A Three-domain Structure for the δ Subunit of the DNA Polymerase III Holoenzyme δ Domain III Binds δ′ and Assembles into the DnaX Complex*

James M. Bullard‡, Arthur E. Pritchard‡, Min-Sun Song§, Bradley P. Glover§, Anna Wieczorek§, Joe Chen§, Nebojsa Janjic‡, and Charles S. McHenry‡§†

From ‡Repidyne, Inc., Denver, Colorado 80206 and the §Department of Biochemistry and Molecular Genetics, University of Colorado Health Sciences Center, Denver, Colorado 80262

Using ψ-BLAST, we have developed a method for identifying the poorly conserved δ subunit of the DNA polymerase III holoenzyme from all sequenced bacteria. This approach, starting with Escherichia coli δ, leads not only to the identification of δ but also to the DnaX and δ′ subunits of the DnaX complex and other AAA⁺ class ATPases. This suggests that, although not an ATPase, δ is related structurally to the other subunits of the DnaX complex that loads the β sliding clamp processivity factor onto DNA. To test this prediction, we aligned δ sequences with those of δ′ and, using the start of δ′ Domain III established from its x-ray crystal structure, predicted the juncture betweenDomains II and III of δ. This putative δ Domain III could be expressed to high levels, consistent with the prediction that it folds independently. δ Domain III, like Domain III of DnaX and δ′, assembles by itself into a complex with the other DnaX complex components. Cross-linking studies indicated a contact of δ with the DnaX subunits. These observations are consistent with a model where two γ subunits and one each of γ, δ′, and δ subunits mutually interact to form a pentameric functional core for the DnaX complex.

All cellular replicases contain three major subassemblies: a sliding clamp processivity factor, an ATP-dependent clamp loader, and a special replicative polymerase that associates with these components. In Escherichia coli, the prototype for bacterial replicases, the DnaX complex clamp loader of the holoenzyme (where holoenzyme means DNA polymerase III holoenzyme) contains two DnaX proteins, γ and γ′, and the homologous δ′ subunit that shares a similar structure (1). δ′ arises by translational frame-shifting, leading to a three-domain protein. Translation of full-length dnaX mRNA leads to a five-domain γ protein that, in addition to sharing the properties of γ, binds to the DnaB helicase and has two polymerase III subassemblies together, effectively organizing the replication fork (2–7). A γ-DnaB contact stimulates helicase activity so that it can separate DNA strands at the physiological fork rate (2). An additional subunit of heretofore apparently unrelated sequence, δ, is essential for DnaX complex function both in vitro and in vivo (8).

In E. coli, the DnaX proteins are active ATPases, whereas δ′ is not, in part because of substitution of an aspartate for the essential lysine in its P-loop (1). Two additional proteins, χ and ψ, are found associated with E. coli DnaX complex (9, 10). Both proteins can interact with either γ or χ in vitro but associate exclusively with χ in the natural holoenzyme isolated from wild-type E. coli (11). χ-ψ increase the affinity of DnaX for δδ′, enabling complex formation at physiological protein concentrations (12). χ also associates with single-stranded DNA-binding proteins, stabilizing replicative complexes at elevated salt concentrations (13, 14).

DnaX exists as a tetramer free in solution by itself, but upon association with δδ′ one DnaX protomer is displaced, forming a DnaX2δδ′ complex (15). With dnaX encoding both a γ and a γ′ subunit, a wide variety of DnaX complexes assemble both in vitro and in vivo from over-produced subunits (15). Comparison of native holoenzyme isolated from wild-type cells with pure 2γ-γ′-containing polymerase III* formed by in vitro reconstitution has demonstrated that authentic holoenzyme has the composition 2γ-γ′ (16). How the assembly pathway is directed exclusively in cells to form 2γ-γ′ holoenzyme with χ-ψ bound exclusively to γ is not yet understood.

The assembly of a pentamer containing two γ subunits and one each of γ, δ, and δ′ creates a structurally asymmetric complex. This asymmetry is manifested by a functionally asymmetric holoenzyme with distinct leading and lagging strand polymerases (17–19). The leading strand can be assembled in the presence of ATPγS1 without hydrolysis. The lagging strand reaction is characterized by an absolute requirement for ATP hydrolysis and the presence of single-stranded DNA-binding proteins (19). Assembly is strictly ordered, with the leading strand occurring first and then the lagging strand. Dissociation, triggered by addition of ATPγS, leads to release of the lagging strand with the leading strand remaining stably bound to its template (19).

