linked by flexible tethers.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Kornberg,A and Baker,T.A. (1991) *DNA Replication*. W.H. Freeman & Co., New York, N.Y.
2. Maki,H., Horiuchi,T. and Kornberg,A. (1985) The polymerase subunit of DNA polymerase III of *Escherichia coli*. I. Amplification of the *dnaE* gene product and polymerase activity of the α subunit. *J. Biol. Chem.*, **260**, 12982–12986.
3. Maki,H. and Kornberg,A. (1985) The polymerase subunit of DNA polymerase III of *Escherichia coli*. II. Purification of the α subunit, devoid of nuclease activities. *J. Biol. Chem.*, **260**, 12987–12992.
4. Scheuermann,R.H., Tam,S., Burgers,P.M.J. and Echols,H. (1983) Identification of the ε-subunit of *Escherichia coli* DNA polymerase III holoenzyme as the *dnaQ* gene product: a fidelity subunit for DNA replication. *Proc. Natl Acad. Sci. USA*, **80**, 7085–7089.
5. 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.
6. McHenry,C.S. and Crow,W. (1979) DNA polymerase III of *Escherichia coli*. Purification and identification of subunits. *J. Biol. Chem.*, **254**, 1748–1753.
7. Kim,D.R. and McHenry,C.S. (1996) *In vivo* assembly of overproduced DNA polymerase III. *J. Biol. Chem.*, **271**, 20681–20689.
8. Maki,H. and Kornberg,A. (1987) Proofreading by DNA polymerase III of *Escherichia coli* depends on cooperative interaction of the polymerase and exonuclease subunits. *Proc. Natl Acad. Sci. USA*, **84**, 4389–4392.
9. 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.
10. Taft-Benz,S.A. and Schaaper,R.M. (2004) The θ subunit of *Escherichia coli* DNA polymerase III: a role in stabilizing the ε proofreading subunit. *J. Bacteriol.*, **186**, 2774–2780.
11. Chikova,A.K. and Schaaper,R.M. (2005) The bacteriophage P1 *hot* gene product can substitute for the *Escherichia coli* DNA polymerase III θ subunit. *J. Bacteriol.*, **187**, 5528–5536.
12. Lobocka,M.B., Rose,D., Rusin,M., Samojedny,A., Lehnherr,H., Yarmolinsky,M.B. and Blattner,F.R. (2004) The genome of bacteriophage P1. *J. Bacteriol.*, **186**, 7032–7068.
13. Lamers,M.H., Georgescu,R.E., Lee,S.-G., O'Donnell,M. and Kuriyan,J. (2006) Crystal structure of catalytic α subunit of *E. coli* replicative DNA polymerase III. *Cell*, **126**, 881–892.
14. Bailey,S., Wing,R.A. and Steitz,T.A. (2006) The structure of *T. aquaticus* DNA polymerase III is distinct from eukaryotic replicative DNA polymerases. *Cell*, **126**, 893–904.
15. Hamdan,S., Carr,P.D., Brown,S.E., Ollis,D.L. and Dixon,N.E. (2002) Structural basis for proofreading during replication of the *Escherichia coli* chromosome. *Structure*, **10**, 535–546.
16. DeRose,E.F., Kirby,T.W., Mueller,G.A., Chikova,A.K., Schaaper,R.M. and London,R.E. (2004) Phage like it HOT: solution structure of the bacteriophage P1-encoded HOT protein, a homolog of the θ subunit of *E. coli* DNA polymerase III. *Structure*, **12**, 2221–2231.
17. Kirby,T.W., Harvey,S., DeRose,E.F., Chalov,S., Chikova,A.K., Perrino,F.W., Schaaper,R.M., London,R.E. and Pedersen,L.C. (2006) Structure of the *Escherichia coli* DNA polymerase III ε-HOT proofreading complex. *J. Biol. Chem.*, **281**, 38466–38471.
18. Keniry,M.A., Park,A.Y., Owen,E.A., Hamdan,S.M., Pintacuda,G., Otting,G. and Dixon,N.E. (2006) Structure of the θ subunit of *Escherichia coli* DNA polymerase III in complex with the ε subunit. *J. Bacteriol.*, **188**, 4464–4473.
19. Pintacuda,G., Park,A.Y., Keniry,M.A., Dixon,N.E. and Otting,G. (2006) Lanthanide labeling offers fast NMR approach to 3D structure determinations of protein-protein complexes. *J. Am. Chem. Soc.*, **128**, 3696–3702.
20. 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.
21. 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.
22. Wootton,J.C. and Drummond,M.H. (1989) The Q-linker: a class of interdomain sequences found in bacterial multidomain regulatory proteins. *Protein Eng.*, **2**, 535–543.
23. Wijffels,G., Dalrymple,B.P., Prosselkov,P., Kongsuwan,K., Epa,C., Lilley,P.E., Jergic,S., Buchardt,J., Brown,S.E. and Alewood,P.F. (2004) Inhibition of protein interactions with the β₂ sliding clamp of *Escherichia coli* DNA polymerase III by peptides from β₂-binding proteins. *Biochemistry*, **43**, 5661–5671.
24. Pratt,J.M. (1984) Coupled transcription-translation in prokaryotic cell-free systems. In Hames,B.D. and Higgins,S.J. (eds), *Transcription and Translation*. IRL Press, Oxford, UK, pp. 179–209.
25. Hamdan,S., Bulloch,E.M., Thompson,P.R., Beck,J.L., Yang,J.-Y., Crowther,J.A., Lilley,P.E., Carr,P.D., Ollis,D.L. and Brown,S.E. (2002) Hydrolysis of the 5′-*p*-nitrophenyl ester of TMP by the proofreading exonuclease (ε) subunit of *Escherichia coli* DNA polymerase III. *Biochemistry*, **41**, 5266–5275.
26. Studier,F.W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Express. Purif.*, **41**, 207–234.
27. 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.
28. Ozawa,K., Headlam,M.J., Schaeffer,P.M., Henderson,B.R., Dixon,N.E. and Otting,G. (2004) Optimization of an *Escherichia coli* system for cell-free synthesis of selectively ¹⁵N-labelled proteins for rapid analysis by NMR spectroscopy. *Eur. J. Biochem.*, **271**, 4084–4093.
29. Gill,S.C. and von Hippel,P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.*, **182**, 319–326.

