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

Received for publication, October 27, 2000
Published, JBC Papers in Press, November 14, 2000, DOI 10.1074/jbc.M009828200

Dexiang Gao and Charles S. McHenry
From the Department of Biochemistry and Molecular Genetics and Program in Molecular Biology, University of Colorado Health Sciences Center, Denver, Colorado 80262

The τ subunit dimerizes Escherichia coli DNA polymerase III core through interactions with the α subunit. In addition to playing critical roles in the structural organization of the holoenzyme, τ mediates intersubunit communications required for efficient replication fork function. We identified potential structural domains of this multifunctional subunit by limited proteolysis of C-terminal biotin-tagged τ proteins. The cleavage sites of each of eight different proteases were found to be clustered within four regions of the τ subunit. The second susceptible region corresponds to the hinge between domain II and III of the highly homologous δ' subunit, and the third region is near the C-terminal end of the τ-δ' alignment (Guenther, B., Onrust, R., Sali, A., O'Donnell, M., and Kuriyan, J. (1997) Cell 91, 335–345). We propose a five-domain structure for the τ protein. Domains I and II are based on the crystallographic structure of δ' by Guenther and colleagues. Domains III–V are based on our protease cleavage results. Using this information, we expressed biotin-tagged τ proteins lacking specific protease-resistant domains and analyzed their binding to the α subunit by surface plasmon resonance. Results from these studies indicated that the α binding site of τ lies within its C-terminal 147 residues (domain V).

The structurally complex DNA polymerase III holoenzyme is responsible for replication of the majority of the chromosome in Escherichia coli. The polymerase of the enzyme and 3' → 5' exonuclease proofreading activities are contained within the heterotrimeric DNA polymerase III core (αεθ) subassembly. The holoenzyme contains seven different auxiliary subunits (β, γ, δ, δ', τ, χ, and ψ) that confer a number of special properties requisite for replicative polymerase function (1–4). These properties include a rapid elongation rate, high processivity, and the ability to communicate with primosomal proteins at the replication fork (5–7). The auxiliary subunits are divided between two functional assemblies: a β2 sliding clamp processivity factor, and the DnaX complex, a multiprotein ATPase that assembles the β2 processivity factor onto the primer-template.

Both the τ and γ subunits of the holoenzyme are expressed from dnaX. Translation of dnaX gene yields the full-length τ subunit (71 kDa) as well as the γ subunit (47 kDa), which corresponds to the N-terminal two-thirds of the τ sequence (8–11). The γ subunit results from −1 translational frameshifting to a frame with an early stop codon. The τ subunit plays central roles in the structure and function of the holoenzyme. Interactions between the τ and α subunits result in the formation of a dimeric DNA polymerase III' (αεθ)2τ2 (12, 13). This dimeric polymerase effectively couples synthesis of the leading and lagging strands (12, 14). The τ subunit binds tightly to α, but the shorter dnaX translation product (γ) does not. This observation suggested that the C-terminal portion unique to τ is critical for its interactions with the α subunit. Indeed, the α subunit and C-τ, an OmpT proteolytic fragment corresponding to the 215 C-terminal residues of τ, bind with a 1:1 stoichiometry (15).

Interactions between the τ subunit and DnaB helicase (DnaB) are critical for rapid movement at the replication fork (16, 17). In systems using the reconstituted DNA polymerase III holoenzyme, τ subunit–DnaB interactions stimulate the rate of helicase unwinding more than 10-fold to levels approaching the rate of fork progression in vivo. The C-terminal region found in τ but lacking in γ has been implicated in replication fork function. The C-τ fragment was shown to interact with DnaB and to effectively couple the leading strand polymerase with DnaB helicase at the replication fork (15).

The τ subunits bind γδδ'χψ to form the DnaX complex, τ2γ(δδ'χψ) (18, 19). The τ subunit also serves as a bridge between α and a γ-Ssb interaction, strengthening the holoenzyme interactions with the single-stranded DNA-binding protein-coated lagging strand at the replication fork (20, 21). As part of the elongation complex, τ protects β2 from removal by exogenous γ complex, increasing the processivity of the replicase to the megabase range (22).

Clearly, τ mediates its functions through interactions with other subunits. To identify distinct structural domains that might mediate these multiple interactions, we performed limited proteolytic digestion of recombinant, biotin-tagged τ. Based on these findings, we constructed plasmids encoding truncated τ fusion proteins lacking one or more putative structural domains. The relative binding of each resultant purified fusion protein to the α subunit was determined by surface plasmon resonance. These studies enabled the identification of domain V (147 C-terminal amino acid residues) as the α subunit-binding domain of τ.

EXPERIMENTAL PROCEDURES

Strains—E. coli DH5α and HB101 were used for initial molecular cloning procedures and plasmid propagation. E. coli BL21 and BL21(A DE3) were employed for protein expression.

Chemicals and Reagents—All proteases were purchased from Roche Molecular Biochemicals or Sigma. d-Biotin was purchased from Sigma. SDS-polyacrylamide gel electrophoresis protein standards were ob-
tained from Amersham Pharmacia Biotech, and prestained molecular mass markers were from Bio-Rad or Life Technologies, Inc. Ni²⁺-NTA\(^{\text{TM}}\) resin, QiAquick Gel extraction kits, QiAquick PCR purification kits, and plasmid preparation kits were purchased from Qiagen (Valencia, CA). The Coomassie Plus Protein Assay Reagent and ImmunoPure Streptavidin are vended by Pierce. CM5 sensor chips (research grade), P-20 surfactant, NHS, EDC, and ethanolamine hydrochloride were obtained from BIAcore, Inc. (Piscataway, NJ).

**Construction of the Fusion Plasmids—** The N- and C-terminal fusion vectors pPA1\(_N\)-N0 and pPA1\(_C\)-CO were constructed as previously described (23). The fusion peptides contained a short 13-amino acid biotinylation sequence, a hexahistidine sequence, and a thrombin cleavage site. The induced fusion proteins are under the control of either the T7 promoter of pET-11C vector (Novagen, Madison, WI) or the p\(_{\text{Al}}\text{lac}0\) promoter/operator (referred to as p\(_{\text{Al}}\)). p\(_{\text{Al}}\) is a semi-synthetic E. coli RNA polymerase-dependent promoter containing two lac operators (23, 24). The dnaX gene was derived from the pRT610A plasmid in which the dnaX gene was modified at the frameshifting site. This modification results in the specific expression of the \(\tau\) subunit; the alternative expression product, \(\gamma\), is not encoded by this construct (25).

