Published online 22 April 2007

Nucleic Acids Research, 2007, Vol. 35, No. 9 ***2825–2832***
doi:10.1093/nar/gkm080

Solution structure of Domains IVa and V of the τ subunit of *Escherichia coli* DNA polymerase III and interaction with the α subunit

Xun-Cheng Su, Slobodan Jergic, Max A. Keniry, Nicholas E. Dixon and Gottfried Otting*

Research School of Chemistry, Australian National University, Canberra ACT 0200, Australia

Received October 19, 2006; Revised January 1, 2007; Accepted January 26, 2007

ABSTRACT

The solution structure of the C-terminal Domain V of the τ subunit of *E. coli* DNA polymerase III was determined by nuclear magnetic resonance (NMR) spectroscopy. The fold is unique to τ subunits. Amino acid sequence conservation is pronounced for hydrophobic residues that form the structural core of the protein, indicating that the fold is representative for τ subunits from a wide range of different bacteria. The interaction between the polymerase subunits τ and α was studied by NMR experiments where α was incubated with full-length C-terminal domain (τ_C16), and domains shortened at the C-terminus by 11 and 18 residues, respectively. The only interacting residues were found in the C-terminal 30-residue segment of τ, most of which is structurally disordered in free τ_C16. Since the N- and C-termini of the structured core of τ_C16 are located close to each other, this limits the possible distance between α and the pentameric $\delta\tau_2\gamma\delta'$ clamp–loader complex and, hence, between the two α subunits involved in leading- and lagging-strand DNA synthesis. Analysis of an N-terminally extended construct (τ_C22) showed that τ_C14 presents the only part of Domains IVa and V of τ which comprises a globular fold in the absence of other interaction partners.

INTRODUCTION

DNA polymerase III (Pol III) is the central enzyme of the *Escherichia coli* replisome. The holoenzyme is composed of several individual proteins, including the subunits α, β, γ, δ, δ′, ε, χ, ψ, τ and θ. 3D structures have been determined for a large fragment of α (1), β (2), the structured parts of ε (3), θ (4) and δ′ (5) and for multimeric complexes including a pentameric $\gamma_3\delta\delta'$ complex (6,7), and the βδ (8), χψ (9) and εθ (10) complexes. To date, no 3D structure of full-length α or τ has been published, though the complete α subunit could presumably be modeled on the corresponding structure from *Thermus aquaticus* (11).

The full-length τ subunit is a 71 kDa protein of 643 amino acid residues. The N-terminal 430 residues of τ (Domains I–III) are identical to γ (12,13) and can substitute for γ in the pentameric complexes with δ and δ′ (14,15). The 213 additional residues at the C-terminus of τ (Domains IVa + V, residues 430–643) have been shown to bind to α with nanomolar affinity (16,17) and to bind to the DnaB helicase, albeit much more weakly (18–20). In addition, a construct comprising Domains IVa and V has been shown to bind to primed DNA (21). The interaction site with α has been mapped to the 147 C-terminal residues of τ (Domain V, residues 499–643; Figure 1) (22) and the binding to DnaB was assigned to residues 430–496 (Domain IVa) (23). Weak DNA-binding activity has also been mapped to Domain IVa (24). The replisome contains two coupled α subunits, one for leading- and one for lagging-strand DNA synthesis (16,25). These two subunits are each linked to the same pentameric clamp–loader complex composed of γ, δ, δ′ and two τ subunits (14). The interaction of τ with α thus presents a critical structural link in the replisome, tying two α subunits to the clamp–loader complex. Here we report the 3D solution structure of the C-terminal Domain V of τ from *E. coli*. In addition, the structure of Domain IVa was probed in a construct comprising Domain IVa and the globular part of Domain V, and the interaction site of α on τ was mapped by NMR experiments using three different constructs of Domain V of τ. The results place restraints on the distances in the replisome between the two α subunits, their proximity to the DnaB helicase, and the flexibility of the complex of τ, α and DnaB.

*To whom correspondence should be addressed. Tel: +61-2-61256507; Fax: +61-2-61250750; Email: gottfried.otting@anu.edu.au
Present address:
Nicholas E. Dixon, Department of Chemistry, University of Wollongong, NSW 2522, Australia.

