Protein Science (2000), *9*:721–733. Cambridge University Press. Printed in the USA.

NMR solution structure of the θ subunit of DNA polymerase III from *Escherichia coli*

MAX A. KENIRY, HILARY A. BERTHON, JI YEON YANG, CAROLINE S. MILES, AND NICHOLAS E. DIXON

Research School of Chemistry, The Australian National University, GPO Box 414, Canberra, ACT 2601, Australia

(RECEIVED September 9, 1999; FINAL REVISION February 4, 2000; ACCEPTED February 18, 2000)

Abstract

The catalytic core of *Escherichia coli* DNA polymerase III contains three tightly associated subunits (α, ϵ, and θ). The θ subunit is the smallest, but the least understood of the three. As a first step in a program aimed at understanding its function, the structure of the θ subunit has been determined by triple-resonance multidimensional NMR spectroscopy. Although only a small protein, θ was difficult to assign fully because approximately one-third of the protein is unstructured, and some sections of the remaining structured parts undergo intermediate intramolecular exchange. The secondary structure was deduced from the characteristic nuclear Overhauser effect patterns, the $^3J_{HN\alpha}$ coupling constants and the consensus chemical shift index. The C-terminal third of the protein, which has many charged and hydrophilic amino acid residues, has no well-defined secondary structure and exists in a highly dynamic state. The N-terminal two-thirds has three helical segments (Gln10–Asp19, Glu38–Glu43, and His47–Glu54), one short extended segment (Pro34–Ala37), and a long loop (Ala20–Glu29), of which part may undergo intermediate conformational exchange. Solution of the three-dimensional structure by NMR techniques revealed that the helices fold in such a way that the surface of θ is bipolar, with one face of the protein containing most of the acidic residues and the other face containing most of the long chain basic residues. Preliminary chemical shift mapping experiments with a domain of the ϵ subunit have identified a loop region (Ala20–Glu29) in θ as the site of association with ϵ.

Keywords: backbone assignment; circular dichroism; DNA polymerase; protein structure; triple-resonance NMR

The DNA polymerase III (pol III) holoenzyme is largely responsible for replication of *Escherichia coli* chromosomal DNA. The enzyme functions at replication forks where it displays remarkable processivity, being capable of extending a DNA primer by many thousands of nucleotides without dissociating (Kornberg & Baker, 1991; Kelman & O'Donnell, 1994). The catalytic core of pol III contains three tightly associated subunits (α, ϵ, and θ). The α subunit is responsible for the DNA polymerase activity (Maki & Kornberg, 1985; Maki et al., 1985), and the ϵ subunit is a potent (3′–5′) proofreading exonuclease (Scheuermann et al., 1983). Although the exact function of θ (76 amino acids; M_r 8,846) has not been determined, there is some evidence to suggest a role in maintaining fidelity (Studwell-Vaughan & O'Donnell, 1993). The ϵ and α subunits form an isolable complex, while θ binds ϵ but not α. This suggests an $\alpha\epsilon\theta$ linear arrangement (Studwell-Vaughan & O'Donnell, 1993).

As part of a program aimed at understanding the molecular mechanism of the core and the accessory subunits of pol III, we have determined the structure of the smallest (θ) subunit by NMR spectroscopy. It is envisaged that the three-dimensional structure of θ will guide a program of directed mutagenesis, which in concert with ^{13}C-filtered NMR experiments will elucidate the interaction and orientation of the θ subunit in the $\epsilon\theta$ complex.

This work details the physical characterization of θ, the *holE* gene product (Studwell-Vaughan & O'Donnell, 1993), by circular dichroism (CD) spectroscopy and the determination of its structure by NMR spectroscopy. The assignment and structure determination of the θ subunit presents a challenge for modern NMR techniques because θ is a protein that is a mixture of distinct folded and unfolded regions, its spectra have poor chemical shift dispersion, and large parts of the protein are subject to intramolecular exchange processes that lead to broadening of several backbone resonances.

Reprint requests to: Max A. Keniry, Research School of Chemistry, The Australian National University, GPO Box 414, Canberra, ACT 2601, Australia; e-mail: max@rsc.anu.edu.au.

Abbreviations: 1D, one-dimensional; 2D, two-dimensional; 3D, three-dimensional; CD, circular dichroism; CPMG, Carr-Purcell-Meiboon-Gill; CSI, chemical shift index; DTT, dithiothreitol; HSQC, heteronuclear single-quantum coherence; IPTG, isopropyl β-D-thiogalactopyranoside; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; PCR, polymerase chain reaction; RMSD, root-mean-square deviation; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TOCSY, total correlation spectroscopy.

Results

Structural assessment by CD

The far-ultraviolet (UV) CD spectrum of θ at pH 7 indicates a high helical content (Fig. 1). The protein unfolds at low pH, and this is distinguished by a change in the CD spectrum (Fig. 1) and several new cross peaks in the ^{1}H–^{15}N HSQC spectrum (data not shown). A numerical analysis of the CD spectra of θ at pH 7 by the technique of Provencher (Provencher & Glöckner, 1981) gave a value of 57% α-helix. At pH 4, the helical content of θ is 14%. The analysis of the helical content of θ at pH 7 agrees with a recent qualitative report of the features of the θ subunit at pH 5.5 (Li et al., 1999).

Preliminary NMR assessment

A variety of sample conditions was examined by 1D and 2D ^{1}H–^{15}N HSQC NMR using the uniformly ^{15}N-labeled protein sample. Low pH conditions, favored for NMR studies, were ruled out by a structural transition at pH ~5.5. Li et al. (1999) presented CD and unassigned NMR spectra similar to those presented here (but obtained at pH 5.5); we considered this pH too close to the pH of the structural transition to warrant further study under these conditions and instead chose to work at pH 6.5. The θ subunit also showed a tendency to aggregate at elevated temperatures as indicated by the appearance of extra peaks over time. The sample was stable for months at 4 °C, but aggregated within a 24 h period above 30 °C. Some samples were found to be stable under low salt conditions at 25 °C over several 3D NMR runs, and although the peaks were broader than desired, these conditions proved to be the best compromise for further study. The effects of aggregation, such as the appearance of additional cross peaks in the ^{15}N–HSQC spectrum, could be reversed by a 10-fold dilution of the sample in phosphate buffer followed by spin concentration with Centricon-3 centrifugal concentrators (Amicon, Beverly, Massachusetts). Nevertheless, the unfavorable solution properties and the poor dispersion in the ^{1}H–^{15}N HSQC spectrum made the production of a uniformly ^{15}N,^{13}C labeled protein essential for near complete sequential assignments to be made.

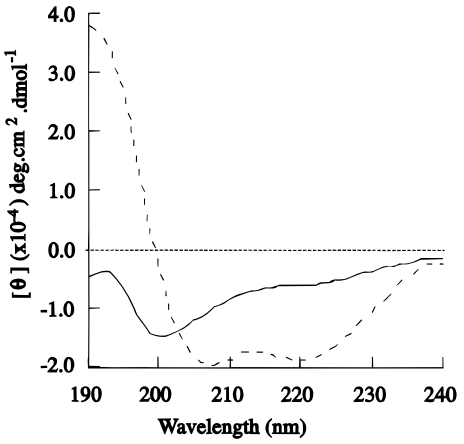


Fig. 1. Far-UV CD spectra of the θ subunit of pol III at pH 7 (– – –) and at pH 4 (—).

Analysis of three ^{15}N–HSQC spectra of θ at 0.25, 0.5, and 1.0 mM concentration in identical buffer conditions showed little or no concentration dependence of the intensity, line width, or ^{15}N chemical shift of the cross peaks. This means that any aggregation that is present at 1.0 mM must be nonspecific.

Resonance assignment and NMR spectroscopy

The 2D ^{15}N–^{1}H HSQC spectrum (Fig. 2) of θ illustrates the degree of overlap of the amide ^{1}H and ^{15}N resonances, and verifies that this protein has a large amount of α-helix, some coiled regions, and very little β-sheet. Also evident are three subsets of cross peaks: a subset of narrow cross peaks that have been assigned to mobile resonances at the N- and C-termini of the protein, a subset of intense but broader cross peaks that have been assigned to resonances in well-defined and stable α-helix, and a subset of broad and less intense cross peaks that belong to regions of the protein that are undergoing intramolecular conformational averaging. Several nuclei in the last subset do not have observable cross peaks in the 2D ^{1}H–^{15}N HSQC spectrum. Each pair of four observable glutamine and asparagine side-chain resonances were assigned (Fig. 2). The single tryptophan side-chain resonance (10.2 ppm, 127.4 ppm) is not shown in this view of the spectrum.