P\(_{\text{Al}}\)-N-\(\tau\) plasmid (see Fig. 1A) encodes the \(\tau\) protein lacking only the first amino acid (methionine) with an N-terminal fusion peptide placed in frame. P\(_{\text{Al}}\)-N-\(\Delta\tau\) was generated by replacing the \(\tau\)-encoding gene dna\(\tau\) from the vector used as the starting material (plasmid P\(_{\text{Al}}\)-N0) with dna\(\tau\) (see Fig. 1A). Oligonucleotides 2782 and 2783 (Table I), which correspond to the codons for \(\tau\) amino acids 2–10, were annealed and inserted into pBluescript (KS\(^{\text{TM}}\)) to generate pDG10. The remainder of the dna\(\tau\) gene sequence was from pDG50, which was derived from pRT610A via elimination of a PstI restriction site. The 2187-base-pair NarI/HindIII restriction fragment from pDG50 was ligated to a similarly digested pDG10 vector to generate pDG100. The 2209-base-pair PstI fragment from pDG100 replaced the corresponding fragment in the N-terminal fusion peptide-containing vector pP\(_{\text{Al}}\)-N-\(\Delta\tau\). Plasmid pP\(_{\text{Al}}\)-N0 encodes the intact \(\tau\) protein tagged with a C-terminal fusion peptide (see Fig. 1B). The Nhel/PstI fragment at the C-terminal end of the dna\(\tau\) sequence within pRT610A was replaced with a PCR-generated fragment to produce pRT610AM, in which the stop codon was replaced with a SpeI cloning site. The 1963-base-pair Nhel/SpeI fragment from this plasmid was used to replace the corresponding fragment in the C-terminal fusion peptide-expressing vector pP\(_{\text{Al}}\)-N0 to produce pP\(_{\text{Al}}\)-\(\text{C}(0)\).

PCR was used to generate plasmid pP\(_{\text{A}}\)-N-\(\Delta\text{I}3\), which lacks the sequences encoding the N-terminal 413 amino acids of \(\tau\). Oligonucleotides N\(-\text{413P1}\) and N\(-\text{413P2}\) (Table I) were used with template pP\(_{\text{A}}\)-N-\(\Delta\tau\) to PCR amplify a truncated dna\(\tau\) fragment. N\(-\text{413P1}\) contained a PstI site in the noncomplementary 5' region followed by a complementary region extending from codons 414–419. N\(-\text{413P2}\) annealed to a region of dna\(\tau\) located 100 bases downstream of Nhel site. To generate pP\(_{\text{A}}\)-N-\(\Delta\text{I}3\), the amplified dna\(\tau\) fragment was ligated into pP\(_{\text{A}}\)-N-\(\Delta\tau\) following digestion of the plasmid with PstI and Nhel.

Plasmid pP\(_{\text{ET11}-\text{N496}}\) lacks the sequences encoding the N-terminal 496 amino acids of \(\tau\). Primers N\(-\text{496P1}\) and N\(-\text{496P2}\) were used with template pP\(_{\text{Al}}\)-N-\(\Delta\tau\) to generate a partial dna\(\tau\) fragment, which contained a PstI site in the noncomplementary 5' region and KpnI restriction site more proximal to the 3' end. The KpnI restriction site was added by amplifying the N-end of the dna\(\tau\) natural stop codon (see Fig. 1A). This PCR fragment replaced the dna\(\tau\) gene in pUtK\(\text{P}1\)-digested pET11-N (23) to generate pET11-N-\(\Delta\text{I}496\).

Plasmid pP\(_{\text{A}}\)-C-\(\Delta\text{I}47\) lacks the sequences encoding the C-terminal 147 amino acids of \(\tau\). PCR primer C-\(\text{147P1}\) annealed to a dna\(\tau\) sequence located 146 bases upstream of the RsrII cloning site. The other PCR primer C-\(\text{147P2}\) was complementary to the dna\(\tau\) from codon 494–496, followed by a noncomplementary SpeI cloning site. The two primers were used with template pRT610A to generate a partial dna\(\tau\) sequence. Following RsrII and SpeI digestion, this fragment replaced the corresponding fragment in pP\(_{\text{A}}\)-\(\text{C}(0)\) (see Fig. 1B) to generate pP\(_{\text{A}}\)-C-\(\Delta\text{I}47\).

**Cell Growth and Induction—** For protein expression applications, E. coli strain BL21 was transformed with plasmids P\(_{\text{A}}\)-\(\text{C}(0)\) or P\(_{\text{A}}\)-N-\(\Delta\tau\). E. coli strain BL21(DE3) was used for each of the other expression plasmids. E. coli bearing plasmid P\(_{\text{A}}\)-N-\(\Delta\tau\) was grown at 37 °C to an optical density of 0.8 in 6 liters of F medium (26) containing 100 \(\mu\text{g}\)/ml ampicillin. Bacteria transformed with each of the other plasmids were grown to the same density under the same conditions, except that the volume of F medium was 2 liters. The induction process was started by the addition of isopropyl-b-D-thio-galactoside (final concentration, 1 mM). Additional ampicillin (100 \(\mu\text{g}\)/ml) and d-biotin (10 \(\mu\text{M}\)) were added to the media at the same time. After 2 h of induction, cells were harvested by centrifugation at 5800 \(\times\) g for 10 min at 4 °C and resuspended in 1 ml of Tris-sucrose buffer (50 mM Tris-HCl, pH 7.5, and 100 \(\mu\text{M}\) dithiothreitol) (50 mM sodium phosphate, pH 7.6, 500 mM NaCl, 10 mM glycerol, 0.5 mM PMSF, 0.5 mM benzamidine, and 1 mM imidazole). Ni²⁺-NTA resin, previously equilibrated with Buffer L, was added to the suspensions for binding. The peak fraction eluted at about 60 mM imidazole. Wash and elution steps for N-\(\Delta\tau\) were the same as above, except for modifications of the precipitation, binding, and washing steps for N-\(\Delta\tau\). Fusion protein complexes were then packed into columns. Columns were washed with 10 column volumes of Buffer L and then with roughly 30 column volumes of Buffer W. The proteins were eluted with 100 mM sodium phosphate, pH 7.6, 500 mM NaCl, 20% glycerol, 0.5 mM PMSF, and 0.5 mM benzamidine) plus 50 mM sodium phosphate, pH 7.6, 500 mM NaCl, 20% glycerol, 0.5 mM PMSF, and 0.5 mM benzamidine) plus 23 mM imidazole.