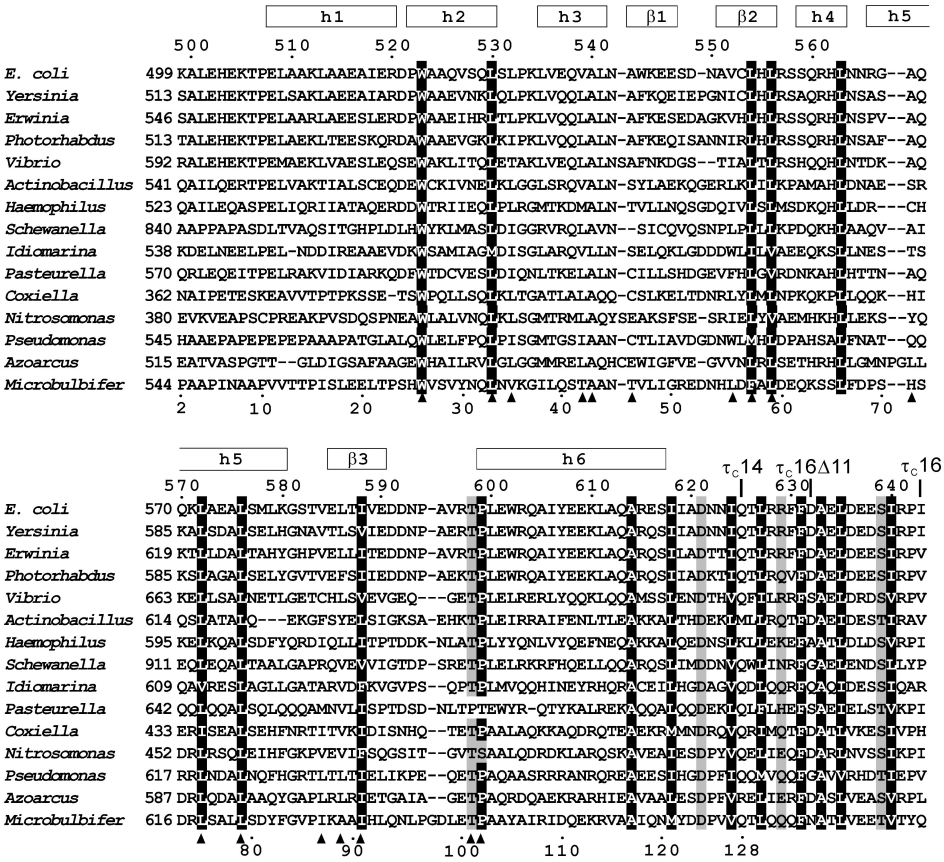

Figure 1. Sequence alignment of τ_C16 with τ proteins from different bacteria. The sequence of τ_C16 comprises the 145 C-terminal residues of the τ subunit of *E. coli* DNA Pol III holoenzyme (residues 499–643 which constitute Domain V). Its sequence numbering is shown at the top. The constructs τ_C14, $\tau_C16\Delta11$ and τ_C16 share the same N-terminus. Their C-termini are indicated above the sequence. The sequence numbering shown below the alignment is that of the NMR sample of τ_C14 included for reference with previous work (41). The regular secondary structure elements of τ_C14 determined in this work are shown at the top. Positions with conserved hydrophobic residues are indicated by white characters on a black background. Conserved hydrophilic residues are shaded gray. Filled triangles below the alignment identify residues with side chains buried in the τ_C14 structure (<5% solvent accessibility). The following sequences from τ proteins are shown (abbreviation, organism, GenBank number): *E. coli*, *Escherichia coli*, 43320; *Yersinia*, *Yersinia pseudotuberculosis*, 51595342; *Erwinia*, *Erwinia carotovora* subsp. *Atroseptica*, 49610641; *Photorhabdus*, *Photorhabdus lummines* subsp. *Laumondii tto1*, 37527701; *Vibrio*, *Vibrio vulnificus cmcp6*, 27361492; *Actinobacillus*, *Actinobacillus pleuropneumoniae serovar 1 STR*, 32035419; *Haemophilus*, *Haemophilus somnus*, 32029123; *Shewanella*, *Shewanella oneidensis MR-1*, 24373576; *Idiomarina*, *Idiomarina loihiensis L2TR*, 56460949; *Pasteurella*, *Pasteurella multocida* subsp. *multocida str. Pm70*, 12720609; *Coxiella*, *Coxiella burnetii RSA 493*, 29541263; *Nitrosomonsa*, *Nitrosomonas europaea ATCC 19718*, 30138336; *Pseudomonas*, *Pseudomonas fluorescens PfO-1*, 48731928; *Azoarcus*, *Azoarcus* sp. *EbN1*, 56476074; *Microbulbifer*, *Microbulbifer degradans 2-40*, 48862900.

MATERIALS AND METHODS

Proteins

The α subunit of *E. coli* DNA Pol III was prepared as described earlier (26). The proteins τ_C16 (N-terminal Met followed by C-terminal residues 499–643 of τ, comprising all of Domain V), $\tau_C16\Delta11$ (τ_C16 with the C-terminal 11 residues removed), τ_C14 (τ_C16 with the C-terminal 18 residues removed) and τ_C22 (N-terminal Met followed by residues 430–626 of τ, corresponding t Domain IVa and τ_C14) were isolated as described in the accompanying paper (24). Stable-isotope enriched samples were produced by expression on a minimal medium (M9) containing $^{15}NH_4Cl$ as the sole nitrogen source or on a $^{13}C/^{15}N$-enriched medium (Silantes, Germany). In the case of τ_C14 the respective yields of purified protein were 9 and 21 mg/l of medium. In the case of τ_C22, the yield was 15 mg of purified ^{15}N-labeled protein per liter of medium.

NMR sample preparation

All samples for NMR measurements were prepared in a buffer containing 10 mM sodium phosphate (pH 6.8), 100 mM NaCl, 1 mM dithiothreitol and 0.1 mM NaN_3, with protein concentrations of up to 1.9 mM. Samples were prepared in 90% H_2O/10% D_2O or 100% D_2O.

NMR measurements

All NMR measurements for the structure determination of τ_C14 were carried out at 30°C, using Varian Inova 600 and Bruker AV 800 NMR spectrometers. The backbone resonance assignment was obtained from the analysis of 3D HNCACB and CBCA(CO)NH together with NOESY-^{15}N-HSQC and 2D NOESY spectra. The side-chain resonances were assigned through the analysis of 3D CC(CO)NH, HNHA, (H)CCH-TOCSY and TOCSY-^{15}N-HSQC and HNCO spectra, together with NOESY-^{15}N-HSQC, ^{13}C-HSQC-NOESY and 2D

NOESY and TOCSY spectra. The protons of aromatic groups were assigned with 2D DQF-COSY, NOESY and TOCSY spectra. A few resonances could be assigned only with the help of NOEs after initial structure calculations. ϕ and Ψ angle restraints were generated by the program TALOS (27) using the chemical shifts of CA, CB, CO, HA and N, and checked against the intensities of intraresidual and sequential NOEs between backbone protons.

In total, 2500 NOE cross-peaks were assigned and integrated, resulting in 1891 meaningful distance restraints. Most of these restraints came from a 2D NOESY spectrum of an unlabeled sample recorded with 40-ms mixing time, using $t_{1max} = 78$ ms, $t_{2max} = 213$ ms and ~3 days of recording time. Stereo-specific resonance assignments were obtained with the program GLOMSA (28), using $^3J(H^\alpha,H^\beta)$ coupling constant information derived from the DQF-COSY spectrum. Stereo-specific assignments of the isopropyl CH_3-groups of valine and leucine residues were achieved by recording constant-time ^{13}C-HSQC spectra of a 10% biosynthetically fractionally ^{13}C-labeled sample (29). Pairs of methyl groups from the same isopropyl group were identified by methyl–methyl NOEs observed in a diagonal-suppressed constant-time ^{13}C-HSQC spectrum with NOE relay (30) recorded of uniformly ^{13}C-labeled protein.