Preliminary resonance assignments were made by analyzing 3D ^{15}N-separated TOCSY–HSQC (Zhang et al., 1994), ^{15}N-separated NOESY–HSQC and ^{15}N-separated HSQC–NOESY–HSQC spectra (Zhang et al., 1994). Some of the helical regions were identified using these spectra, but there were several points of ambiguity including several missing cross peaks that made even a nearly full assignment by the single label strategy improbable. Further progress was made by analysis via a sequential walk of the ^{15}N,^{13}C–HNCACB and ^{15}N,^{13}C–CACB(CO)NH spectra from a ^{15}N,^{13}C–double-labeled sample. At the end of the analysis, there were resonances from several amino acids in two regions of the protein that could not be found. The residues bordering these stretches invariably had broad amide cross peaks in the ^{1}H–^{15}N HSQC spectrum and are at the C-terminal ends of the helices. This leads to the conclusion that the signals are attenuated either by solvent effects or some form of conformational averaging. It was clear that further progress and the determination of the structure could not be achieved with the standard experiments. The C^{β}-decoupled–^{15}N,^{13}C–HNCA and C^{β}-decoupled–^{15}N,^{13}C–HN(CO)CA experiments (Matsuo et al., 1996) complemented the standard ^{15}N,^{13}C–HNCACB and the ^{15}N,^{13}C–CACB(CO)NH experiments by providing some missing cross peaks and sequential assignments. Improved quality spectra, including several additional cross peaks for those residues with broad amide resonances, were obtained by adding a CPMG sequence (Mueller et al., 1995; Mulder et al., 1996) to some standard experiments such as the sensitivity enhanced ^{15}N–NOESY HSQC, ^{15}N–HSQC–NOESY–HSQC, and ^{15}N,^{13}C–HNCO experiments. The enhancement of previously unobserved cross peaks by the addition of a series of refocusing pulses during the period of coherence transfer by scalar coupling confirms the existence of conformational exchange in some regions of θ. Those peaks where the intensity change was most noticeable were the broadest peaks in the spectrum and include Asp19, Leu20, Gly24, Val25, Ala26, Glu29, Arg30, Tyr31, Asn32, Met33, Arg54, Ser63, and Val64.

All the amide ^{1}H and ^{15}N resonances of the backbone except those of Met1, Ala22, Ala23, Phe27, Lys28, and Leu55 to Leu61 have been assigned. This amounted to 86% of the amino acid residues. In addition, the proton and nitrogen resonances of the

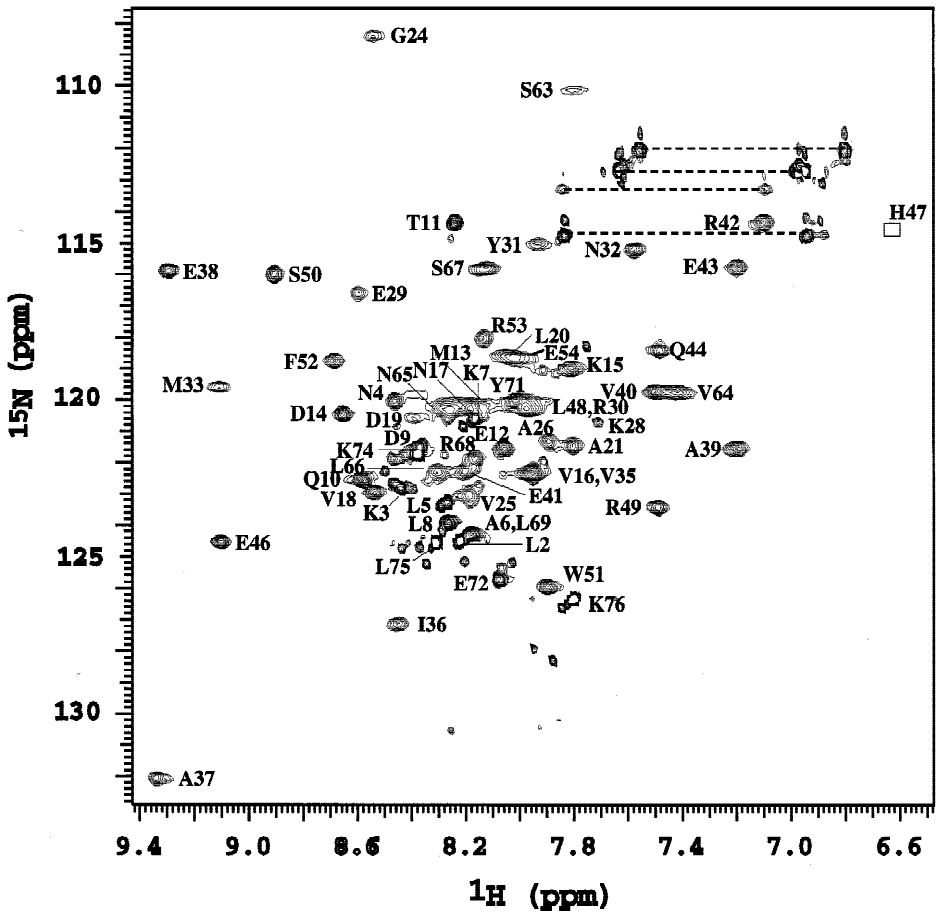


Fig. 2. The sensitivity enhanced ^{1}H–^{15}N HSQC spectrum of the θ subunit of pol III. Assignments for the backbone amides are labeled. Cross peaks connected by dashed lines correspond to glutamine and asparagine side-chain NH_2 groups.

imino group of the indole ring ($NH^{\epsilon 1}$) of Trp51 and four of the six side-chain amino groups of glutamine and asparagine (Asn17, Asn32, Asn65, and Gln10) were observed and assigned based on cross peaks from the ^{15}N–NOESY HSQC spectrum. Side-chain ^{1}H and ^{13}C assignments were obtained via the ^{13}C–HCCH–TOCSY, ^{15}N,^{13}C–C(CO)NH, and ^{15}N,^{13}C–H(CCO)NH experiments.

Secondary structure

The $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, and ^{13}CO chemical shifts are correlated with secondary structure (Wishart et al., 1995). The Chemical Shift Index (CSI) method of Wishart and Sykes (Wishart & Sykes, 1994b; Wishart et al., 1995) provides a convenient procedure for determining the secondary structure elements even when other markers such as the characteristic NOE patterns are absent. The deviations of the $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, and ^{13}CO chemical shifts from their random coil values and the consensus CSI index [calculated using the public domain program, CSI (Wishart & Sykes, 1994a)] are depicted in Figure 3. The consensus CSI clearly shows that all the well-defined secondary structure occurs in the N-terminal two-thirds of θ, while the C-terminal one-third of the protein contains no α-helix or β-structure.

The secondary structure of θ was further characterized by analysis of sequential, medium- and long-range NOEs involving the backbone H^{α}, H^{β}, and amide protons observed in a 80 ms mixing-time, 3D sensitivity-enhanced ^{15}N–NOESY–HSQC. The complete analysis is summarized in Figure 4, and a strip-plot of helix-1 is shown in Figure 5. By combining the CSI data and the secondary structure analysis of Figure 4, we were able to identify three helical regions in θ, two of which are separated by a short loop or kink centered at Pro45. Helix-1 starts at Gln10 and terminates at a broad resonance, Ala26, according to the CSI. The region from Gln10 to Asp19 is characterized by medium to strong $NN(i, i + 1)$ cross peaks and either weak or nonexistent $\alpha N(i, i + 1)$ cross peaks (Fig. 4). Furthermore, helix-1 shows three $\alpha N(i, i + 2)$, four $\alpha N(i, i + 3)$, four $\alpha N(i, i + 4)$ cross peaks, and small $^{3}J_{HN\alpha}$. The section from Leu20 to Met33 is characterized by a series of either broad or nonexistent cross peaks and consequently presented some impediment to a definitive analysis. The section from Leu20 to Ala26 has a negative CSI and weak to medium $NN(i, i + 1)$ cross peaks, but the $\alpha N(i, i + 2)$, $\alpha N(i, i + 3)$, and $\alpha N(i, i + 4)$ cross peaks are absent. This section of the backbone is clearly undergoing conformational exchange but does have a nascent tendency to be helical. Glu29 to Met33 also has medium $NN(i, i + 1)$ cross peaks but the CSI does not indicate a helix. Pro34 to Ala37 is a short extended structure characterized by strong sequential $\alpha N(i, i + 1)$ cross peaks and nonexistent $NN(i, i + 1)$ cross peaks. A second helical region, denoted helix-2, starting at Glu38 and