### Table I

| Oligonucleotide | Use | Sequence \(^a\) |
|----------------|-----|----------------|
| 2782           | N-\(\Delta\tau\) | AAGCTTGTGGTGGAGGCTATACGCTGACCTGACCGCAACNGGAAGGCACCATCTGCGGCGAAGCTAGCGGCAGAAGC |
| 2783           | N-\(\Delta\tau\) | TACCACTTCGAGACTGTAATGGGGCGG |
| 3221           | C(0)r | TACCAATTTVGGCTTGGCTCAGTGG |
| 3222           | C(0)r | GGACTAGTCGCCTTCGG |
| C-\(\text{147P1}\) | C-\(\text{147P1}\) | TACCAATTTVGGCTTGGCTCAGTGG |
| C-\(\text{147P2}\) | C-\(\text{147P2}\) | GGACTAGTCGCCTTCGG |
| N-\(\text{413P1}\) | N-\(\text{413P1}\) | TACCAATTTVGGCTTGGCTCAGTGG |
| N-\(\text{413P2}\) | N-\(\text{413P2}\) | GGACTAGTCGCCTTCGG |
| N-\(\text{496P1}\) | N-\(\text{496P1}\) | TACCAATTTVGGCTTGGCTCAGTGG |
| N-\(\text{496P2}\) | N-\(\text{496P2}\) | GGACTAGTCGCCTTCGG |

\(^a\) The underlined sequences are complementary to portions of the dna\(\tau\) gene.

\(\text{NTA}, \text{nitrilotriacetic acid; CAPS}, 3\)-cychoexylyamin-1-propanesulfonic acid; PCR, polymerase chain reaction; NHS, N-hydroxysuccinimide; EDC, 1-ethyl-3-[3-dimethylaminopropyl]-carbodiimide; PMSF, phenylmethylsulfonyl fluoride; RU, response units.
C-Δ147r proteins were each eluted in single steps with 150 mM imidazole in Buffer W.

**SDS-Polyacrylamide Gel Electrophoresis**—Proteins were separated by overnight electrophoresis at 65 V on a 10–17.5% SDS-polyacrylamide gradient gel (0.75 × 18 × 16 cm). Gels were stained with a 0.1% solution of Coomassie Brilliant Blue R-250 in 20% methanol and 10% acetic acid and then destained in a solution of 10% methanol and 10% acetic acid.

**Biotin Blots**—After separation by SDS-polyacrylamide gel electrophoresis, proteins were transferred onto polyvinylidene difluoride membranes at 500 mA for 3 h in 25 mM Tris-HCl, 192 mM glycine, pH 8.3, 20% methanol, and 0.01% SDS. Membranes were dipped in methanol and then air-dried for 20 min. Membranes were incubated with alkaline phosphatase-conjugated streptavidin (2 μg/ml) in TBS + 0.05% Tween 20 plus 0.5% nonfat milk for 1 h at room temperature and then washed three times in TBS +0.05% Tween20. Blots were developed in a substrate solution containing nitroblue tetrazolium chloride (0.33 mg/ml) and 5-bromo-4-chloro-3-indolylphosphate toluidine salt (0.165 mg/ml) in 0.1 mM Tris-HCl (pH 9.5), 0.1 mM NaCl, and 50 mM MgCl2. Reactions were stopped by washing the membranes with distilled water.

**Limited Proteolysis**—C(0)147r digests were carried out in 50 mM HEPES-KOH (pH 7.4), 150 mM NaCl, 10% glycerol, and 0.1 mM EDTA. Proteolytic digests of C-Δ13r were in 50 mM Tris-HCl (pH 7.6), 100 mM NaCl, 5% glycerol, and 10 mM MgCl2. At different time points, 15-μl aliquots from reaction mixtures were removed, mixed with 8 μl of stop buffer (0.18× Tris, pH 6.8, 30% sucrose, 6% SDS, 180 mM dithiothreitol), and then immediately boiled for 2 min. Each aliquot contained 3 μg of protein. Digestion products were separated by SDS-polyacrylamide gel electrophoresis and then stained with Coomassie Brilliant Blue or transferred onto a polyvinylidene difluoride membrane for biotin blots.

**Protein Sequencing**—After digestion, the selected biotinylated fragments were purified from others by using Ni2+-NTA chromatography with the same buffers used for N-Δ1r except that urea was added to the binding and washing buffer at 8 M final concentration. These purified biotinylated fragments were resolved by SDS-polyacrylamide gel electrophoresis and transferred onto polyvinylidene difluoride membranes in 10 mM CAPS (pH 11.0) and 10% methanol at constant current (0.4 A) for 3 h. The membranes were washed with 20% aqueous methanol and then subjected to N-terminal sequence analysis using standard Edman chemistry (James McManaman, University of Colorado Cancer Center Protein Core Laboratory).

**Protein Determinations**—Concentrations of all purified proteins were determined by UV spectroscopy using their extinction coefficients. Concentrations of r fusion proteins were determined by the Pierce Coomassie Plus Protein Assay Reagent according to the manufacturer’s instructions. Bovine serum albumin (fat-free; Sigma) was used as a standard.