^{15}N-relaxation data of τ_C22 were recorded at 25°C using standard experiments (31). The relaxation delays used were 9, 17, 26, 43, 60, 78, 95 and 112 ms in the R_2 experiment and 3, 30, 100, 200, 350, 600, 900 and 1300 ms in the R_1 experiment. The complete set of R_1 and R_2 relaxation data was recorded in about 18 h per protein sample, using $t_{1max} = 50$ ms and $t_{2max} = 100$ ms for each spectrum.

Structure calculations

The cross-peaks in the NOESY spectra of τ_C14 were assigned and integrated using the program XEASY (32). The NMR structure was calculated using the program DYANA (28), starting from 200 random conformers that were annealed in 40 000 steps using torsion-angle dynamics. The 20 conformers with the lowest residual restraint violations were energy minimized in a shell of water using the program OPAL with standard parameters (33).

Table 1 shows an overview of the restraints used and structural statistics. The Ramachandran plot was analyzed using PROCHECK-NMR (34). No residue was consistently in a forbidden region in all 20 conformers.

Secondary structure elements and root mean square deviation (r.m.s.d.) values were calculated using the program MOLMOL (35). Side-chain solvent accessibilities were measured with a spherical probe of 1.4 Å radius and calculated in percent of the accessibilities measured for a fully extended side chain of residue X in a helical Gly-X-Gly peptide (36). The values obtained were averaged over the 20 NMR conformers.

The chemical shifts and coordinates of the structure have been deposited in the BioMagResBank (accession code 6869) and PDB (accession code 2AYA).

Table 1. Structural statistics for the NMR conformers of τ_C14

Parameter	Value
Number of assigned NOE cross-peaks	2500
Number of non-redundant NOE upper-distance limits	1891
Number of scalar coupling constants (H^α–H^β)[a]	64
Number of dihedral-angle restraints	241
Intra-protein AMBER energy (kcal/mol)	-5388 ± 321
Maximum NOE-restraint violations (Å)	0.09 ± 0.00
Maximum dihedral-angle restraint violations (°)	2.0 ± 0.5
r.m.s.d. to the mean for N, C^α and C′ (Å)[b]	0.55 ± 0.14
r.m.s.d. to the mean for all heavy atoms (Å)[b]	0.99 ± 0.11
Ramachandran plot appearance[c]	
Most favored regions (%)	90.2
Additionally allowed regions (%)	9.5
Generously allowed regions (%)	0.3
Disallowed regions (%)	0.0

[a]Stereo-specific resonance assignments were obtained for 49 pairs of $C^\beta H_2$, one pair of $C^\gamma H_2$ and two pairs of $C^\delta H_2$ protons. [b]For residues 509–609. [c]From PROCHECK-NMR (34).

RESULTS

NMR resonance assignments

The ^{15}N-HSQC spectrum of τ_C14 presented well-dispersed resonances characteristic of a folded protein, with line widths consistent with it being monomeric (Figure 2). Virtually complete resonance assignments were obtained for the polypeptide backbone between residues 500 and 625, except that the amide cross-peak of His562 was missing and the amide cross-peaks of Lys534 and Arg566 were very broad, so that no NOEs could be observed with these H^N resonances. ^{15}N-relaxation measurements of τ_C16Δ11 indicated that the loop region with Lys534 was affected by chemical exchange (37), while the line broadening of Arg566 may have been due to fast amide proton exchange. A constant-time ^{13}C-HSQC spectrum with NOE relay and suppression of the diagonal peaks proved particularly useful for the identification of the ^{13}C–^{1}H methyl cross-peaks belonging to the same isopropyl group, since constant-time ^{13}C-HSQC spectra can be recorded with much higher resolution than HCCH-TOCSY experiments.

Structure of τ_C14 and comparison with the KH fold

The structure of τ_C14 comprises six helices and a three-stranded mixed parallel/antiparallel β-sheet (Figures 1 and 3). The amino- and carboxyl-terminal helices point away from the structural core of the protein (Figure 3). Their orientation towards the terminal ends becomes increasingly uncertain, as the NMR structure of this part of the protein is based exclusively on short-range distance restraints.

A search of the protein data bank with the program DALI (38) revealed only remote structural homologues. The best match was with the ribosome-binding factor A from *Mycoplasma pneumoniae* (MNP156), with a r.m.s.d. of 2.5 Å for 66 superimposed residues. This protein, like the next best matches detected by DALI, assumes a KH type-II fold (39). The KH fold does not include helices corresponding to helices 1 and 6 in τ_C14. Furthermore, the KH fold includes a helix-turn-helix motif (identified by