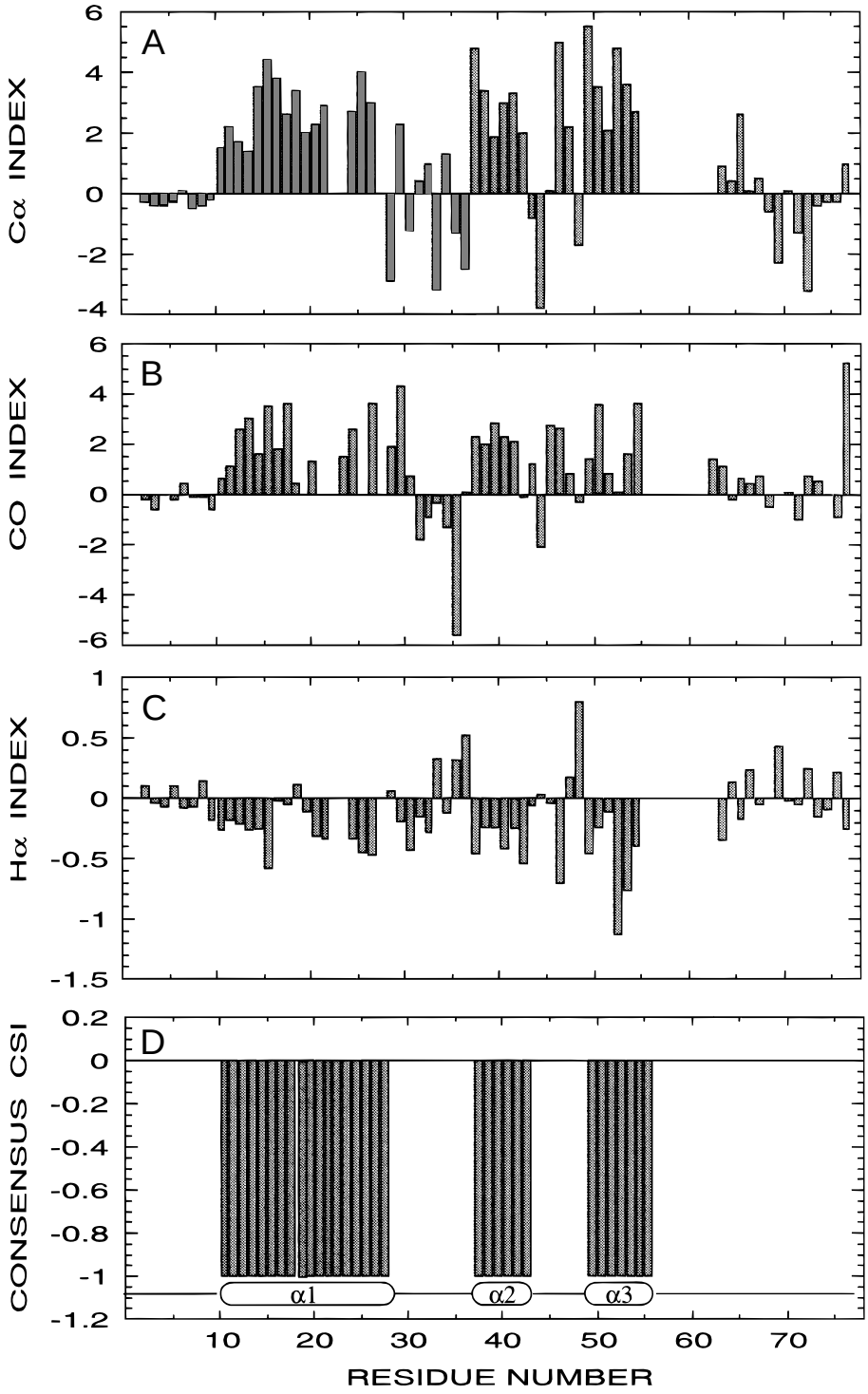


Fig. 3. The H^{α}, $^{13}C^{\alpha}$, and ^{13}CO chemical shift indices and the consensus CSI of the θ subunit of pol III. Helical regions of the protein are characterized by negative deviations for ΔH^{α} and positive deviations for ΔC^{α} and ΔCO. A consensus CSI combines the information from each of these indices. A negative consensus CSI is indicative of helix and a positive consensus CSI suggests β-structure.

ending at Glu43 has three $\alpha N(i, i + 2)$, six $\alpha N(i, i + 3)$, three $\alpha N(i, i + 4)$ cross peaks, and small $^{3}J_{HN\alpha}$. A kink or a short loop located at Gln44 to His47 is centered about a proline (Pro45) and immediately precedes helix-3, which stretches from Leu48 to Glu54. Helix-3 is characterized by one $\alpha N(i, i + 2)$, four $\alpha N(i, i + 3)$, three $\alpha N(i, i + 4)$ cross peaks, and small $^{3}J_{HN\alpha}$. This third helical region also has a very broad resonance at its C-terminus. It could be argued that helix-2 and helix-3 are one single helical region

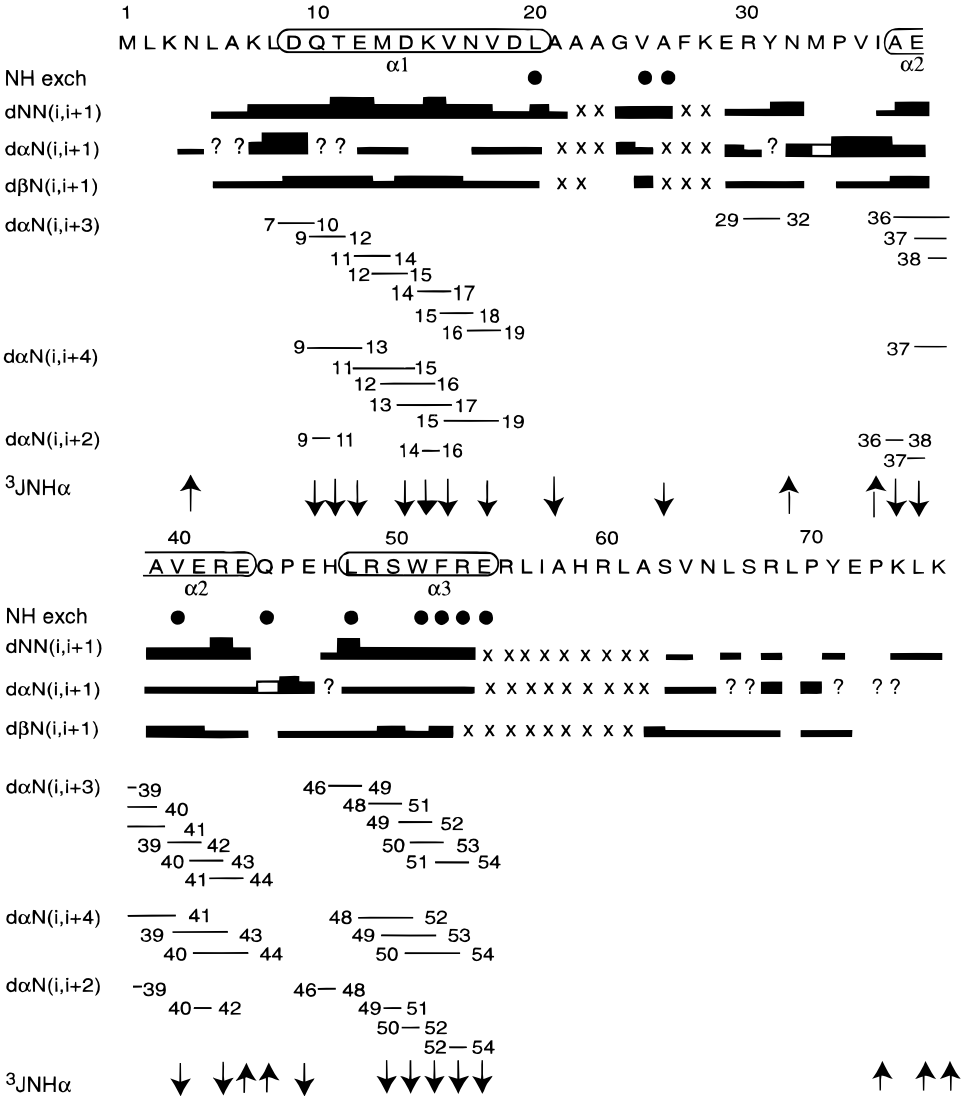


Fig. 4. A summary of the NMR data defining the secondary structure of the θ subunit of pol III at 25 °C and pH 6.5. The primary sequence is displayed at the top. Filled circles immediately below the sequence identify residues with sufficiently slow amide proton exchange to enable observation of an amide cross peak in a ^{1}H–^{15}N HSQC spectrum 1 h after dissolution in D_2O. Sequential NOE connectivities, d_{NN}, $d_{\alpha N}$, and $d_{\beta N}$, are indicated by bars linking the residues concerned and are classified as strong, medium, and weak according to the thickness of the bar. Crosses represent instances where a connectivity cannot be found because of overlap, the amide resonance is broadened because of conformational averaging or, in the case of proline, the connectivity does not exist. Medium-range NOE connectivities, $d_{\alpha N}(i, i + 2)$, $d_{\alpha N}(i, i + 3)$, and $d_{\alpha N}(i, i + 4)$, are linked by lines. $^{3}J_{HN\alpha}$ values greater than 7.5 Hz are indicated by an upright arrow, $^{3}J_{HN\alpha}$ values less than 6 Hz are indicated by an inverted arrow.

punctuated by a kink centered at Pro45. None of the resonances from Arg55 through to Leu61 are observed in the ^{1}H–^{15}N HSQC spectrum. The cross peaks reappear at Ala62, which is very weak and broad, and gradually get narrower and stronger for the resonances nearest the C-terminus.

Structure calculations

The input restraint data were obtained from four spectra. The NH_i to NH_j ($j = i + 1$) NOE data were obtained from 3D ^{15}N–NOESY–HSQC and HSQC–NOESY–HSQC spectra. NH_i to CH_j NOE data were obtained from the 3D ^{15}N–NOESY–HSQC spectrum. Some CH_i to CH_j NOE data were obtained from a 3D ^{13}C–NOESY–HSQC spectrum. All of these spectra were recorded in 90% H_2O/10% D_2O buffer. All NOE data involving aromatic protons were obtained from a 2D NOESY spectrum of an unlabeled sample of θ subunit in 100% D_2O buffer.

The final 511 NOEs (Table 1) were complemented with 300 angle restraints derived from sequential and intraresidue NOEs and 18 $^{3}J_{HN\alpha}$ coupling constants. A total of 11 $C^{\beta}H2$ groups were unambiguously stereoassigned out of 54 spin systems.