**DNA Polymerization Assays**—Activities of r fusion proteins were measured by their requirement for reconstitution of holozyme activity and by measuring DNA synthesis from a primed M13 ori template (26). Assay mixtures (25 μl) contained 500 pmol of M13Gori (as nucleotides), 185 units (40 ng) of DnaG primase, 1.6 μg of E. coli single-stranded DNA binding protein, 250 fmol each of DNA polymerase III core (αεθ), β, β′, γ, and the test r fusion protein (20–50 fmol). Reactions were performed in a buffer containing 50 mM HEPES-KOH (pH 7.5), 10% (v/v) glycerol, 100 mM potassium glutamate, 10 mM dithiothreitol, 10 mM Mg(CH3COO)2, 200 μg/ml bovine serum albumin, 0.02% (v/v) Tween 20, 48 μM dATP, 48 μM dGTP, 48 μM dCTP, 18 μM [3H]dTPP (specific activity, 520 cpm/pmol TTP), and 200 μM rNTP. Assay mixtures were incubated at 30 °C for 5 min, quenched by trichloroacetic acid precipitation, and then filtered through GF/C filters (26). One unit is defined as the amount of enzyme catalyzing the incorporation of 1 pmol of dNTPs/min at 30 °C.

**Surface Plasmon Resonance**—A BLAcore™ instrument was used for protein binding analyses. CM5 research grade sensor chips were used for all experiments. The carboxymethyl dextran matrix by injecting over the chip in 10 mM sodium acetate (pH 4.5) buffer at 0.2 and 0.1 mg/ml, respectively. IgG was used to partially block the negatively charged carboxyl groups on the sensor chip surface.

**Protein Determinations**—Concentrations of all purified proteins were determined by UV spectroscopy using their extinction coefficients. Concentrations of r fusion proteins were determined by the Pierce Coomassie Plus Protein Assay Reagent according to the manufacturer’s instructions. Bovine serum albumin (fat-free; Sigma) was used as a standard.

**DNA Polymerization Assays**—Activities of r fusion proteins were measured by their requirement for reconstitution of holozyme activity and by measuring DNA synthesis from a primed M13Gori template (26). Assay mixtures (25 μl) contained 500 pmol of M13Gori (as nucleotides), 185 units (40 ng) of DnaG primase, 1.6 μg of E. coli single-stranded DNA binding protein, 250 fmol each of DNA polymerase III core (αεθ), β, β′, γ, and the test r fusion protein (20–50 fmol). Reactions were performed in a buffer containing 50 mM HEPES-KOH (pH 7.5), 10% (v/v) glycerol, 100 mM potassium glutamate, 10 mM dithiothreitol, 10 mM Mg(CH3COO)2, 200 μg/ml bovine serum albumin, 0.02% (v/v) Tween 20, 48 μM dATP, 48 μM dGTP, 48 μM dCTP, 18 μM [3H]dTPP (specific activity, 520 cpm/pmol TTP), and 200 μM rNTP. Assay mixtures were incubated at 30 °C for 5 min, quenched by trichloroacetic acid precipitation, and then filtered through GF/C filters (26). One unit is defined as the amount of enzyme catalyzing the incorporation of 1 pmol of dNTPs/min at 30 °C.

**Surface Plasmon Resonance**—A BLAcore™ instrument was used for protein binding analyses. CM5 research grade sensor chips were used for all experiments. The carboxymethyl dextran matrix by injecting over the chip in 10 mM sodium acetate (pH 4.5) buffer at 0.2 and 0.1 mg/ml, respectively. IgG was used to partially block the negatively charged carboxyl groups on the sensor chip surface.

**Unreacted NHS ester groups were inactivated using 1 M ethanolamine-HCl (pH 8.5). Typically, 2000 response units (RU) of streptavidin were purified from others by using Ni2+-NTA chromatography with the same buffers used for N-Δ1r except that urea was added to the binding and washing buffer at 8 M final concentration. These purified biotinylated fragments were resolved by SDS-polyacrylamide gel electrophoresis and transferred onto polyvinylidene difluoride membranes in 10 mM CAPS (pH 11.0) and 10% methanol at constant current (0.4 A) for 3 h. The membranes were washed with 20% aqueous methanol and then subjected to N-terminal sequence analysis using standard Edman chemistry (James McManaman, University of Colorado Cancer Center Protein Core Laboratory).

**Protein Determinations**—Concentrations of all purified proteins were determined by UV spectroscopy using their extinction coefficients. Concentrations of r fusion proteins were determined by the Pierce Coomassie Plus Protein Assay Reagent according to the manufacturer’s instructions. Bovine serum albumin (fat-free; Sigma) was used as a standard.

**DNA Polymerization Assays**—Activities of r fusion proteins were measured by their requirement for reconstitution of holozyme activity and by measuring DNA synthesis from a primed M13Gori template (26). Assay mixtures (25 μl) contained 500 pmol of M13Gori (as nucleotides), 185 units (40 ng) of DnaG primase, 1.6 μg of E. coli single-stranded DNA binding protein, 250 fmol each of DNA polymerase III core (αεθ), β, β′, γ, and the test r fusion protein (20–50 fmol). Reactions were performed in a buffer containing 50 mM HEPES-KOH (pH 7.5), 10% (v/v) glycerol, 100 mM potassium glutamate, 10 mM dithiothreitol, 10 mM Mg(CH3COO)2, 200 μg/ml bovine serum albumin, 0.02% (v/v) Tween 20, 48 μM dATP, 48 μM dGTP, 48 μM dCTP, 18 μM [3H]dTPP (specific activity, 520 cpm/pmol TTP), and 200 μM rNTP. Assay mixtures were incubated at 30 °C for 5 min, quenched by trichloroacetic acid precipitation, and then filtered through GF/C filters (26). One unit is defined as the amount of enzyme catalyzing the incorporation of 1 pmol of dNTPs/min at 30 °C.