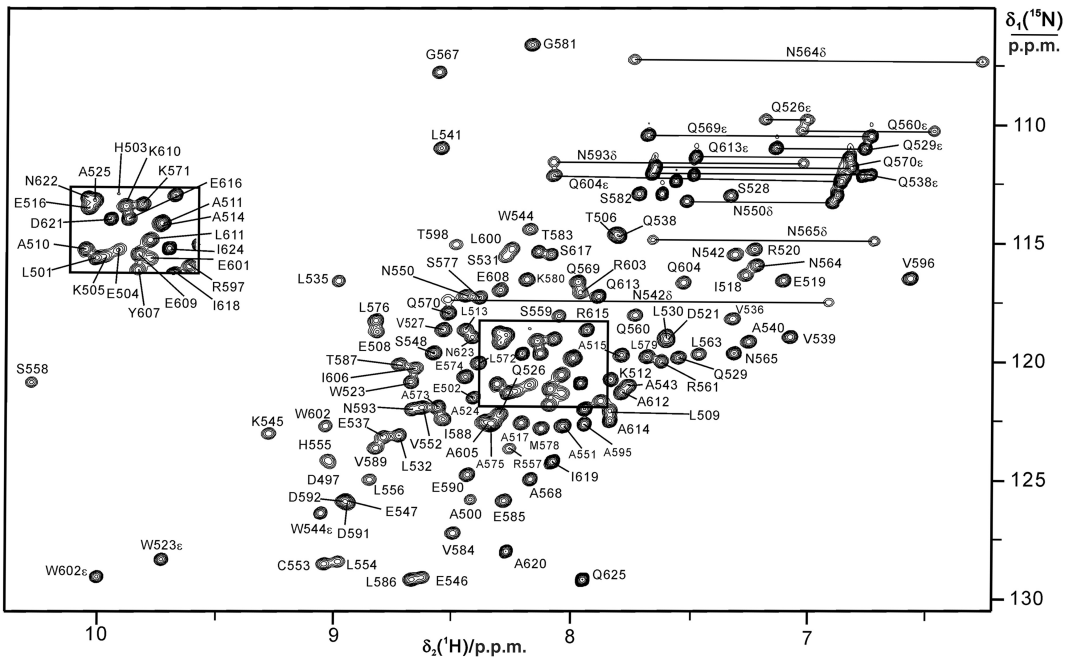

Figure 2. ^{15}N-HSQC spectrum of τ_C14. The spectrum was recorded at 30°C with a 1.8 mM solution of uniformly ^{15}N/^{13}C-labeled τ_C14 in 90% H_2O/10% D_2O at pH 6.8 in phosphate buffer (10 mM phosphate, 100 mM NaCl, 1.0 mM dithiothreitol and 0.1 mM NaN_3). The cross-peaks are assigned using one-letter amino acid symbols and residue numbers. Side-chain resonances are identified by Greek characters. Cross-peaks from side-chain amides are connected by horizontal lines.

an arrow in Figure 3B) that is involved in RNA binding and contains a characteristic GxxG turn signature (39). The τ_C14 structure does not feature a similar motif. In particular, helix 4 is a short 3_{10} helix with clear $H^{\alpha}(i)-H^{N}(i+2)$ NOEs rather than a long α-helix as in the KH fold. There is also no evidence for an extended, positively charged surface in this or any other region of the protein which could favor binding to the phosphodiester backbone of RNA or DNA.

The fold of τ_C14 includes N- and C-terminal helices that protrude from the globular core of the structure into the solvent (Figure 3). These helices appear to be more than mere extensions of the KH fold since they are attached to the rest of the protein in a very specific manner: helix 6 partly extends the helix dipole of helix 3 and helices 1 and 6 form intimate contacts that define their relative orientation. For example, ring currents from the side chain of Trp602 in helix 6 cause significant up-field shifts of the resonances of Arg520 $C^{\beta}H_2$ in helix 1.

Helices 1 and 6 border flexible polypeptide segments near the N- and C-termini, respectively. The virtual absence of medium-range NOEs for the residues preceding Pro507 indicated that these residues are structurally disordered. Also the C-terminal end of helix 6 is conformationally unstable, as the mobility of the polypeptide chain gradually increases towards the C-terminus, as evidenced by the observation of exchange peaks between the water resonance and the backbone amide protons for Ala612 and the following residues in NOESY and TOCSY experiments. This conclusion was further supported by the observation of increasing signal intensities for Ala612 and subsequent residues in experiments with long periods of T_2 relaxation. Helix 6 seems to end at residue Ser617, as the chemical shifts of Ile618 were indicative of an extended backbone conformation and the chemical shifts and NOEs of the following residues were characteristic of random coil conformations.

Conservation of the domain fold

A search for homologous proteins using BLAST (40) exclusively revealed τ domains from other bacterial species (Figure 1). Our data show that all these domains are structurally conserved and closely related to the structure of *E. coli* τ_C14. First, any insertions and deletions occur exclusively in the loop regions between regular secondary structure elements. Second, almost all of the strictly conserved hydrophobic residues (highlighted in Figure 3D for residues between Leu509 and Glu609) are buried in the 3D structure with very low solvent accessibility, as expected for residues that are critical to maintain the fold of the protein. The only exception is Leu564, the side chain of which is somewhat solvent exposed (about 20%). Finally, residues with buried side chains in the structure of τ_C14 are consistently uncharged in the sequence alignment of Figure 1. The only exception is Arg925 in the *Shewanella* protein located at the position of Val584 in *E. coli* τ_C14. Since the side chain of arginine is longer than that of valine, its charge is quite possibly solvent exposed in the *Shewanella* protein.

Structure of τ_C16 (Domain V)

Many conserved residues are located after residue 613 (Figure 1). All those present in the NMR structure of *E. coli* τ_C14 are highly solvent exposed, indicating that their conservation is of functional importance. Samples