The structure of δ′, as determined by x-ray crystallography, has served as a model for the structure of the DnaX proteins (1). δ′ and DnaX appear as three-domain C-shaped molecules with the DnaX ATP binding site in Domain I along with a sensor I motif and a small connector Domain II containing sensor II. Other elements conserved between RFC subunits (eukaryotic clamp loader) and DNAx and δ′ reside in Domains I and II (1). Domain III in both DNAx and δ′ appears to be the

* This work was supported by National Institutes of Health Research Grant GM 35695. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
† To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Genetics, University of Colorado Health Sciences Center, Denver, CO 80262. E-mail: charles.mchenry@uchsc.edu.

1 The abbreviations used are: ATPγS, adenosine 5′-O-(thiotriphosphate); RFC, replication factor C; RU, response unit; NCBI, National Center for Biotechnology Information; ψ-BLAST, position-specific iterated basic local alignment search tool; DTT, dithiothreitol; nr, nondundant; NTA, nitrilotriacetic acid; NHS, N-hydroxysuccinimide; EDC, 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide.

Received for publication, September 10, 2001, and in revised form, January 22, 2002.

Published, JBC Papers in Press, January 24, 2002, DOI 10.1074/jbc.M108708200
major protein interaction domain as it pertains to interactions within the DnaX complex assembly. DnaX Domain III binds δ' and other DnaX proteomes, ψ-δ, and can serve to form the core of the DnaX complex in the presence of δδ' (20, 21). δ' Domain III binds δ and DnaX and effectively supports DnaX δδ' complex formation (22, 23).

Eukaryotic RFC clamp loader complexes contain five homologous RFC subunits. In contrast, DnaX complex appeared to contain only four related subunits, τ2ψγδ, and a dissimilar δ subunit. δ is poorly conserved and not readily recognizable in diverse bacteria when searches are performed using simple search tools like BLAST with a single divergent bacteria when searches are performed using simple search tools like BLAST with a single divergent bacterial genome. However, the conserved element found between the DnaX Walker A and B motifs, as identified by DnaX identification Rule 2, is not present in δ. Thus, Rule 1, δ must lack this motif to be distinguished from DnaX and validated as an authentic δ sequence. Furthermore, to permit the unbiased identification of legitimate δ' candidates preliminarily identified by ψ-BLAST, additional rules must be applied. Rule 2, seven of the 12 conserved residues must be present in the δ' RFC-like motif Val as defined by the E. coli δ' sequences 124[AI/ANAI/SI/IV/LI/V]/Kh/XL/XVI/LX/XX/EEPP126. Three of the six conserved residues must be present in the SRC motif as defined by E. coli δ' sequence 119[LI/V]/LX/XX/VTI/XV/SRC129.

Summary of Rules Used for Identification of δ

Many of the elements present in DnaX are also found in δ'. However, the conserved element found between the DnaX Walker A and B motifs, as identified by DnaX identification Rule 2, is not present in δ'. Thus, Rule 1, δ must lack this motif to be distinguished from DnaX and validated as an authentic δ' sequence. Furthermore, to permit the unbiased identification of legitimate δ' candidates preliminarily identified by ψ-BLAST, additional rules must be applied. Rule 2, seven of the 12 conserved residues must be present in the δ' RFC-like motif Val as defined by the E. coli δ' sequences 124[AI/ANAI/SI/IV/LI/V]/Kh/XL/XVI/LX/XX/EEPP126. Three of the six conserved residues must be present in the δ' SRC motif as defined by E. coli δ' sequence 119[LI/V]/LX/XX/VTI/XV/SRC129.

Summary of Rules Used for Identification of δ'

To be confirmed as δ from a list of candidates identified by ψ-BLAST, proteins must meet the following criteria in at least three of the six conserved regions. In conserved regions 1–6, the indicated number of conserved residues must be matched in alignments generated using CLUSTAL W. 1) 5 of possible 7 as defined by the E. coli δ sequence 124[AI/ANAI/SI/IV/LI/V]/Kh/XL/XVI/LX/XX/EEPP126. 2) 5 of possible 11 from 66[LI/V]/LX/XX/XX/VI/XI/VI/LI/V]/Kh/XL/XVI/LX/XX/EEPP126. 3) 5 of possible 10 from 97[LI/V]/LX/XX/XX/VI/XI/VI/LI/V]/Kh/XL/XVI/LX/XX/EEPP126. 4) 9 of possible 18 from 124[AI/ANAI/SI/IV/LI/V]/Kh/XL/XVI/LX/XX/XX/VI/XI/VI/LI/V]/Kh/XL/XVI/LX/XX/EEPP126. 5) 11 of possible 23 from 119[LI/V]/LX/XX/XX/VI/XI/VI/LI/V]/Kh/XL/XVI/LX/XX/EEPP126. 6) 5 of possible 12 conserved residues from 209[LI/V]/LX/XX/XX/VI/XI/VI/LI/V]/Kh/XL/XVI/LX/XX/EEPP126. 7) 3 of the seven conserved residues must be present in the δ' SRC motif as defined by E. coli δ' sequence 119[LI/V]/LX/XX/VTI/XV/SRC129.