**Surface Plasmon Resonance**—A BLAcore™ instrument was used for protein binding analyses. CM5 research grade sensor chips were used for all experiments. The carboxymethyl dextran matrix by injecting over the chip in 10 mM sodium acetate (pH 4.5) buffer at 0.2 and 0.1 mg/ml, respectively. IgG was used to partially block the negatively charged carboxyl groups on the sensor chip surface. Unreacted NHS ester groups were inactivated using 1 M ethanolamine-HCl (pH 8.5). Typically, 2000 response units (RU) of streptavidin were immobilized. The biotinylated r proteins were then injected over the immobilized streptavidin in HBS buffer (10 mM HEPES, pH 7.4, 150
mm NaCl, 3.4 mM EDTA, and 0.005% P-20 surfactant). For kinetic analyses, less than 200 RU of protein were recovered at 85% purity after Ni$_2^+$-NTA chromatography. Lysates (FrI) were prepared from 21 g of overproducing cells. Both C(0)$_\tau$ and N-$\Delta1\tau$ were purified via Ni$_2^+$-NTA metal chelating chromatography (Table II). The hexahistidine sequence within the fusion peptide was added to the new N terminus. The fusion peptide is located at the truncation site of the truncated fusion proteins (48 and 24 kDa) which become more intense with increased digestion time (arrows indicate position in lane 1; the same products migrate faster in following lanes because of electrophoresis irregularity). Lane 15, undigested C(0)$_\tau$.

### RESULTS

**Expression and Purification of the C(0)$_\tau$ and N-$\Delta1\tau$ Proteins**—We constructed plasmids P$_{A1}$-C(0)$_\tau$ (Fig. 1A) and P$_{A1}$-N-$\Delta1\tau$ (Fig. 1B), which encode C(0)$_\tau$ and N-$\Delta1\tau$, respectively. Our nomenclature system for truncated fusion proteins indicates the number of terminal residues deleted following the Δ; the preceding N or C indicates the terminus from which the amino acids were deleted. The fusion peptide is located at the truncated terminus. N-$\Delta1\tau$, for example, indicates that one amino acid was deleted from the N terminus of the $\tau$ sequence and that the fusion peptide was added to the new N terminus. The expressed C(0)$_\tau$ and N-$\Delta1\tau$ represented ~5% of the total cell protein, as determined by densitometric scans of Coomassie-stained gels.

C(0)$_\tau$ and N-$\Delta1\tau$ were purified via Ni$_2^+$-NTA metal chelating chromatography (Table II). The hexahistidine sequence within the fusion peptide specifically interacts with Ni$_2^+$ chelated to the column resin. Lysates (FrI) were prepared from 21 g of C(0)$_\tau$ or 50 g of N-$\Delta1\tau$ expression cells. Both C(0)$_\tau$ and N-$\Delta1\tau$ were recovered at $>85\%$ purity after Ni$_2^+$-NTA chromatography. The activity peaks of the eluted fractions of both C(0)$_\tau$ and N-$\Delta1\tau$ corresponded to each of their protein peaks (data not shown). Both C(0)$_\tau$ and N-$\Delta1\tau$ were fully active compared with wild-type $\tau$ protein in DNA polymerization assays. C(0)$_\tau$ and N-$\Delta1\tau$ were the only biotinylated proteins in the corresponding eluted fractions examined by the biotin blot analysis (data not shown).

**Limited Proteolysis of C(0)$_\tau$**—Limited proteolyses were performed to identify protease-sensitive interdomain hinges of the $\tau$ protein. Different proteases that encompass a broad spectrum of substrate specificities were tested: chymotrypsin, endoprotease Glu-C (SV8), papain, subtilisin, trypsin, thermolysin, endoprotease Asp-N, and endoprotease Lys-C. We first investigated the effects of varying the protease: C(0)$_\tau$ ratios and incubation times on the observed proteolytic products. Varying incubation times distinguished the initial cleavage products and also established the differences between stable

| Protein name | Fraction | Total protein (mg) | Activity (units/mg) | Specific activity (units/mg$^{-1}$) |
|--------------|----------|--------------------|--------------------|-----------------------------------|
| C(0)$_\tau$  | I        | 300                | 150                | 2000                             |
|              | II       | 52                 | 110                | 2200                             |
|              | III      | 8                  | 47                 | 5700                             |
| N-$\Delta1\tau$ | I        | 1280               | 40                 | 320                              |
|             | II       | 220                | 29                 | 1300                             |
|             | III      | 35                 | 20                 | 6000                             |

**Fig. 2.** Proteolysis of C(0)$_\tau$ with chymotrypsin. C(0)$_\tau$ was subjected to limited proteolysis by two concentrations of chymotrypsin for varying times. Each lane contains 3 μg of C(0)$_\tau$ protein. After digestion, products were separated by 10–17.5% SDS-polyacrylamide gel and stained with Coomassie Blue. Arrows on the left indicate two examples of cleavage products (48 and 24 kDa) which become more intense with increased digestion time (arrows indicate position in lane 1; the same products migrate faster in following lanes because of electrophoresis irregularity). Lane 15, undigested C(0)$_\tau$.

**Fig. 3.** Biotin blots of C(0)$_\tau$ and C-$\Delta13\tau$ proteolysis products. After digestion, samples were boiled immediately with the addition of SDS sample buffer, resolved on 10–17.5% SDS-polyacrylamide gel, and stained gels. Arrows indicate two examples of cleavage products (48 and 24 kDa) which become more intense with increased digestion time (arrows indicate position in lane 1; the same products migrate faster in following lanes because of electrophoresis irregularity). Lane 15, undigested C(0)$_\tau$.
and unstable fragments. Results from a typical experiment employing chymotrypsin proteolysis are shown in Fig. 2. At short incubation times, 56-, 52-, 48-, and 24-kDa products were observed along with full-length \( \tau \) (Fig. 3, lanes 1–3 and 8–10). At longer incubation times, the 56- and 52-kDa products were diminished, whereas the 48- and 24-kDa products and several small bands (<20 kDa) became more intense (Fig. 2, lanes 4–7 and 11–14).

Similar experiments were performed for six proteases, and optimal time points were selected for each of them. After separation on SDS-polyacrylamide gels, the digested products were transferred to membranes. Biotin blot analyses were used to identify terminal fragments (Fig. 3A). Several cleavage products resulted from each protease digestion. Several bands with similar mobilities were generated via digestion with different proteases, suggesting that certain regions of C(0)\( \tau \) were subject to cleavage by multiple proteases. For example, bands of roughly 38 kDa were obtained by digestion with thermolysin, papain, or subtilisin. Bands migrating at ~30 kDa were obtained with either chymotrypsin or subtilisin, and products of about 22 kDa were obtained after digestion with chymotrypsin, SV8, or papain. These observations suggested that C(0)\( \tau \) contains several protease-sensitive regions.