All the structures were calculated using the program DYANA 1.5 (Güntert et al., 1997). An iterative process involving several rounds of calculations was employed to assign previously ambiguous NOE cross peaks. Typically each DYANA run was composed of 4,000 simulated annealing steps followed by 1,000 energy min-

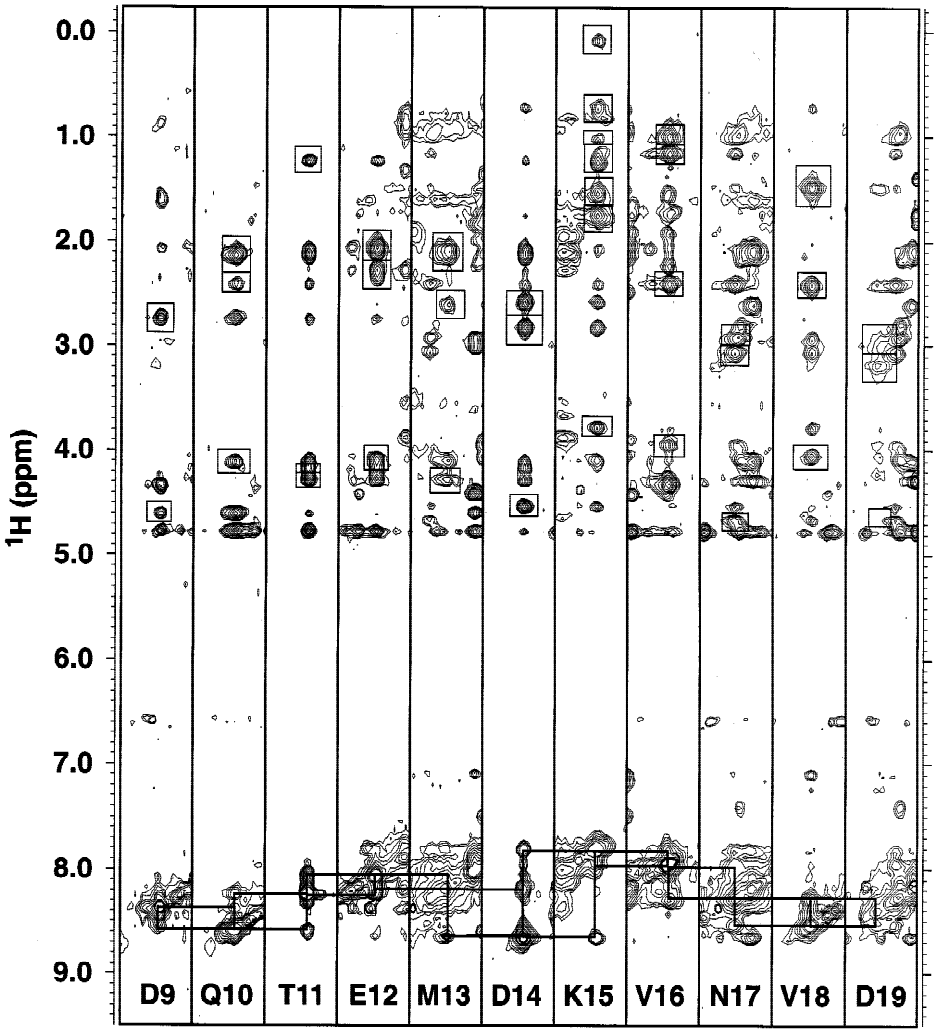


Fig. 5. Strips from the 3D sensitivity enhanced ^{15}N–NOESY–HSQC spectrum of the θ subunit of pol III displaying the sequential and long-range contacts observed for the helical region from D9 to D19. The boxed cross peaks are intraresidue contacts.

imization steps. After the final round of calculations, the 20 structures with the lowest target function were retained from a total of 400 structures calculated. The best fit superposition of the 20 structures for the structured region of θ (Leu8 to Glu54) is shown in Figure 6A. The C-terminal part of the molecule (Leu55 to Lys76) is not shown because its structure is poorly defined due to a lack of restraints in the region between Leu55 and Ala62, although some long-range contacts were observed between Val64–Asn65 and residues Ala26 and Glu29 in loop L1. The remaining part of the C-terminus is clearly unstructured.

Each structure satisfies the experimental restraints with no upper bound violations greater than 0.5 Å, no lower bound violations greater than 0.1 Å, and no dihedral violation greater than 5°. The backbone conformation is well-defined except for seven residues at the N-terminus and 22 residues at the C-terminus, which show a lack of interresidue NOEs. The backbone RMSD of the protein core (residues 8–54) for the 20 best structures is 1.67 ± 0.41 Å, and the heavy atom RMSD for the equivalent segment is 2.53 ± 0.44 Å. The backbone dihedral angles are well defined ($S = 0.8$) for residues that comprise the hydrophobic core of the protein. The resolution is sufficient to define the backbone coordinates of the protein core, which is displayed as a ribbon diagram in Figure 6B. The quality of the structures was analyzed by the programs AQUA and PROCHECK-NMR (Laskowski et al., 1996).

A survey of the structural statistics and residual violations of the experimental restraints is listed in Table 1, and a summary of the distribution of NOE restraints is displayed in Figure 7A. Residues 8–19 and 31–54 have large numbers of medium- and long-range NOEs, and the lowest RMS differences, whereas residues 1–7, 20–30, and 55–76 have very few medium- and long-range NOEs and, hence, large RMS differences. These latter regions have no recognizable secondary structure and are poorly defined by these structures. The mean pairwise RMSD for the helical segments of θ are 1.13 ± 0.24 Å for the backbone heavy atoms (N, C^{α}, C) and 1.7 ± 0.6 Å for all the heavy atoms. The corresponding values for the whole molecule are much larger, at 5.9 ± 1.5 and 6.8 ± 1.4 Å, respectively, reflecting the poor definition of the N- and C-termini and loop L1. Well-defined backbone angles are found for residues Leu8–Leu20, Val25–Ala26, and Tyr31–Glu54, correlating with the regions of low RMSD.

Table 1. *Structural statistics and RMSDs for 20 NMR structures of the θ subunit of DNA polymerase III*

Structural restraints	
Distance restraints	
Total	511
Intraresidue	167
Sequential	172
Medium range $\|i - j\| < 5$	119
Long range $\|i - j\| = 5$	53
Dihedral angle restraints	300
Hydrogen bond restraints	8
Statistics for structure calculations	
Average restraint violations	
NOE violation (Å)	0.23
Dihedral angle violation (°)	<0.5
Atomic RMSD values (Å), all residues (1–76)	
Backbone	5.9 ± 1.5
All heavy atoms	6.8 ± 1.4
Aomic RMSD values (Å), residues 8–54	
Backbone	1.67 ± 0.41
All heavy atoms	2.53 ± 0.44
Atomic RMSD values (Å), α-helices	
Backbone	1.13 ± 0.25
All heavy atoms	1.71 ± 0.59

A Ramanchandran plot, generated by the program PROCHECK-NMR (Laskowski et al., 1996), of the averaged minimized structure shows 49% of the meaningful residues fall into the most favored regions of torsion angle space, 29% of residues are in other allowed regions, 14% of residues fall in the so-called generously allowed region of the Ramanchandran plot, and 8% fall in the disallowed regions. All the residues in the latter two regions are in disordered regions of the protein (particularly the C-terminus) and have no meaningful sequential or medium-range NOE data to restrict the protein backbone to any region of conformational space.

Description of the solution structure

The well-defined N-terminal region of the θ subunit of DNA polymerase III (residues 8–54) is an elongated structure consisting of three helices that fold in an arch-like formation. The base of the arch is formed by helix-1 (Gln10–Asp19), and the apex is formed by helix-2 (Glu38–Glu43) and helix-3 (His47–Glu54), which are separated by a short stretch of amino acids containing a proline. Asp19 and Glu54 may not be the true termini of helices 1 and 3, respectively, because conformational averaging at the C-termini of these two helices broadens the proton resonances, thereby reducing the number of NMR restraints and consequently the definition of the structure. The first seven residues of the N-terminus are poorly defined as is most of the C-terminus. There is also a poorly defined loop region between Leu20 and Arg30.

Helix-1 makes an angle of 18° with helix-3, and this orientation may be determined by a hydrogen bond between the $NH^{\epsilon 1}$ proton of Trp51 and the backbone carbonyl group of Lys15, and by a salt bridge between Lys15 and Glu54 (Fig. 8). The side chain of Lys15 sits directly above the aromatic ring of Trp51, which explains the substantial upfield ring current shifts of the side-chain protons of Lys15. Similarly, the side chain of Val18 is in the plane of the aromatic ring of Trp51, which accounts for the downfield shift of the valine side-chain proton resonances. Val18, Pro45, His47, Leu48, and Trp51 occupy the hydrophobic core between helix-1 and helix-3 (Fig. 8). The ring of His47 is oriented at right angles to that of Trp51. The side chain of Leu48 is oriented above the plane of this ring accounting for the upfield shift of the δ proton resonances. A long flexible loop L1 (colored yellow in Fig. 6B) leads to a structured region that changes the direction of the chain. Tyr31 is important for stabilizing this region of the protein. The aromatic proton resonances of Tyr31 are magnetically inequivalent indicating that this aromatic ring reorients slowly on the NMR timescale.