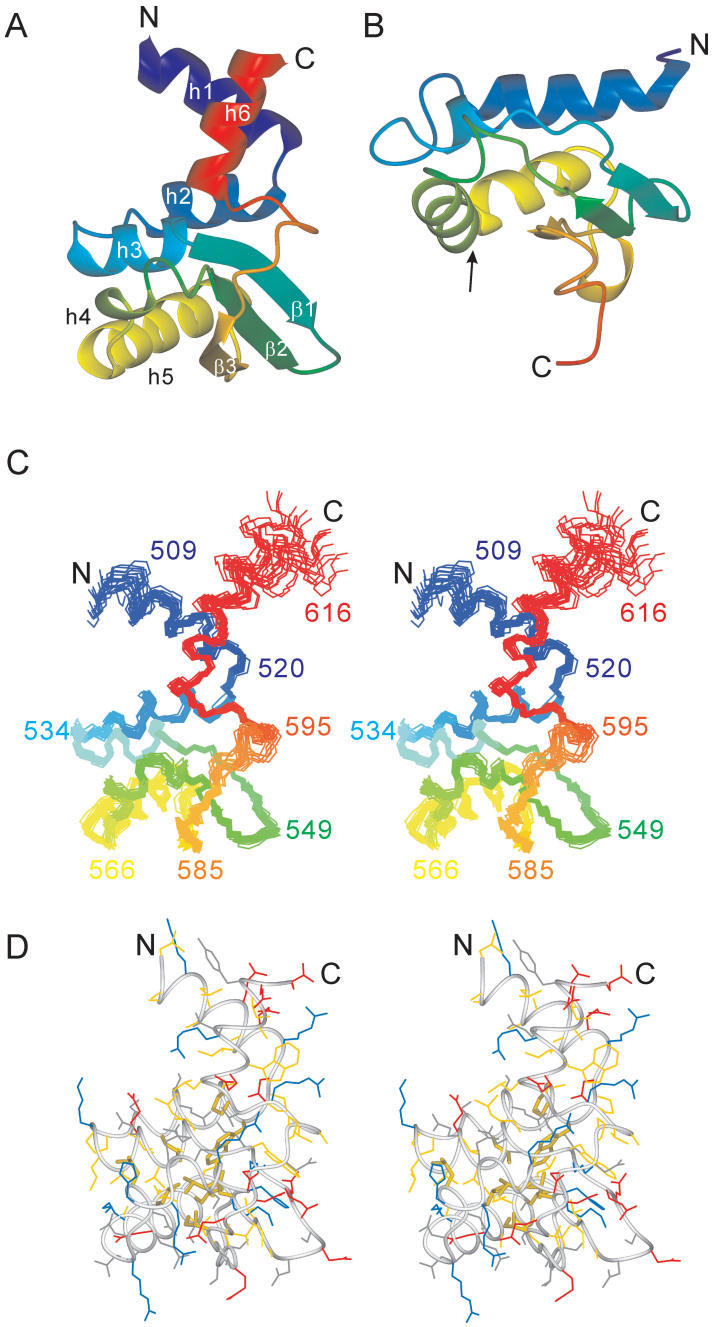

Figure 3. Solution structure of τ_C14 and comparison with the KH domain of protein MPN156. (**A**) Ribbon representation of τ_C14, including residues 509–609. The secondary structure elements are labeled as in Figure 1. (**B**) Ribbon representation of MPN156 from *Mycoplasma pneumoniae* (PDB code: 1PA4). The arrow identifies the RNA-binding helix-loop-helix motif of the KH-domain type II fold. (**C**) Stereo view of a superposition of the backbone atoms of 20 NMR conformers, including residues 507–617. (**D**) Stereo view of a heavy-atom representation of the conformer closest to the mean structure of τ_C14 (only residues 509–609 are shown). The side-chains are color coded in blue (Lys, Arg, His), red (Asp, Glu), yellow (Ala, Cys, Ile, Leu, Met, Phe, Pro, Trp, Val) and gray (Asn, Gln, Ser, Thr, Tyr). The side chains of the strictly conserved hydrophobic residues Trp523, Leu530, Leu554, Leu556, Leu563, Leu572, Leu576, Ile588 and Pro599 are highlighted with broader lines. This figure was prepared using the program MOLMOL (35).

comprising the entire Domain V of τ, τ_C16, were highly sensitive to proteolytic cleavage (24). The ^{15}N-HSQC spectra of ^{15}N-labeled τ_C16 samples that had been freshly prepared in a cell-free protein synthesis system (41) showed increasingly narrow and more intense signals for the segment following Ser617, indicating substantially increased mobility. In particular, there was no evidence for any ordered association of this unstructured segment with the folded part of τ_C14 (41).

Interaction with the α subunit and relaxation data analysis

The interaction of Domain V of τ with the α subunit of Pol III was probed by monitoring changes observed in

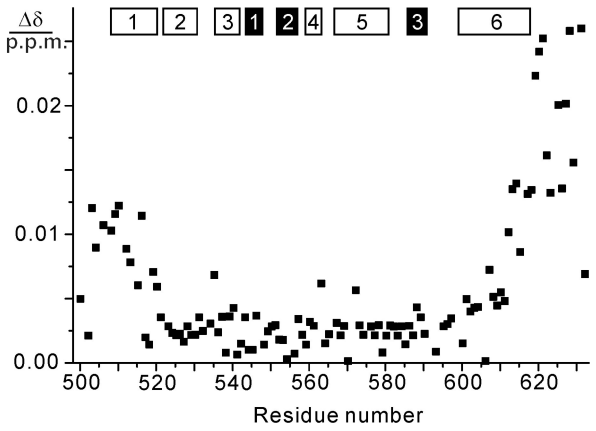

Figure 4. Changes in amide chemical shifts of ^{15}N-labeled $\tau_C16\Delta11$ observed upon binding of α at pH 6.8 and 25°C. The data were recorded with concentrations of $\tau_C16\Delta11$ and α of 90 and 120 μM, respectively. The chemical shift change Δδ was calculated as $(\{[\Delta\delta(^1H)]^2+[0.1\Delta\delta(^{15}N)]^2\}/2)^{1/2}$, where $\Delta\delta(^1H)$ and $\Delta\delta(^{15}N)$ denote the chemical shift changes in p.p.m. observed in the ^{1}H and ^{15}N dimensions of the ^{15}N-HSQC spectrum. The locations of the α-helices and β-strands in τ_C14 are indicated by open and filled bars, respectively.

the ^{15}N-HSQC spectra of ^{15}N-labeled τ_C14, τ_C16 and $\tau_C16\Delta11$ upon addition of purified α. No spectral changes were observed at all when α was added to τ_C14, indicating no binding (data not shown). This is consistent with gel filtration and surface plasmon resonance (SPR) data that also showed that τ_C14 does not interact with α (24). Upon addition of increasing amounts of α, the ^{15}N-HSQC spectrum of τ_C16 became weaker without changing its appearance (data not shown). This indicated tight binding, where only the spectrum of free τ_C16 is observed because the signals of bound τ_C16 are broadened beyond detection. Therefore, the last 18 residues of τ present in τ_C16 but not τ_C14 are critical for the α–τ interaction. The results are in accord with the gel filtration and SPR data (24).