To facilitate mapping of the cleavage sites closer to the N terminus of \( \tau \), C-\( \Delta 2137 \), which is equivalent to the \( \gamma \) protein plus the fusion peptide at its C terminus, was subjected to limited proteolysis. The conditions for each protease were optimized as described above. Products with apparent molecular masses of 45 kDa were obtained after digestion with endoprotease Asp-N, endoprotease Lys-C, or chymotrypsin (Fig. 3B). These and other cleavage products vanished after longer endoprotease Lys-C digestion because of excessive proteolysis. Intensely staining cleavage products of both 27 and 26 kDa were obtained by digestions with either endoproteinase Asp-N or papain. Products of about 8 kDa were obtained after digestion with either SV8 or chymotrypsin (Fig. 3B).

Seven biotinylated fragments were selected for N-terminal amino acid sequencing analysis for the identification of their cleavage sites. The 22- and 30-kDa chymotryptic fragments were chosen for analysis, as was the 38-kDa fragment resulting from digestion with papain (Fig. 3A). Several products obtained from some of the more specific proteases were also subjected to sequence analyses. These included the 27- and 45-kDa products generated via endoproteinase Asp-N treatment, as well as the 8-kDa product from the SV8 digestion (Fig. 3B). Further, a 20-kDa cleavage product generated by chymotrypsin digestion after prolonged incubation time (data not shown) was also evaluated. Eight corresponding cleavage sites were identified (Fig. 4). Cleavage sites N-terminal to amino acid residues 106 (Asp) and 109 (Asp) corresponded to endoproteinase Asp-N digestion products that migrated at about 45 kDa; residue 222 (Asp) was the proteolytic site resulting in the 27-kDa fragment (Fig. 3B). Papain cleaved between residues 382 (Ala) and 383 (Val) to generate the 38-kDa fragment (Fig. 3A). The 8-kDa fragment generated by SV8 (Fig. 3B) was due to cleavage between residues 407 (Glu) and 408 (Thr). The 30- and 22-kDa chymotryptic fragments (Fig. 3A) resulted from cleavage C-terminal to residues 413 (Leu) and 478 (Trp), respectively. The 20-kDa chymotryptic fragment was due to cleavage after residue 496 (Ala). These cleavage sites are consistent with the substrate specificities of the respective proteases used.

The probable cleavage site resulting in the 26-kDa endoproteinase Asp-N restriction product was deduced based on the substrate specificity of that protease and the distribution of aspartate residues in the sequence of \( \tau \). Taking an experimental error of ±10% for molecular mass determination via SDS-polyacrylamide gel electrophoresis into account, the cleavage site resulting in the 27-kDa fragment would be predicted to be roughly 9 residues C-terminal to that of the 26-kDa fragment (Asp\( ^{225} \)). There are two aspartates C-terminal to Asp\( ^{225} \), those at positions 229 and 245. Thus, the 26-kDa endoproteinase Asp-N fragment is likely due to cleavage N-terminal to Asp\( ^{225} \). Similarly, in consideration of the substrate preferences of papain in light of the known cleavage site of the 27-kDa endoproteinase Asp-N product, it seems likely that the 26-kDa papain digestion product is due to cleavage of bond(s) involving Glu\( ^{228} \) and/or Glu\( ^{230} \). These probable cleavage sites (228, 229, and 230) together with the eight cleavage sites determined by sequence analysis all cluster at four regions of the \( \tau \) subunit: amino acid residues 106–109, 222–230, 383–413, and 478–496.

The N-terminal half of \( \tau \) shares high sequence similarity with \( \delta \) (28–30), another component of the DnaX complex. The crystal structure of \( \delta \) contains three domains, and sequence
alignment predicts a similar three-domain structure for the N-terminal half of τ (30). The majority of the cleavage sites we observed agree and extend this prediction to a five domain model for τ: amino acid residues 1–179 as domain I, 180–221 as domain II, 230–382 as domain III, 413–496 as domain IV, and 497–643 as domain V (Fig. 4).

Expression and Purification of Biotin and Hexahistidine-tagged DnaX—C-τ, the unique C-terminal portion of τ, is required for binding to α (15). C-τ contains the majority of predicted domain IV and the entire domain V. The fusion protein N-Δ413τ (domains IV and V) was expressed and purified so that it could be used as a tool for mapping the α binding domain of the τ subunit. We also expressed fusion proteins corresponding to domain V by itself (N-Δ496τ), and domains I-IV (C-Δ147τ).

The expression levels of C-Δ147τ, N-Δ413τ, and N-Δ496τ were ~3, 1, and 5% of total cell protein, respectively. All three of these fusion proteins were soluble. After purification by Ni²⁺-NTA chromatography, C-Δ147τ and N-Δ496τ were obtained at greater than 80% purity, and N-Δ413τ was obtained.
at over 65% purity as determined by scanning densitometry (Fig. 5). 11.5 mg of C-Δ147τ, 8 mg of N-Δ496τ, and 5.3 mg of N-Δ413τ were purified from 1250, 330, and 1200 mg of total protein from cell lysates, respectively. C-Δ147τ, N-Δ413τ, and N-Δ496τ were the only biotinylated proteins in the corresponding eluted fractions examined by the biotin blot analysis. Thus, the biotinylated τ fusion proteins were presumed to be the only proteins captured onto the BIAcore sensor chip during the immobilization step. C-Δ147τ is as active as C(0)τ in DNA polymerization assays.

The C-terminal Domain of τ Binds to α—The interaction between α and C(0)τ was first assessed using BIAcore methodology. Streptavidin was chemically coupled to the CM5 sensor chip, and C(0)τ was immobilized via biotin-streptavidin interaction. Dilutions of α were injected over and bound the immobilized C(0)τ (Fig. 6A). The off rate (k_{off}) was determined after saturating α with α to eliminate artifacts arising from reassociation. The calculated K_{d} was ~4 nM (Table III).