30. Ozawa,K., Dixon,N.E. and Otting,G. (2005) Cell-free synthesis of ^{15}N-labeled proteins for NMR studies. *IUBMB Life*, **57**, 615–622.

31. Apponyi,M., Ozawa,K., Dixon,N.E. and Otting,G. (2008) Cell-free protein synthesis for analysis by NMR spectroscopy. In Kobe,B., Guss,M. and Huber,T. (eds), *Methods in Molecular Biology, Vol. 426, Structural Proteomics: High-Throughput Methods.* Humana Press, Totowa, USA, pp. 257–268.

32. Kigawa,T., Yabuki,T., Yoshida,Y., Tsutsui,M., Ito,Y., Shibata,T. and Yokoyama,S. (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett.*, **442**, 15–19.

33. Ozawa,K., Jergic,S., Crowther,J.A., Thompson,P.R., Wijffels,G., Otting,G. and Dixon,N.E. (2005) Cell-free *in vitro* protein synthesis in an autoinduction system for NMR studies of protein-protein interactions. *J. Biomol. NMR*, **32**, 235–241.

34. Bradford,M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248–254.

35. Wu,P.S.C., Ozawa,K., Lim,S.P., Vasudevan,S.G., Dixon,N.E. and Otting,G. (2007) Cell-free transcription/translation from PCR amplified DNA for high-throughput NMR studies. *Angew. Chemie Int. Ed.*, **46**, 3356–3358.

36. Schoepfer,R. (1993) The pRSET family of T7 promoter expression vectors for *Escherichia coli. Gene*, **124**, 83–85.

37. Scheuermann,R.H. and Echols,H. (1984) A separate editing exonuclease for DNA replication: the subunit ε of *Escherichia coli* DNA polymerase III holoenzyme. *Proc. Natl Acad. Sci. USA*, **81**, 7747–7751.