In contrast, the previous study showed that removal of eleven residues from the C-terminus of τ_C16 produced a protein ($\tau_C16\Delta11$) that could no longer be observed by SPR to interact with α, indicating $K_D > 10\,\mu M$ for the α–$\tau_C16\Delta11$ complex (24). However, at the higher concentrations used for NMR experiments, $\tau_C16\Delta11$ could be observed to bind to α with a K_D of about 0.9 mM, resulting in an observable NMR spectrum of $\tau_C16\Delta11$ even in the presence of an excess of α (37). This indicates the existence of fast exchange between free and bound $\tau_C16\Delta11$. Averaging of the NMR signals between free and bound $\tau_C16\Delta11$ was further indicated by the fact that the signals shifted continuously for different ratios of $\tau_C16\Delta11$ to α.

As expected for weak binding, only small chemical shift changes were observed, most of which were within the line width of the cross-peaks (Figure 4). These changes were predominantly observed for the C-terminal 20 residues of $\tau_C16\Delta11$ including part of helix 6, but also among the N-terminal residues including helix 1. Extensive line broadening was observed only for the residues following Gln613. Notably, Ala614 is one of the few totally conserved residues in τ_C16 in the alignment of Figure 1. These data confirm the importance of the conserved C-terminal residues for binding and suggest that the chemical shift changes observed in helix 1 may be caused by small structural adjustments in helix 6 (since the two helices are close together, as shown in Figure 3A) rather than by direct contacts with α.

Binding of $\tau_C16\Delta11$ to α was also manifested by exchange broadening of the ^{15}N-HSQC cross-peaks of the residues following Ile619. In contrast, the presence of α caused no exchange broadening for the N-terminal residues of $\tau_C16\Delta11$. In summary, these data stress the importance of the C-terminal unstructured segment of τ and of the C-terminal residues of helix 6 of Domain V for binding to α, whereas no other parts of τ_C16 seem to contact α.

Structure of τ_C22 (Domain IVa + V)

The 69 additional residues in τ_C22 N-terminal of τ_C14, which represent Domain IVa, were found to assume no defined conformation. ^{15}N-HSQC cross-peaks of τ_C14 superimposed perfectly with ^{15}N-HSQC cross peaks of τ_C22 (Supplementary Data, Figure S1), showing that the additional residues present in τ_C22 did not affect the structure of τ_C14 or interact with it in any specific manner. In agreement with a random coil conformation of Domain IVa, the additional ^{15}N-HSQC cross-peaks present in the spectrum of τ_C22 were observed at ^{1}H chemical shifts characteristic of random coil conformations. Finally, uniformly large $R_1(^{15}N)$ and small $R_2(^{15}N)$ relaxation rates were measured for Domain IVa, and the averaged R_2/R_1 ratio was 5.4 for Domain IVa and 19.8 for the τ_C14 domain in the τ_C22 construct, indicating that all residues of Domain IVa were highly mobile (Supplementary Data, Figure S2).

DISCUSSION

Escherichia coli τ_C14 is the first example of a C-terminal domain from a τ subunit of Pol III holoenzyme from any organism for which a 3D structure has been determined. As demonstrated by amino acid sequence conservation of buried hydrophobic residues, the fold is representative of all known bacterial τ_C domains (Figure 1).

The core of the fold of Domain V of τ displays overall structural similarity with KH type-II domains. Yet, pronounced structural differences in the region of the helix-turn-helix motif that is required for RNA binding in KH domains indicate that the overall structural similarity does not extend to functional similarity. The fold of Domain V of τ can thus be considered to be a unique fold that delivers a fixed-angle relationship between an N-terminal and a C-terminal α-helix (Figure 3).

The apparent absence of significant amino acid conservation on the surface of the domain suggests that interactions between the folded part of the domain and other proteins are not conserved. In stark contrast, the highest sequence conservation is found among the thirty C-terminal residues of τ that are required for binding

to the Pol III α subunit. Three of the six positions of complete sequence conservation are found among these thirty residues (Figure 1). Remarkably, this polypeptide segment is highly solvent-exposed in the free protein and most of it is highly mobile. The sequence conservation in this segment is thus entirely determined by its interaction with a site in α. The conserved amino acids may be optimal for maximum binding affinity across a relatively small protein–protein interface. Alternatively, the sequence conservation may indicate interactions with different proteins during the lifecycle of τ which would be less tolerant to compensating mutations emerging simultaneously in τ and α. At present, it is not known whether a second binding partner exists for this segment of τ.

The τ_C16 domain binds to α considerably more tightly than a peptide comprising the 32 C-terminal residues of τ (24). Yet, no evidence could be found that any other residues of τ interact directly with α. This apparent inconsistency could be explained if the pre-formed secondary structure of helix 6 were important for binding to α. A correctly pre-formed structure has been shown before to greatly enhance binding affinity (42). In addition, it is not clear how much the binding affinities are affected by the generation of charged N- and C-termini in the peptide and in $\tau_C16\Delta11$, respectively, which are not present in the τ_C16 domain.

Interestingly, most of the conserved amino acid residues between Ala614 and Phe630 are hydrophobic and spaced along the amino acid sequence in a manner such that they could lie on the same face of an amphipathic α-helix (Figure 1). This suggests that these residues form a helix upon binding to α (24). A recent report showed that the point mutations R557C and P599L in full-length τ affect the interaction with α (43). Yet, the side chains of these residues are buried with, respectively, 15 and 2% solvent accessibility in the solution structure of τ_C14. This suggests that these mutations affect the interaction with α indirectly by disrupting the structure of τ.