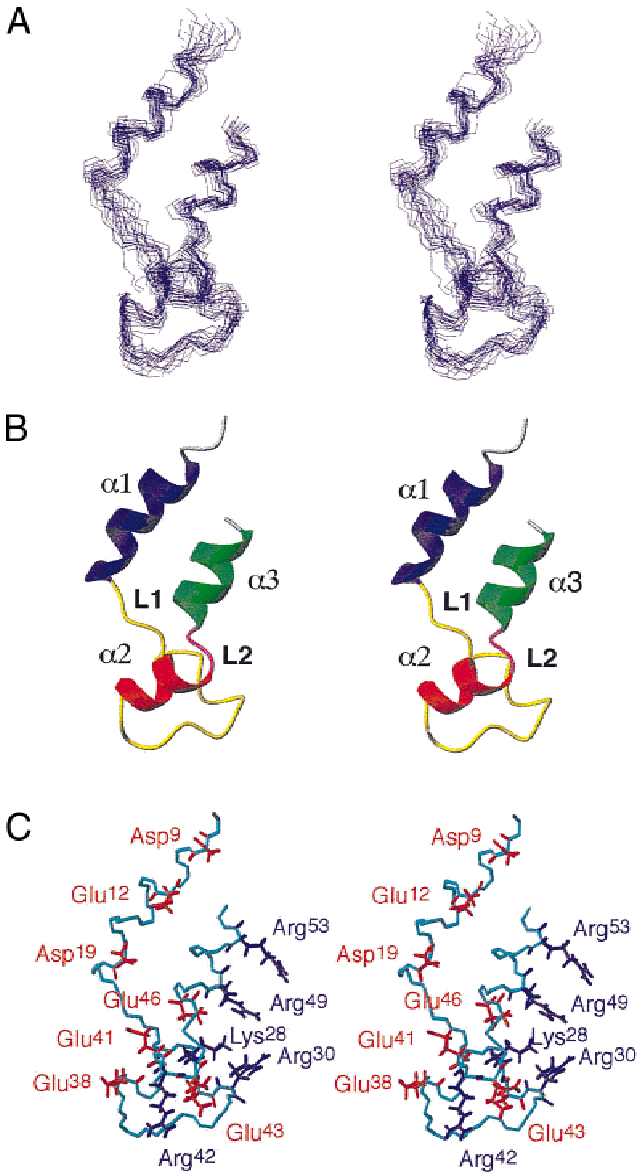
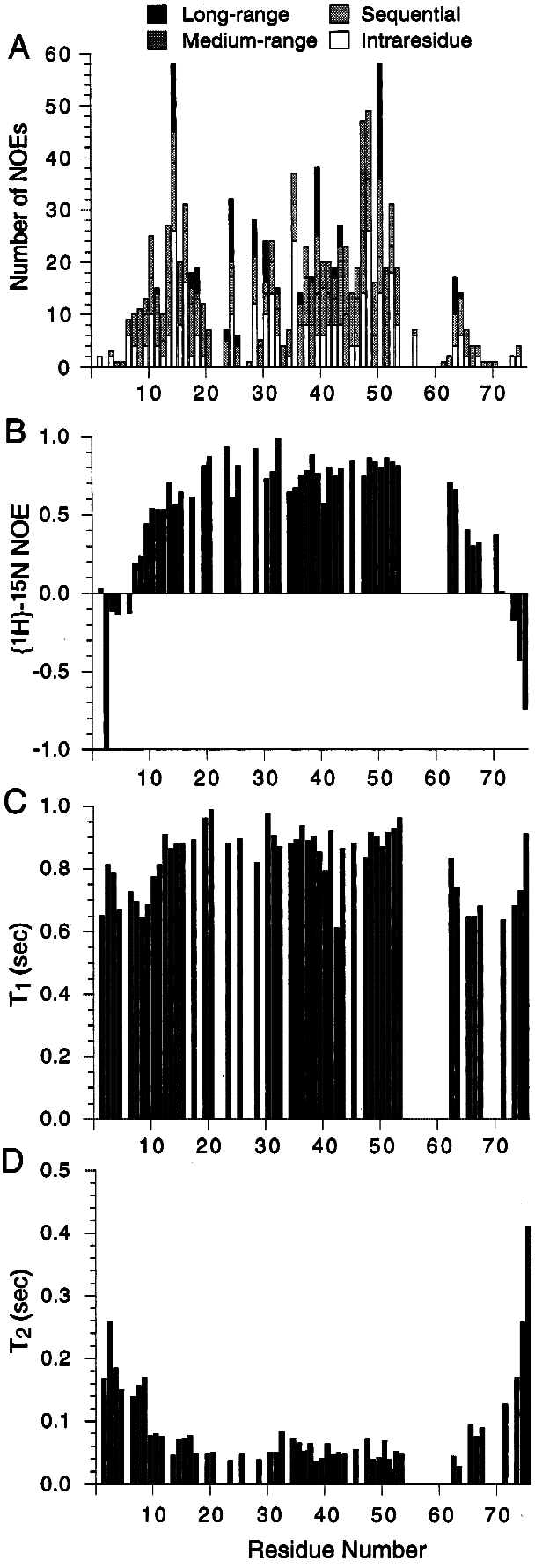


Fig. 6. A: Stereo-superposition of the backbone heavy atoms N, C^{α}, C′, and O from the final ensemble of 20 structures for the θ subunit of pol III. Only the relatively well-determined region from Leu8 to Glu54 is shown. **B:** Ribbon diagram of the conformer closest to the mean structure of the θ subunit. **C:** Some significant charged side chains of the θ subunit (the conformer closest to the mean structure is displayed). Basic side chains (Lys and Arg) are colored blue, and acidic side chains (Glu and Asp) are colored red.



A ^{15}N-filtered NOESY experiment verified that the hydroxyl proton of Tyr31 is slowly exchanging and thus possibly involved in a hydrogen bond. NOEs between the aromatic protons of Tyr31 and the amide protons of Glu43 and Gln44, and substantial ring current shifts of the side chains of these two residues verify that the protein backbone has a turn structure in the intervening residues. Sequential NOEs between the amide proton of Asn32 and the aromatic protons of Tyr31 verify that the aromatic ring is oriented toward Asn32 and Met33. Inspection of the final structures shows that the hydroxyl group of Tyr31 may form a hydrogen bond with the sulfur atom of Met33. There is a short stretch of extended structure from Pro43 to Ala37 in which several hydrophobic side chains are oriented toward the solvent. In association with the loop L1, which has a preponderance of hydrophobic residues, this face of θ forms a hydrophobic surface (Fig. 9B).

Helix-2 and helix-3 are separated by a type III turn segment (L2 in Fig. 6B) such that the axis of helix-2 is at a 43° angle to the axis of helix-3. The turn is stabilized by a hydrogen bond between the amide proton of Leu48 and the carbonyl of Pro45. There is also a cavity in the protein structure bounded by the three helices and part of the loop L1. Because of substantial averaging of several residues within this loop, it is difficult to define the exact size of the cavity.

One notable feature of this protein is the three distinct surfaces. The hydrophobic surface is noted above. The surface of helix-1 has several acidic residues (Asp9, Glu12, Asp14, and Asp19), which are shown in red in Figure 6C, and, on the opposite face, there are several arginine and lysine side chains (Lys28, Arg30, Arg42, Arg49, and Arg53), which are shown in blue. The amphipathic and bipolar nature of the structure section of θ is also depicted by a surface representation of the electrostatic charge distribution. The acidic nature of helix-1 and the basic cleft are shown in Figure 9A (which shows the same view as Figs. 6 and 8) and the hydrophobic surface formed by loop L1 and the extended structure are shown by a view rotated by 90° (Fig. 9B).

The peptide bonds preceding Pro34 and Pro45 are in the *trans* conformation as shown by strong $H^{\alpha}(i)$–$H^{\delta}(i+1)$ NOEs. Pro70 and Pro73 are part of the very mobile C-terminus and do not appear to favor a single conformation. There is doubling of cross peaks for Glu72, Lys74, Leu75, and Lys76, suggesting *cis-trans* isomerization at Pro70 and Pro73.

Backbone dynamics

The hydrogen–deuterium exchange rate of amide protons is indicative of the stability of the secondary structure elements. In many proteins, the ^{1}H–^{15}N cross peaks of the amide groups of residues in α-helices are visible for several hours after the exchange of H_2O with D_2O. The θ subunit by itself in solution appears to have an

Fig. 7. A: Distribution of NOE restraints for the θ subunit of pol III. The height of each bar denotes the number of NOEs. The bars signify intraresidue, sequential, medium-range, and long-range NOEs as shown in the legend above the graph. **B:** The heteronuclear ^{15}N{^{1}H} NOE values for the θ subunit of pol III. **C:** ^{15}N T_1 relaxation times for the θ subunit of pol III. **D:** ^{15}N T_2 relaxation times for the θ subunit of pol III. Where no data are given, the amide cross peak was either not observed or it was severely overlapped. The relaxation data were measured at 25 °C, pH 6.5 and at a field strength of 14.1 T.

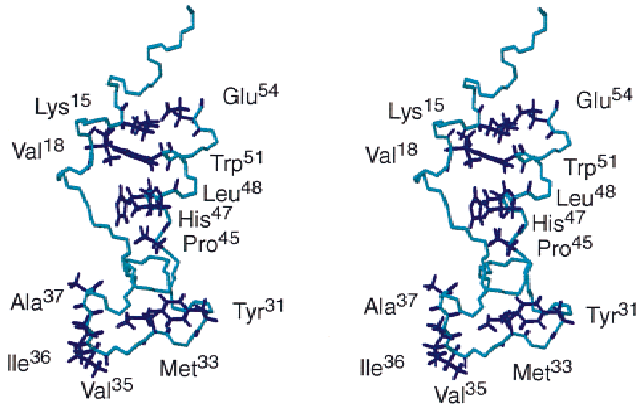


Fig. 8. Stereo-diagram of the θ subunit (Leu8 to Glu54) showing the location of selected side chains (the conformer closest to the mean structure is displayed).

abnormal degree of flexibility, even though the N-terminal two-thirds does have identifiable secondary structure elements. The exchange rate for each of the helical regions is unusually rapid. Only nine amide resonances were observable within 1 h of the exchange of the solvent, 90% H_2O/10% D_2O with 100% D_2O, and no amide resonances were observable after 3 h. Several of these more slowly exchanging resonances are in helix-3 (Trp51, Phe52, Arg53, and Glu54) and some include broad and weak resonances (Leu20, Ala21, and Glu54), which leads to the conclusion that the intermediate exchange process arises from conformational averaging and not from intermediate exchange of the amide and solvent protons.