38. DeRose,E.F., Darden,T., Harvey,S., Gabel,S., Perrino,F.W., Schaaper,R.M. and London,R.E. (2003) Elucidation of ε−θ subunit interface of *Escherichia coli* DNA polymerase III by NMR spectroscopy. *Biochemistry*, **42**, 3635–3644.

39. Schmitz,C., John,M., Park,A.Y., Dixon,N.E., Otting,G., Pintacuda,G. and Huber,T. (2006) Efficient χ-tensor determination and NH assignment of paramagnetic proteins. *J. Biomol. NMR*, **35**, 79–87.

40. Pintacuda,G., Keniry,M.A., Huber,T., Park,A.Y., Dixon,N.E. and Otting,G. (2004) Fast structure-based assignment of ^{15}N HSQC spectra of selectively ^{15}N-labeled paramagnetic proteins. *J. Am. Chem. Soc.*, **126**, 2963–2970.

41. Lehtinen,D.A. and Perrino,F.W. (2004) Dysfunctional proofreading in the *Escherichia coli* DNA polymerase III core. *Biochem. J.*, **384**, 337–348.

42. Hamdan,S., Brown,S.E., Thompson,P.R., Yang,J.-Y., Carr,P.D., Ollis,D.L., Otting,G. and Dixon,N.E. (2000) Preliminary X-ray crystallographic and NMR studies on the exonuclease domain of the ε subunit of *Escherichia coli* DNA polymerase III. *J. Struct. Biol.*, **131**, 164–169.

43. Ishihama,Y., Schmidt,T., Rappsilber,J., Mann,M., Hartl,F.U., Kerner,M.J. and Frishman,D. (2008) Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genomics*, **9**, 102.

44. 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.

45. Allegrozzi,M., Bertini,I., Janik,M.B.L., Lee,Y.-M., Liu,G. and Luchinat,C. (2000) Lanthanide-induced pseudocontact shifts for solution structure refinements of macromolecules in shells up to 40 Å from the metal ion. *J. Am. Chem. Soc.*, **122**, 4154–4161.

46. Wieczorek,A. and McHenry,C.S. (2006) The NH_2-terminal php domain of α subunit of the *Escherichia coli* replicase binds the ε proofreading subunit. *J. Biol. Chem.*, **281**, 12561–12567.

47. Brautigam,C.A. and Steitz,T.A. (1998) Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.*, **8**, 54–63.

48. Steitz,T.A. (1999) DNA polymerases: structural diversity and common mechanism. *J. Biol. Chem.*, **274**, 17395–17398.

49. Franklin,M.C., Wang,J. and Steitz,T.A. (2001) Structure of the replicating complex of a Pol α family DNA polymerase. *Cell*, **105**, 657–667.

50. Falzone,C.J., Benkovic,S.J. and Wright,P.E. (1990) Partial proton NMR assignments of the *Escherichia coli* dihydrofolate reductase complex with folate: evidence for a unique conformation of bound folate. *Biochemistry*, **29**, 9667–9677.

51. Pintacuda,G., John,M., Su,X.-C. and Otting,G. (2007) NMR structure determination of protein-ligand complexes by lanthanide labeling. *Acc. Chem. Res.*, **40**, 206–212.

52. Jergic,S., Ozawa,K., Williams,N.K., Su,X.-C., Scott,D.D., Hamdan,S.M., Crowther,J.A., Otting,G. and Dixon,N.E. (2007) The unstructured C-terminus of the τ subunit of *Escherichia coli* DNA polymerase III holoenzyme is the site of interaction with the α subunit. *Nucleic Acids Res.*, **35**, 2813–2824.

53. Su,X.-C., Jergic,S., Keniry,M.A., Dixon,N.E. and Otting,G. (2007) Solution structure of domains IVa and V of the τ subunit of *Escherichia coli* DNA polymerase III and interaction with the α subunit. *Nucleic Acids Res.*, **35**, 2825–2832.