The close proximity of the N- and C-termini in τ_C14 presents a restraint for the maximum spatial separation between the two α subunits in the replisome that simultaneously replicate the leading and lagging DNA strands. First, the N-terminal domains of the τ subunits are located close together in the natural $\delta\tau_2\gamma\delta'$ clamp loader complex (44). Second, the monomeric protein formed by the C-terminal 213 residues of τ binds to DnaB with a dissociation constant greater than 1 μM (20), which is too weak to maintain the complex under physiological conditions. DnaB is, however, a hexamer and binding to tetrameric full-length τ was reported to be much tighter (23). Since the replisome contains two τ molecules but only a single DnaB hexamer (25), both τ proteins are available for binding to the same DnaB hexamer. In view of the results that (i) the DnaB-binding site has been mapped to Domain IVa (residues 430–496) of τ (23), (ii) the structured part of Domain V starts at residue Leu509, and (iii) Domain V is rigidly attached to α, the complex consisting of DnaB, two τ and two α subunits could be structurally quite well defined. This would bring the α subunits into close proximity, in particular as both the DnaB hexamer (310 kDa) and α (130 kDa) are large proteins. These proposals are further supported by an atomic force microscopy study of the DnaB–τ interaction in *Bacillus*, where two τ subunits seem to bind to two neighboring C-terminal hexamerization domains of DnaB (45).

While the residues of Domain IVa were all mobile in the construct comprising Domains IVa and V (τ_C22), significant parts of Domain IVa (which is only 69 residues long) would presumably be immobilized by the interaction with DnaB. Further immobilization may result from interactions with DNA which depend on the positively charged Domain IVa rather than the overall negatively charged Domain V (24). The *E. coli* τ segment connecting the C-terminus of Domain III (Pro368) with the N-terminus of Domain IVa (Lys430), however, comprises a linker which could be designed for flexibility. The amino acid sequence of this linker contains eleven proline residues between residues 370 and 410 and no regular secondary structure is predicted for this segment. Therefore, τ may connect the clamp–loader complex with DnaB via a long flexible linker, whereas the relative orientations of DnaB and the α subunits are more rigidly defined.

SUPPLEMENTARY DATA

Supplementary Data is available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Kiyoshi Ozawa for initial sample characterization by selective ^{15}N-labeling and TOCSY spectra. The project and the purchase of the 800 MHz NMR spectrometer were supported by grants from the Australian Research Council. GO acknowledges an ARC Federation Fellowship. Access to the NMR Facility at the Australian National University (ANU) is is gratefully acknowledged. Funding to pay the Open Access publication charge was provided by ANU.

Conflict of interest statement. None declared.

REFERENCES

1. Lamers,M.H., Georgescu,R.E., Lee,S.G., O'Donnell,M. and Kuriyan,J. (2006) Crystal structure of the catalytic α subunit of *E. coli* replicative DNA polymerase III. *Cell*, **126**, 881–892.
2. Kong,X.-P., Onrust,R., O'Donnell,M. and Kuriyan,J. (1992) Three-dimensional structure of the β subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell*, **69**, 425–437.
3. 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.
4. Mueller,G.A., Kirby,T.W., DeRose,E.F., Li,D., Schaaper,R.M. and London,R.E. (2005) Nuclear magnetic resonance solution structure of the *Escherichia coli* DNA polymerase III θ subunit. *J. Bacteriol.*, **187**, 7081–7089.
5. Guenther,B., Onrust,R., Sali,A., O'Donnell,M. and Kuriyan,J. (1997) Crystal structure of the δ′ subunit of the clamp–loader complex of *E. coli* DNA polymerase III. *Cell*, **91**, 335–345.
6. Jeruzalmi,D., O'Donnell,M. and Kuriyan,J. (2001) Crystal structure of the processivity clamp loader gamma (γ) complex of *E. coli* DNA polymerase III. *Cell*, **106**, 429–441.