The interaction between α and N-Δ143τ was characterized by an association rate similar to that observed for C(0)τ-α binding; however, its dissociation rate was at least 1000-fold slower (Fig. 6B). This indicated that residues sufficient for specific binding to the α subunit lie within the C-terminal 230 amino acid residues of τ. C-Δ147τ (domains I-IV) and N-Δ496τ (domain V) were then immobilized on the streptavidin chips separately to further limit the α-binding region. Binding to the α subunit was observed with N-Δ496τ but not C-Δ147τ (Fig. 6, C and D), strongly suggesting that domain V functions as the α-binding component of the τ subunit. The off rates (k_{off}) of both α-N-Δ496τ and α-N-Δ413τ interactions were extremely slow (Fig. 6, B and C). Within the 30-min dissociation period, the response unit changes were within the machine noise level (10 RU), precluding the calculation of a dissociation rate constant, but permitting limits to be placed on the off rates and the III–V domains (Table III).

The interaction between α and N-Δ1τ was examined to investigate whether the weaker interaction between C(0)τ-α might be due to interference by the proximity of the peptide tag component of C(0)τ to the α-binding domain. The N-Δ1τ-α interaction was characterized by a dissociation rate similar to that observed for the N-Δ496τ-α interaction but much slower than that detected for the C(0)τ-α interaction (Fig. 6E). This suggests that the C-terminal tag interferes with the α-τ interaction. We also examined the C(0)τ-α interaction in the presence of the auxiliary subunits δ, δ', χψ, and ATP. The presence of δ, δ', χψ, and ATP did not make obvious changes in the rate of C(0)τ-α interaction, suggesting that the δδ'χψ-τ interaction and α-τ interaction are independent events.

The calculated stoichiometry of N-Δ496τ-α (1:0.65) was similar to that obtained for the full-length τ standards (1:0.75 and 1:0.65 for N-Δ1τ and C(0)τ, respectively). This indicates that domain V, expressed alone, is properly folded and functional.

DISCUSSION

We report here an effective approach to mapping heterologous protein-protein interacting domains. We improved upon our previously reported mapping method (23) by combining it with limited proteolysis to more precisely identify probable interdomain linkers. This modification improves the probability that the expressed proteins will contain intact domains. Using C-terminal biotin-tagged proteins in our limited proteolysis approach allowed us to distinguish terminal fragments from internal fragments by using streptavidin-conjugated detection reagents. The biotin tag also enabled rapid purification of the terminal fragment for sequence analysis and precise localization of the cleavage sites.

We chose eight proteases with different substrate specificities to identify potential interdomain hinges. As exposed sequences, interdomain hinges are generally susceptible to cleavage by more than one protease. Digestion was monitored as a function of time to help distinguish relatively stable products. Cleavage products common to more than one protease were selected for identification of the precise sites of cleavage. Eleven cleavage sites were obtained following N-terminal amino acid sequencing of the selected proteolysis products. The cleavage sites were clustered within four regions: 106–109, 222–230, 383–413, and 478–496.

The δ' subunit shares a 34% sequence similarity with the N-terminal region of both τ and γ and aligns to the N-terminal 370-amino acid sequence of DnaX proteins (28, 30), which corresponds to over half of the τ sequence. There are several known motifs conserved between these two proteins, including the Walker A (GXXXGTK), Walker B (DEXX), zinc module (CXXCXXCXXC), and SRC sequences (Fig. 4) (30). The δ' crystal structure showed three domains in a C-shaped configuration, suggesting that the N-terminal half of τ also contains

---

**Table III**

| Name   | Domain | Association rate constant (k_{on}) | Dissociation rate constant (k_{off}) | Dissociation constant (K_{d}) | Stoichiometry ratio (α/τ) |
|--------|--------|------------------------------------|-------------------------------------|-------------------------------|--------------------------|
| C(0)τ  | I–V    | 3.7 x 10^5                         | 1.3 x 10^{-3}                       | 4 x 10^{-9}                  | 0.65                     |
| N-Δ1τ  | I–V    | 9.5 x 10^5                         | <1 x 10^{-5}                       | 1 x 10^{-11}                 | 0.75                     |
| N-Δ413τ| IV, V  | 2.8 x 10^6                         | 1 x 10^{-5}                        | <4 x 10^{-12}                | 0.67                     |
| N-Δ496τ| V      | 2.9 x 10^6                         | <1 x 10^{-5}                       | <4 x 10^{-12}                | 0.65                     |
| C-Δ147τ| I–IV   | No detectable binding              |                                     |                               |                          |

a The stoichiometry of α/τ was determined by the following equation: stoichiometry = α response units/τ response units × mM/μM. In each case, molecular weight values corresponding to monomeric forms were used.

b The k_{on} cannot be determined from the dissociation phase because the derivative is too small. The limits were set according to the BIAcore System Manual.

c The discrepancy between the 10 μM (for N-Δ1τ-α) obtained here and the 70 μM (for αΔ1-τ) previously reported (23) could be due to inaccuracies in the determination of association and dissociation rates. The association rates were fast enough and the dissociation rates slow enough that both approached the limits of the technique used. The differences observed between the two carefully performed studies are likely an indication of the precision. The studies here are not intended to determine precise K_{d} values but to compare interaction quantitatively to be certain that significant binding energy is not being lost along with a deleted domain.
α Binding Domain of τ

three domains, based on their sequence homology (30). Their alignment suggested that τ amino acid residues 1–179 were domain I and that residues 180–221 were domain II. Less certain was the assignment of domain III to the residues following 226. Our results support the prediction that the N-terminal end of τ and δ' have similar structures. One of the τ protein cleavage sites (Asp222-Gln223) is located within the stretch predicted to be a hinge between domains II and III. The following cleavage site (Ala382-Val383) occurs just 13 residues C-terminal to the end of the τ-δ' alignment and may define the C-terminal boundary of the third domain of τ. We observed no cleavage between domains I and II, suggesting that this hinge was not accessible for proteolysis under the conditions employed in these studies. Based on the most N-terminal cleavage site identified (Arg105-Asp106), we constructed a plasmid encoding urchin amino acid residues 1–105. The resultant protein was not stable in expression cells (data not shown). Both the residue Asp106 of τ and the corresponding residue Glu95 of δ' are predicted to be in a similar position of a helical region revealed by the PHD program (data not shown). The residue Glu95 is on the surface of domain I of δ' (30). Above results suggest that τ residues 1–105 do not form an intact structural domain by themselves.