Two of the helical regions are terminated by residues that appear to be undergoing conformational averaging. The resonances of Asp19, Leu20, and Ala21 at the C-terminus of helix-1 are weak and broad, as is that of Glu54 at the C-terminus of helix-3. The overall conclusion is that each of the α-helices exists in a relatively dynamic state and suggests that, in the absence of the ϵ subunit, θ is not completely folded.

Backbone flexibility investigated by ^{15}N T_1, T_2, and $^{15}N\{^1H\}$ heteronuclear NOE

Analysis of the backbone dynamics of proteins via relaxation data is a useful adjunct to structural studies because it provides information about the mobility of distinct regions of a protein that are not directly available from the ^1H–^1H NOE data. The ^{15}N T_1 and T_2 values were obtained for 50 of the 72 protonated amide nitrogens of θ and are displayed in Figures 7C and 7D, respectively. The mean ^{15}N T_1 and T_2 values for residues in the helical regions are 0.85 and 0.058 s, respectively, while the mean T_1 and T_2 for the whole protein is 0.79 and 0.071 s, respectively. The heteronuclear $^{15}N\{^1H\}$ NOE is more sensitive to the effects of internal motion than T_1 and T_2 and serves as a monitor of the flexibility of various segments of the backbone of the θ subunit. It is affected by the overall tumbling of the protein by fast internal motions and slower restricted reorientation of the protein backbone. A maximal NOE of 0.83 indicates that the N–H bond vectors are rigid and reorienting at the same rate as the overall tumbling of the molecule (Grasberger et al., 1993). At the other end of the scale, a NOE of −3.5 indicates rapid and very large amplitude motions (Grasberger et al., 1993). Short stretches of negative NOEs are frequently observed at the N- and/or C-termini of proteins where there is no long-lived secondary structure. The $^{15}N\{^1H\}$ NOE data of θ are shown in Figure 7B. The average NOE of θ is 0.51, and the average NOE for the well-defined secondary structure is 0.72.

Several negative NOEs and long T_2 values at the termini indicate that the first seven residues at the N-terminus and the final

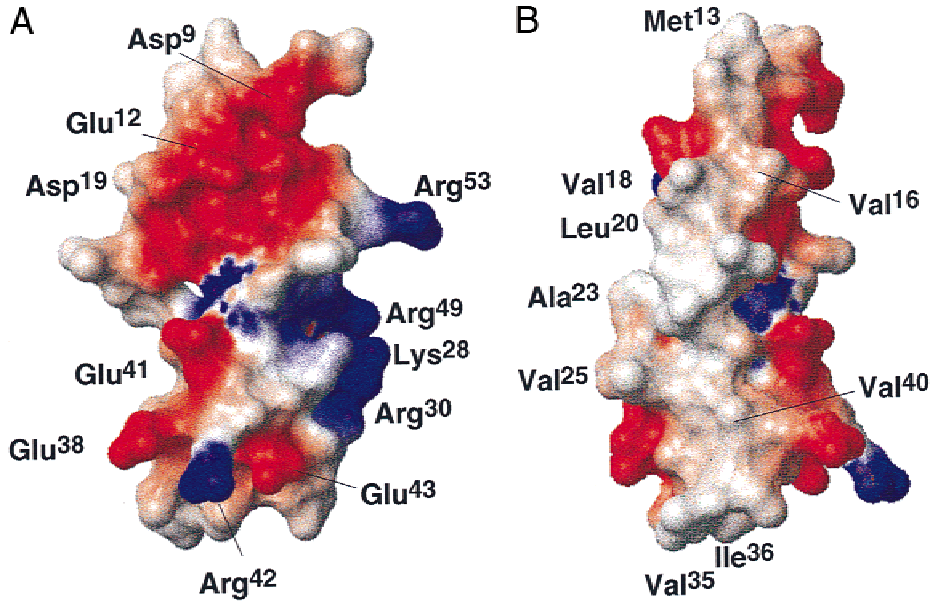


Fig. 9. A: Surface representation showing the charge distribution mapped onto the θ subunit (Leu8 to Glu54). The regions of basic potential are shown in blue and acidic in red. The conformer closest to the mean structure is displayed and the molecule has the same orientation as shown in Figures 6 and 8. **B:** A 90° rotation of the view shown in **A** depicting the hydrophobic surface formed by loop L1 and the extended region from Pro34 to Ala37.

seven residues at the C-terminus are undergoing rapid, large-amplitude motions. The resonances of these nuclei at the termini of the protein chain are also readily identified in the ^{1}H–^{15}N and ^{1}H–^{13}C HSQC spectra by their narrow linewidth. The rest of the protein has positive NOEs and T_2 values $<$100 ms. It is notable that the NOEs in the loop L1 are large and positive and, although the coordinates in this region have low precision, the residues in this loop are not undergoing large, rapid internal motions. This is in contrast to the substantial decrease in NOE values observed in proteins with a flexible loop region, for example, calmodulin (Barbato et al., 1992). This supports the possibility that at least the N-terminal end of the loop may be a molten helix. The positive NOEs and short T_2 values are maintained up to residues Val64 and Asn65. Because of the relative rigidity of Val64–Asn65, and because none of the resonances from Leu55 to Ala62 are observed in the spectra, it is likely that this region of the protein is undergoing some form of restricted motion on the intermediate timescale.

Discussion

Although the θ subunit resides in the catalytic core of pol III, which has been extensively studied (Maki & Kornberg, 1985; McHenry, 1988; McHenry, 1991; Studwell-Vaughan & O'Donnell, 1991), very little is known about its function, and until this work, very little was known about its structure. We determined the 3D structure of the protein by NMR methods as a necessary starting point for understanding its interaction with the ϵ subunit. Although a small protein, it presented problems for NMR solution studies because of its helical character, the dynamic nature of parts of the backbone, and its unfavorable solution properties, all of which have also been reported by Li et al. (1999).

The fold of the θ subunit appears to be unique. A search of the Protein Data Bank (PDB) using the program DALI (Holm & Sander, 1993) could not identify any protein fold with a meaningful Z score with either the complete protein structure (residues 1–76) or the region of the protein that is structured (residues 8–54).

There are two interesting features near the two central prolines of θ. Pro45 breaks a long helical stretch from Glu38 to Glu54. The secondary structure near Pro45 has features reminiscent of the break in helix-2 of the DNA binding domain of the heat shock transcription factor, HSF (Damberger et al., 1994). A proline in the middle of a helix usually produces a bend but, as with HSF, there is a substantial distortion that produces a large $^{3}J_{NH\alpha}$ at Glu43 and Gln44, and a break in the consistently negative consensus CSI. A stretch of residues (Glu29 to Met33), with features reminiscent of a helical structure, is abruptly terminated by Pro34 and changes to a short extended region (Val35 to Ala37) that has strong $\alpha N(i, i+1)$ cross peaks.

Two aromatic residues appear to have crucial importance to the structural integrity of the θ subunit. Trp51 is buried between helix-1 and helix-3, and the hydrogen bond involving its slowly exchanging indole proton stabilizes the interaction between these two helices. His47 also lies between helix-1 and helix-3, and because we observe, near pH 5.5, a large structural transition that induces the loss of most of the helical secondary structure, the uncharged state of this residue appears necessary for the structural integrity of θ. The only other amino acid residue that titrates in the range pH 5 to 6 is His59, which is not in the hydrophobic core of θ and, furthermore, is undergoing conformational averaging in the intermediate timescale because no resonances are observed from this residue.

Preliminary experiments, in which ^{15}N,^{13}C-labeled θ was bound to unlabeled N-terminal domain of the ϵ subunit, have demonstrated that more ^{1}H–^{15}N cross peaks with a greater dispersion are observed in the ^{15}N HSQC spectrum of the complex. Preliminary assignments suggest that the flexible loop from Leu20 to Asn32 is immobilized by the interaction with ϵ, because several resonances within loop L1 are shifted. Loop L1 is highly hydrophobic and, furthermore, contains a predominance of short chain hydrophobic residues. One stretch, in particular, Ala21, Ala22, Ala23, Gly24, Val25, and Ala26, may have evolved to promote very close contact with the ϵ subunit. Further work is underway to map the ϵ binding surface of θ and to determine the changes in structure and dynamics induced by the binding process.

The θ subunit is bound to the ϵ subunit, the 3′–5′ exonuclease in pol III, which is in turn bound to the α subunit, the 5′–3′ polymerase. In analogy to DNA polymerase I, θ might be thought to be a 5′–3′ exonuclease, but exonuclease activity has never been observed for θ, nor does pol III possess 5′–3′ exonuclease activity (O'Donnell & Kornberg, 1985). The only known effect that has been ascribed to θ is the stimulation of removal by ϵ of a G-T mispair at a primer terminus (Studwell-Vaughan & O'Donnell, 1993). An interesting parallel is seen in the 8 kDa domain of DNA polymerase β (pol β), which binds to the 5′-phosphate of a downstream primer via a lysine-rich pocket (Pelletier et al., 1996). θ has a lysine- and arginine-rich region (Figs. 6C, 9A), which may also be a phosphate recognition site. This basic domain is also on a face of θ that is opposite to the hydrophobic loop that interacts with ϵ. It has been suggested that the 8 kDa domain of pol β increases fidelity by bending the template DNA, thereby minimizing the indiscriminate base-stacking forces between the primer terminus and the incoming base (Pelletier et al., 1996). Further experimentation is necessary to determine whether θ functions in a similar manner or has some other phosphate recognition function.