7. Kazmirski,S.L., Podobnik,M., Weitze,T.F., O'Donnell,M. and Kuriyan,J. (2004) Structural analysis of the inactive state of the *Escherichia coli* DNA polymerase clamp–loader complex. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 16750–16755.
8. Jeruzalmi,D., Yurieva,O., Zhao,Y., Young,M., Stewart,J., Hingorani,M., O'Donnell,M. and Kuriyan,J. (2001) Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III. *Cell*, **106**, 417–428.
9. Gulbis,J.M., Kazmirski,S.L., Finkelstein,J., Kelman,Z., O'Donnell,M. and Kuriyan,J. (2004) Crystal structure of the χ:Ψ subassembly of the *Escherichia coli* DNA polymerase clamp–loader complex. *Eur. J. Biochem.*, **271**, 439–449.
10. 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.
11. 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.
12. Flower,A.M. and McHenry,C.S. (1990) The γ subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 3713–3717.
13. Tsuchihashi,Z. and Kornberg,A. (1990) Translational frameshifting generates the γ subunit of DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 2516–2520.
14. Gao,D. and McHenry,C.S. (2001) τ binds and organizes *Escherichia coli* replication proteins through distinct domains. Domain III, shared by γ and τ binds $\delta\delta'$ and $\chi\Psi$. *J. Biol. Chem.*, **276**, 4447–4453.
15. Glover,B.P., Pritchard,A.E. and McHenry,C.S. (2001) τ binds and organizes *Escherichia coli* replication proteins through distinct domains. Domain III, shared by γ and τ, oligomerizes DnaX. *J. Biol. Chem.*, **276**, 35842–35846.
16. Studwell-Vaughan,P.S. and O'Donnell,M. (1991) Constitution of the twin polymerase of DNA polymerase III holoenzyme. *J. Biol. Chem.*, **266**, 19833–19841.
17. Onrust,R., Finkelstein,J., Naktinis,V., Turner,J., Fang,L. and O'Donnell,M. (1995) Assembly of a chromosomal replication machine: two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. I. Organization of the clamp loader. *J. Biol. Chem.*, **270**, 13348–13357.
18. Kim,S., Dallman,H.G., McHenry,C.S. and Marians,K.J. (1996) τ couples the leading- and lagging-strand polymerases at the *Escherichia coli* DNA replication fork. *J. Biol. Chem.*, **271**, 21406–21412.
19. Yuzhakov,A., Turner,J. and O'Donnell,M. (1996) Replisome assembly reveals the basis for asymmetric function in leading and lagging strand replication. *Cell*, **86**, 877–886.
20. Dallmann,H.G., Kim,S., Pritchard,A.E., Marians,K.J. and McHenry,C.S. (2000) Characterization of the unique C terminus of the *Escherichia coli* τ DnaX protein. Monomeric C-τ binds α and DnaB and can partially replace τ in reconstituted replication forks. *J. Biol. Chem.*, **275**, 15512–15519.
21. Leu,F.P., Georgescu,R. and O'Donnell,M. (2003) Mechanism of the *E. coli* τ processivity switch during lagging-strand synthesis. *Mol. Cell*, **11**, 315–327.
22. Gao,D. and McHenry,C.S. (2001) τ binds and organizes *Escherichia coli* replication proteins through distinct domains. Partial proteolysis of terminally tagged τ to determine candidate domains and to assign domain V as the α binding domain. *J. Biol. Chem.*, **276**, 4433–4440.
23. Gao,D. and McHenry,C.S. (2001) τ binds and organizes *Escherichia coli* replication proteins through distinct domains. Domain IV, located within the unique C terminus of τ, binds the replication fork helicase, DnaB. *J. Biol. Chem.*, **276**, 4441–4446.
24. 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. (2006) 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.*, accompanying paper.
25. Benkovic,S.J., Valentine,A.M. and Salinas,F. (2001) Replisome-mediated DNA replication. *Annu. Rev. Biochem.*, **70**, 181–208.
26. Wijffels,G., Dalrymple,B.P., Prosselkov,P., Kongsuwan,K., Epa,V.C., Lilley,P.E., Jergic,S., Buchardt,J., Brown,S.E. *et al.* (2004) Inhibition of protein interactions with the β_2 sliding clamp of *Escherichia coli* DNA polymerase III by peptides from β_2-binding proteins. *Biochemistry*, **43**, 5661–5671.
27. 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.
28. Güntert,P., Mumenthaler,C. and Wüthrich,K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, **273**, 283–298.
29. Neri,D., Szyperski,T., Otting,G., Senn,H. and Wüthrich,K. (1989) Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry*, **28**, 7510–7516.
30. Wu,J., Fan,J., Pascal,S.M. and Yang,D. (2004) General method of suppression of diagonal peaks in heteronuclear-edited NOESY spectroscopy. *J. Am. Chem. Soc.*, **126**, 15018–15019.
31. Farrow,N.A., Muhandiram,R., Singer,A.U., Pascal,S.M., Kay,C.M., Gish,G., Shoelson,S.E., Pawson,T., Forman-Kay,J.D. *et al.* (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry*, **33**, 5984–6003.
32. Bartels,C., Xia,T., Güntert,P., Billeter,M. and Wüthrich,K. (1995) The program XEASY for computer-supported NMR spectral analysis. *J. Biomol. NMR.*, **5**, 1–10.
33. Luginbühl,P., Güntert,P., Billeter,M. and Wüthrich,K. (1996) The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR.*, **8**, 136–146.
34. Laskowski,R.A., Rullmann,J.A.C., MacArthur,M.W., Kaptein,R. and Thornton,J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR.*, **8**, 477–486.
35. Koradi,R., Billeter,M. and Wüthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–55.
36. Sevilla-Sierra,P., Otting,G. and Wüthrich,K. (1994) Determination of the nuclear magnetic resonance structure of the DNA-binding domain of the *P22 c2* repressor (1 to 76) in solution and comparison with the DNA-binding domain of the *434* repressor. *J. Mol. Biol.*, **235**, 1003–1020.
37. Su,X.C., Jergic,S., Ozawa,K., Burns,N.D., Dixon,N.E. and Otting,G. (2007) Measurement of dissociation constants of high-molecular weight protein-protein complexes by transferred ^{15}N-relaxation. *J. Biomol. NMR.* In press.
38. Holm,L. and Sander,C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.*, **233**, 123–138.
39. Grishin,N.V. (2001) KH domain: one motif, two folds. *Nucleic Acids Res.*, **29**, 638–643.
40. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
41. Wu,P.S.C., Ozawa,K., Jergic,S., Su,X.-C., Dixon,N.E. and Otting,G. (2006) Amino acid type identification in ^{15}N-HSQC spectra by combinatorial selective ^{15}N-labelling. *J. Biomol. NMR.*, **34**, 13–21.
42. Liepinsh,E., Berndt,K.D., Sillard,R., Mutt,V. and Otting,G. (1994) Solution structure and dynamics of PEC-60, a protein of the Kazal-type inhibitor family, determined by nuclear magnetic resonance. *J. Mol. Biol.*, **239**, 137–153.
43. Pham,P.T., Zhao,W. and Schaaper,R.M. (2006) Mutator mutants of *Escherichia coli* carrying a defect in the DNA polymerase III τ subunit. *Mol. Microbiol.*, **59**, 1149–1161.
44. Bullard,J.M., Pritchard,A.E., Song,M.-S., Glover,B.P., Wieczorek,A., Chen,J., Janjic,N. and McHenry,C.S. (2002) A three-domain structure for the δ subunit of the DNA polymerase III holoenzyme. δ domain III binds δ' and assembles into the DnaX complex. *J. Biol. Chem.*, **277**, 13246–13256.
45. Haroniti,A., Anderson,C., Doddridge,Z., Gardiner,L., Roberts,C.J., Allen,S. and Soultanas,P. (2004) The clamp–loader-helicase interaction in *Bacillus*. Atomic force microscopy reveals the structural organisation of the DnaB-τ complex in *Bacillus*. *J. Mol. Biol.*, **336**, 381–393.