Two cleavage sites (Glu407-Thr408 and Leu413-Ala414) occur C-terminal to Ala382-Val383. These cleavage sites were identified via sequence analyses of different proteolytic fragments (Fig. 3A). 30 residues lie between the scissile bonds at Val383 and Leu413. The presence of at least three cleavage sites within this 30-amino acid stretch suggests that it is highly exposed to proteases and may exist as a nonstructured region of the protein. Therefore, no domain was assigned to the stretch flanked by residues 383 and 413. This analysis does not preclude the formation of structures resulting from association with other molecules not present in our proteolysis experiments domains IV and V were assigned to the remaining sequence at the C terminus of τ. Domain IV is composed of 17 amino acid residues in common with the τ translation product and 66 residues unique to the C-terminal sequence of the τ. Domain V corresponds to 147 amino acid residues of the C-terminal end of τ. α interacts with the C-terminal region unique to τ containing domains IV and V. To further limit the α binding domain of τ, BlAcore analysis was performed. Domain V alone bound α with domains III–V in the pm range. The interaction between α and N-Δ1τ was of a similar magnitude, indicating that domain V contains all of the binding energy for α-τ interaction. This also suggests that our domain assignments for the C-terminal sequence of τ accurately reflect both the structural and functional integrity of domain V. C-terminal tagged τ bound α with 1000-fold lower affinity, whereas tagging at the N terminus did not reduce α binding activity in the BlAcore assay. This observation is consistent with localization of the τ binding sites for α at or near the C terminus of domain V.

Based on our proteolysis data and sequence alignment with the homologous protein δ', considered in light of the δ' crystal structure, we assigned five structural domains to the τ subunit. Final proof of the domain assignments awaits more complete determination of the structure of the τ subunit.

Acknowledgment—We thank Dr. John Kurian for sharing the δ' structure and coordinates with us in advance of publication and Dr. Deborah Wilkinson-Fitzgerald for writing and editorial assistance on this and the following two papers.

REFERENCES

1. McHenry, C. S. (1988) Annu. Rev. Biochem. 57, 519–550
2. McHenry, C. S. (1991) J. Biol. Chem. 266, 19127–19130
3. Kornberg, A. (1988) J. Biol. Chem. 263, 1–4
4. Kelman, Z., and O'Donnell, M. (1995) Annu. Rev. Biochem. 64, 171–200
5. Fay, P. J., Johanson, K. O., McHenry, C. S., and Bambara, R. A. (1981) J. Biol. Chem. 256, 976–983
6. Wu, C. A., Zechnner, E. L., Reems, J. A., McHenry, C. S., and Marians, K. J. (1992) J. Biol. Chem. 267, 4074–4083
7. Wu, C. A., Zechnner, E. L., and Marians, K. J. (1992) J. Biol. Chem. 267, 4030–4044
8. McHenry, C. S., Griepp, M. A., Tomaszewicz, H., and Bradley, M. (1989) in Molecular Mechanism in DNA Replication and Recombination, pp. 115–126, Alan R. Liss, Inc., New York
9. Tsukihashi, Z., and Kornberg, A. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2516–2520
10. Binkowa, A. L., and Walker, J. R. (1990) Nucleic Acids Res. 18, 1725–1729
11. Flower, A. M., and McHenry, C. S. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 3713–3717
12. McHenry, C. S. (1982) J. Biol. Chem. 257, 2657–2663
13. Studwell-Vaughan, P. S., and O'Donnell, M. (1993) J. Biol. Chem. 268, 11785–11791
14. Kim, S., Dallmann, H. G., McHenry, C. S., and Marians, K. J. (1996) J. Biol. Chem. 271, 21406–21412
15. Dallmann, H. G., Kim, S., Marians, K. J., and McHenry, C. S. (2000) J. Biol. Chem. 275, 15512–15519
16. Kim, S., Dallmann, H. G., McHenry, C. S., and Marians, K. J. (1996) Cell 84, 643–650
17. Yuzhakov, A., Turner, J., and O'Donnell, M. (1996) Cell 86, 877–886
18. Glover, B. P., and McHenry, C. S. (2000) J. Biol. Chem. 275, 9017–9020
19. Pritchard, A. E., Dallmann, H. G., Glover, B. P., and McHenry, C. S. (2000) EMBO J. 19, 6536–6545
20. Kelman, Z., Yuzhakov, A., Andjelicovic, J., and O'Donnell, M. (1998) EMBO J. 17, 2436–2449
21. Glover, B. P., and McHenry, C. S. (1998) J. Biol. Chem. 273, 23476–23484
22. Kim, S., Dallmann, H. G., McHenry, C. S., and Marians, K. J. (1996) J. Biol. Chem. 271, 4315–4318
23. Kim, D. R., and McHenry, C. S. (1996) J. Biol. Chem. 271, 20690–20698
24. Lanzer, M., and Bujard, H. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8973–8977
25. Dallmann, H. G., Thimmig, R. L., and McHenry, C. S. (1995) J. Biol. Chem. 270, 29555–29562
26. Cull, M. G., and McHenry, C. S. (1995) Methods Enzymol. 262, 22–35
27. Olsen, M. W., Dallmann, H. G., and McHenry, C. S. (1995) J. Biol. Chem. 270, 29570–29577
28. Carter, J. R., Franken, M. A., Aebbersold, R., and McHenry, C. S. (1993) J. Biol. Chem. 268, 20690–20698
29. O'Donnell, M., Onrust, R., Dean, F. B., Chen, M., and Hurwitz, J. (1993) Nucleic Acids Res. 21, 1–3
30. Guenther, B., Onrust, R., Sal, A., O'Donnell, M., and Kurian, J. (1997) Cell 91, 335–345