The NMR structure of the θ subunit provides the basis for the rational design of further biochemical experiments, particularly those that are intended to elucidate its interaction with the ϵ subunit of pol III.

Materials and methods

Construction of plasmid pCM869 for the overproduction of θ

The plasmid pETθ (Studwell-Vaughan & O'Donnell, 1993), a derivative of vector pET3c (Studier et al., 1990) containing the *holE* gene with an *Nde*I site at the start codon, was a gift from Dr. Mike O'Donnell. It was used as template for a PCR reaction, using as primers oligodeoxynucleotides 5′-GGAGACCACAACGTTTC (complementary to vector sequences upstream of the gene) and 5′-CTCGAATTCTTATTAAGTTTGGGCT-CGTAAGG. Use of the latter primer generated an *Eco*RI site (underlined) after the *holE* stop codon (italicized). The PCR mixture (50 μL) contained pETθ (77 pg), both primers (1 μM each), $MgSO_4$ (1 mM), dNTPs (10 μM each), and Vent DNA polymerase (New England Biolabs, Beverly, Massachusetts). The thermal cycler was programmed as follows: 94 °C for 5 min; 94 °C for 15 s, 55 °C for 15 s and 72 °C for 45 s (three times); 94 °C for 15 s, 72 °C for 15 s and 72 °C for 45 s (32 times); 20 min at 72 °C and 5 min at 25 °C. The amplified DNA was purified using Wizard PCR Preps (Promega, Madison, Wisconsin) and digested with *Nde*I and *Eco*RI. The resulting 233-bp fragment was inserted between the corresponding sites in the pET derivative pETMCSI (Miles et al., 1997) to generate plasmid

pCM869. Ampicillin-resistant transformants of strain BL21::DE3 (Studier et al., 1990) were selected at 37 °C and plasmid DNAs were isolated on a small scale (expected size 4,877 bp). Overproduction of the 8.8 kDa θ subunit following induction by IPTG at 37 °C was confirmed by SDS-PAGE. The sequence of *holE* from pCM869 was determined: one silent mutation had been introduced during the PCR reaction, the codon for Ile36 being changed from ATC to ATT.

Bacterial growth

For preparation of uniformly ^{15}N- and/or ^{13}C-labeled θ, *E. coli* strain BL21::DE3/pCM869 was grown in a minimal medium containing (per liter): K_2HPO_4 (10.6 g), $NaH_2PO_4 \cdot H_2O$ (6.1 g), vitamin B_1 (1 mg), $MgSO_4$ (1 mM), trace salts and ampicillin (50 mg), supplemented with $^{15}NH_4Cl$ (1 g) and/or ^{13}C-glucose (0.65 g). Growth was at 37 °C until $A_{595} = 0.5$, when production of θ was induced by the addition of IPTG to 1 mM. After 3 h at 37 °C, cells were harvested by centrifugation, yielding 2 g (wet weight) per liter.

Protein purification

Cells (2 g) were resuspended in lysis buffer (50 mM Tris·HCl pH 7.5, 10% w/v sucrose, 200 mM NaCl, 20 mM spermidine, 2 mM DTT; 50 mL) and lysed using a French press operated at 12 kpsi. The supernatant obtained following centrifugation (38,700 × *g*, 45 min, 4 °C; Fraction I) was dialyzed overnight at 4 °C against two changes of Buffer A (20 mM Tris·HCl pH 7.5, 20% glycerol, 0.25 mM EDTA, 2 mM DTT). After clarification by centrifugation, Fraction I was loaded onto a column (75 mL) of DEAE-Fractogel (Merck, Darmstadt, Germany), previously equilibrated in Buffer A at a flow rate of 0.6 mL/min. The eluted θ subunit, which did not bind to the resin, was sufficiently pure for direct use in NMR experiments (verified by SDS-PAGE). The protein in fractions containing θ was precipitated with ammonium sulfate (0.5 g/mL), resuspended after centrifugation in 7 mL of 10 mM Na·phosphate pH 6.5 (Buffer B) and dialyzed extensively against the same buffer. The yield of purified ^{13}C,^{15}N-labeled θ from a 1 L culture was 58 mg. The molecular weight of similarly prepared unlabeled θ was determined to be 8,845.5 ± 0.4 (cf. calculated value: 8,846.31) by electrospray mass spectrometry, essentially as described (Miles et al., 1997).

Circular dichroism

CD spectra of θ in 10 mM Na·phosphate pH 7.0 were recorded on a Jobin-Yvon (Longjumeau, France) CD6 spectrodichrograph using a cell of pathlength 0.1 m. Each spectrum resulted from an average of five successive individual scans. The peptide concentration was estimated from A_{280} assuming a value of 8,250 L/mol·cm for ϵ_{280}. The pH of the protein solution was adjusted from 7 to 4 by addition of small aliquots of 1 M HCl.

NMR spectroscopy

NMR measurements were performed at 25 °C on each of the ^{15}N-labeled, ^{13}C,^{15}N-double labeled and unlabeled θ. Each labeled protein sample was dissolved to a concentration of 1 mM in Buffer B containing 10% D_2O. Unlabeled θ was dissolved in Buffer B prepared in 100% D_2O. Initial 2D and 3D NMR experiments were carried out on a Varian VXR-500S NMR spectrometer. Subsequent experiments were performed on Varian INOVA-500 and Varian INOVA-600 spectrometers. Data from the following experiments were collected on the ^{15}N-labeled protein: sensitivity-enhanced ^{15}N–HSQC (Kay et al., 1992), ^{15}N–HMQC–J (Szyperski et al., 1992), 2D ^{15}N–HSQC–TOCSY (Zhang et al., 1994) with a mixing time of 70 ms, 3D sensitivity-enhanced ^{15}N–TOCSY–HSQC (Zhang et al., 1994) with a mixing time of 70 ms, 3D sensitivity-enhanced ^{15}N–NOESY–HSQC (Zhang et al., 1994) with mixing times of 80 and 140 ms, 3D sensitivity-enhanced ^{15}N–HSQC–NOESY–HSQC (Zhang et al., 1994) with a mixing time of 140 ms, 3D sensitivity-enhanced ^{15}N–HSQC–TOCSY–NOESY–HSQC (Zhang et al., 1994) with a NOESY mixing time of 100 ms and a TOCSY mixing time of 35 ms. Typically, datasets of 96 × 44 × 512 complex points were acquired in the *t*1, *t*2, and *t*3 domains, respectively, over a ^{1}H sweep width of 7,000 Hz and an ^{15}N sweep width of 1,700 Hz. FIDs in *t*1 were extended by 96 complex points by linear prediction using 12 coefficients followed by zero filling to 256 complex points and FIDs in *t*2 were extended by 44 complex points using eight coefficients following by zero filling to 128 complex points. T_1 and T_2 relaxation times and ^{15}N{^{1}H} steady-state NOEs were determined from a series of 2D ^{1}H–^{15}N correlation spectra at 14.1 T as described by Farrow et al. (1994). ^{15}N T_1 values were measured from spectra recorded with relaxation delays of 5, 60, 140, 240, 370, 530, 760, and 1,150 ms. ^{15}N T_2 values were determined from spectra with CPMG relaxation delays of 16, 32, 48, 63, 94, 126, and 158 ms. ^{15}N{^{1}H} steady-state NOE values were obtained by recording spectra with and without ^{1}H presaturation of duration 4 s and a recycle delay of 2 s. Saturation was achieved by a series of 120° pulses at 5 ms intervals over the 4 s saturation time. The steady-state saturation was determined by monitoring the intensity of the first increment while varying the saturation time. All the above experiments employed pulsed field gradients.

The following experiments were collected on the ^{15}N,^{13}C double-labeled protein: HNCA (Yamazaki et al., 1994), HN(CO)CA (Yamazaki et al., 1994), C^β-decoupled HNCA (Matsuo et al., 1996), C^β-decoupled HN(CO)CA (Matsuo et al., 1996), HNCO (Muhandiram & Kay, 1994), HCACO (Grzesiek & Bax, 1993), HN(CA)CO (Engelke & Rüterjans, 1995), HNCO-NOESY (Löhr & Rüterjans, 1997), HNCACB (Yamazaki et al., 1994), CBCA(CO)NH (Grzesiek & Bax, 1992), HCCH-TOCSY (Kay et al., 1993), H(CCO)NH (Grzesiek et al., 1993), and C(CO)NH (Grzesiek et al., 1993). 2D NOESY with a mixing time of 100 ms and 3D TOCSY–NOESY with a NOESY mixing time of 150 ms and a TOCSY mixing time of 50 ms (Oschkinat et al., 1989) were performed on the unlabeled sample. Typically, 512 complex points over a sweep width of 7,000 Hz were collected for the directly acquired ^{1}H dimension, 96 complex points over a 7,000 Hz sweep width for the indirectly acquired ^{1}H dimension, and 44 complex points over a sweep width of 1,700 Hz for the indirectly acquired ^{15}N dimension. Where ^{13}C was detected in the indirectly acquired dimension, 96 complex points over a sweep width of 6,000 Hz were acquired for the aliphatic carbon region, 60 complex points over 3,000 Hz for the C^α region, and 32–48 complex points over 2,400 Hz for the carbonyl region. In the ^{13}C HCCH–TOCSY and ^{13}C-separated NOESY experiments, 32 complex points over a sweep width of 3,600 Hz was acquired for the ^{13}C dimension. Free induction decays in the indirectly acquired domain were extended by linear prediction and zero filling in a manner similar to that described above.

The amide exchange rate was probed by exchanging the buffer solvent with 100% D_2O by rapidly passing the protein solution through a NAP-10 column (Pharmacia, Uppsala, Sweden) pre-equilibrated with Buffer B in 100% D_2O and then quickly acquiring 2D ^1H–^{15}N HSQC experiments at 15 and 195 min post-buffer exchange at 20 °C. No further experiments were collected because only nine amide resonances were observed at the first time point and no resonances were observed at the second time point.

Spectral analysis and experimental restraints

Multidimensional Fourier transformation employing linear prediction was performed with Varian VNMR versions 5.2 and 6.1 software (Varian Associates, Palo Alto, California). After transformation, the processed data from all experiments were converted into XEASY format (Bartels et al., 1995). All peak picking, spectral analysis, and semiautomated resonance assignments were performed with the XEASY software package (Bartels et al., 1995). Three-dimensional NOESY cross peaks were integrated using the program SPSCAN 1.0.53 (R.W. Glaser & K. Wüthrich, http://www.mol.biol.ethz.ch/wuthrich/software/spscan/). CALIBA (part of the DYANA suite of programs) was used to convert NOESY cross-peak volumes into upper and lower distance bounds. Hydrogen bond restraints were included for slowly exchanging NH groups for which a single hydrogen bond acceptor was identified in preliminary structure calculations and was consistent with the secondary structure determined previously. Lower bound distance restraints between the hydrogen bond donor and acceptor and the hydrogen bond donor antecedent and the acceptor of 1.5 and 2.3 Å, respectively, were employed. Upper bound distance restraints for the same bonds were 2.3 and 3.3 Å, respectively. The $^3J_{HN\alpha}$ coupling constants, estimated from the ^{15}N–HNHA experiment by the method of Kuboniwa et al. (1994), were considered quantitative and given the limits $\pm$1.0 Hz. $^3J_{HN\alpha}$ values in the 5.5–8 Hz range were not used in the calculations. χ_1 angle restraints and H^β stereospecific assignments were made via estimates of the $^3J_{NH\beta}$ values obtained from relative HN–H^β cross-peak intensities in a 3D HNHB spectrum using the program HABAS. $^3J_{NH\beta}$ values were treated qualitatively and given the limits of $\pm$2.0 Hz. The coordinates of the θ subunit of DNA polymerase III have been deposited with the Protein Data Bank (accession number 1DU2).

Acknowledgments

The authors would like to acknowledge Professor M. O'Donnell for providing pETθ; Prof. K. Wüthrich (ETH, Zürich) for the XEASY and MOLMOL software packages; Prof. Brian Sykes (University of Alberta) for the CSI software package; Prof. Lewis Kay (University of Toronto) for the code for the sensitivity-enhanced ^{15}N pulse sequences that were used in the initial stages of this work; E.A. Owen for assistance with the CD spectroscopy and the DYANA calculations; and C. Braybrook for the ESI-MS determination. The authors have benefited from the use of NMR facilities at the Australian National University NMR Centre.

References

Barbato G, Ikura M, Kay LE, Pastor RW, Bax A. 1992. Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: The central helix is flexible. *Biochemistry 31*:5269–5278.

Bartels C, Xia T-H, Billeter M, Güntert P, Wüthrich K. 1995. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J Biomol NMR 6*:1–10.

Damberger FF, Pelton JG, Harrison CJ, Nelson HCM, Wemmer DE. 1994. Solution structure of the DNA-binding domain of the heat shock transcription factor determined by multidimensional heteronuclear magnetic resonance spectroscopy. *J Biomol NMR 3*:1806–1821.

Engelke J, Rüterjans H. 1995. Sequential protein backbone resonance assignments using an improved 3D-HN(CA)CO pulse scheme. *J Magn Reson Ser B 109*:318–322.

Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. 1994. Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry 33*:5984–6003.

Grasberger BL, Gronenborn AM, Clore GM. 1993. Analysis of the backbone dynamics of interleukin-8 by ^{15}N-relaxation measurements. *J Mol Biol 230*:364–372.

Grzesiek S, Anglister J, Bax A. 1993. Correlation of backbone amide and aliphatic side-chain resonances in $^{13}C/^{15}N$-enriched proteins by isotropic mixing of ^{13}C magnetization. *J Magn Reson Ser B 101*:114–119.

Grzesiek S, Bax A. 1992. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc 114*:6291–6293.

Grzesiek S, Bax A. 1993. The origin and removal of artifacts in 3D HCACO spectra of proteins uniformly enriched with ^{13}C. *J Magn Reson Ser B 102*:103–106.

Güntert P, Mumenthaler C, Wüthrich K. 1997. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol 273*:283–298.

Holm L, Sander C. 1993. Protein structure comparison by alignment of distance matrices. *J Mol Biol 233*:123–138.

Kay LE, Keifer P, Saarinen T. 1992. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc 114*:10663–10665.

Kay LE, Xu GY, Singer AU, Muhandiram DR, Forman-Kay JD. 1993. A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^1H and ^{13}C correlations in H_2O samples of proteins. *J Magn Reson Ser B 101*:333–337.

Kelman Z, O'Donnell M. 1994. DNA replication: Enzymology and mechanisms. *Curr Opin Genet Dev 4*:185–195.

Kornberg A, Baker TA. 1991. *DNA replication*, 2nd ed. New York: Freeman.

Kuboniwa H, Grzesiek S, Delaglio F, Bax A. 1994. Measurement of H^N-H^α J couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods. *J Biomol NMR 4*:871–878.

Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM. 1996. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR 8*:477–486.

Li D, Allen DL, Harvey S, Perrino FW, Schaaper RM, London RE. 1999. A preliminary CD and NMR study of the *Escherichia coli* DNA polymerase III θ subunit. *Proteins 36*:111–116.

Löhr F, Rüterjans H. 1997. Unambiguous NOE assignments in proteins by a combination of through-bond and through-space correlations. *J Biomol NMR 9*:371–378.

Maki H, Horiuchi T, 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.

Maki H, 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.

Matsuo H, Kupce E, Li HJ, Wagner G. 1996. Increased sensitivity in HNCA and HN(CO)CA experiments by selective C^β decoupling. *J Magn Reson Ser B 113*:91–96.

McHenry CS. 1988. DNA polymerase III holoenzyme of *Escherichia coli*. *Annu Rev Biochem 57*:519–550.

McHenry CS. 1991. DNA polymerase III holoenzyme: Component structure and mechanism of a true replicative complex. *J Biol Chem 266*:19127–19130.

Miles CS, Weigelt J, Stamford NPJ, Dammerova N, Otting G, Dixon NE. 1997. Precise limits of the N-terminal domain of DnaB helicase determined by NMR spectroscopy. *Biochem Biophys Res Commun 231*:126–130.

Mueller L, Legault P, Pardi A. 1995. Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino protons. *J Am Chem Soc 117*:11043–11048.

Muhandiram DR, Kay LE. 1994. Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J Magn Reson Ser B 103*:203–216.

Mulder FAA, Spronk CAEM, Slijper M, Kaptein R, Boelens R. 1996. Improved HSQC experiments for the observation of exchange broadened signals. *J Biomol NMR 8*:223–228.

O'Donnell ME, Kornberg A. 1985. Dynamics of DNA polymerase III holoenzyme of *Escherichia coli* in replication of a multiprimed template. *J Biol Chem 260*:12875–12883.

Oschkinat H, Cieslar C, Gronenborn AM, Clore GM. 1989. Three-dimensional

homonuclear Hartmann-Hahn-nuclear Overhauser enhancement spectroscopy in H_2O and its application to proteins. *J Magn Reson 81*:212–216.

Pelletier H, Sawaya MR, Wolfle W, Wilson SH, Kraut J. 1996. Crystal structure of human polymerase β complexed with DNA: Implications for catalytic mechanism, processivity and fidelity. *Biochemistry 35*:12742–12761.

Provencher SW, Glöckner J. 1981. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry 20*:33–37.

Scheuermann RH, Tam S, Burgers PMJ, 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.

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol 185*: 60–89.

Studwell-Vaughan PS, O'Donnell M. 1991. Constitution of the twin polymerase of DNA polymerase III holoenzyme. *J Biol Chem 266*:19833–19841.

Studwell-Vaughan PS, O'Donnell M. 1993. DNA polymerase III accessory proteins. V. θ encoded by *holE*. *J Biol Chem 268*:11785–11791.

Szyperski T, Güntert P, Otting G, Wüthrich K. 1992. Determination of scalar coupling constants by inverse Fourier transformation of in-phase multiplets. *J Magn Reson 99*:552–560.

Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. 1995. ^{1}H, ^{13}C and ^{15}N random coil chemical shifts of common amino acids. I. Investigation of nearest-neighbour effects. *J Biomol NMR 5*:67–81.

Wishart DS, Sykes BD. 1994a. The ^{13}C Chemical Shift Index: A simple method for the identification of protein secondary structure using ^{13}C chemical shift data. *J Biomol NMR 4*:171–180.

Wishart DS, Sykes BD. 1994b. Chemical shifts as a tool for structure determination. *Methods Enzymol 239*:363–392.

Yamazaki T, Lee W, Arrowsmith CH, Muhandiram DR, Kay LE. 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.

Zhang O, Kay LE, Olivier JP, Forman-Kay JD. 1994. Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J Biomol NMR 4*:845–858.