Reagents and Materials

BLAcore CM5 sensor chips (research grade), P-20 surfactant, N-hydroxysuccinimide, N-ethyl-N-(3-diethylaminopropyl)carbodiimide, and ethanalamine hydrochloride were obtained from BLAcore Inc. (Piscataway, NJ).

Proteins

Bovine serum albumin was purchased from Sigma, bovine IgG from Bio-Rad, and streptavidin from Pierce. γ and δ' were purified as described previously (8, 31). Reconstituted γ complex and δ complex were prepared as previously described (11). The DNA polymerase III holoenzyme was purified according to Cull and McHenry (32) with modifications as described (11). The monoclonal antibody directed against the δ subunit (cell line 281A1) was produced in collaboration with the University of Colorado Health Sciences Center Tissue Culture and Monoclonal Antibody Facility. The δ antibody was purified from 500 ml of hybridoma supernatant by 50% ammonium sulfate reprecipitation, resuspended in 50 ml of dialyzed against phosphate-buffered saline (80 mM NaH2PO4, 20 mM NaH2PO4 (pH 7.5), 100 mM NaCl), and used in a 1:100 dilution. The benchmark prestained protein ladder was from Invitrogen. BS3 was purchased from Pierce.

Buffers

The buffers used were: Buffer NT (50 mM potassium phosphate (pH 7.8), 10% glycerol, 1 mM imidazole, 0.5 mM NaCl, 0.5 mM DTT); Buffer N (50 mM potassium phosphate (pH 7.8), 10% glycerol, 1 mM imidazole, 0.5 mM NaCl, 0.5 mM DTT); Buffer AS2 (50 mM Tris-HCl (pH 7.5), 20% glycerol, 1 mM EDTA, 0.1 mM MgCl2, 5% (v/v) glycerol, 5 mM DTT); Buffer C (5 mM sodium citrate (pH 5.0)); and Buffer H (50 mM HEPES (pH 7.4), 100 mM NaCl, 5% glycerol).
Gel Filtration

Samples containing various combinations of δ Domain III, δ', and γ were incubated in a total volume of 500 µl in buffer HMD at 15 °C for 20 min and then injected into a 24-ml Superose 6 HR 10/10 FPLC gel filtration column. Fractions of 0.4 ml were collected at a flow rate of 6 ml/h. When present, the amounts of the components in the load volume were: 0.41 nmol (0.75 µM) γ; 0.56 nmol (1.0 µM) δ, and 0.28 nmol (0.52 µM) δ Domain III. Gel filtration column fractions (0.4 ml) were precipitated using trichloroacetic acid, dissolved in sample buffer as described (23), and loaded onto an SDS-polyacrylamide 4–20% gradient gel.

Use of Surface Plasmon Resonance to Measure δ-δ' Interaction

A BIAcore instrument was used to measure Domain III-δ' binding. The carboxymethyl dextran matrix of the CM5 sensor chip was activated by the NHS/EDC coupling reaction as described previously (22). The matrix was activated using a 100-µl injection of a mixture of 0.2 M EDC and 0.05 M NHS in water to maximize the conversion of the carboxyl groups of the sensor chip surface to NHS esters. For immobilization of the biotinylated fusion protein Domain III of δ, streptavidin and bovine IgG were sequentially captured onto the activated matrix by injecting each over the chip in 10 mM sodium acetate buffer (pH 4.5) at 0.2 and 0.1 mg/ml, respectively. Bovine 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). The fusion peptide-tagged δ Domain III was then injected over the immobilized streptavidin in HBS buffer. The binding of δ Domain III to the native δ' subunit was performed in HBS buffer + 5 mM DTT. A flow rate of 5 µl/min was used for kinetic analysis. Kinetic parameters were determined using the BLAevaluation 2.1 software (BIAcore). Apparent stoichiometries were estimated using Equation 1.

\[
\text{Stoichiometry} = \frac{[R_u \text{analyte}][R_u \text{ligand}]}{([M_{\text{ligand}}][R_u \text{analyte}])}, \quad (\text{Eq. 1})
\]

where RU is the response unit obtained at binding saturation and M is molecular weight.

Estimation of Apparent \(K_D\) for Interaction of δ with DnaX,δ'

The interaction of δ Domain III in the presence of molar excesses of γ and δ' was modeled as δ Domain III + δ' DnaX \→ δ Domain III-δDnaX, where the apparent \(K_D\) = [δ Domain III][δ'DnaX]/[δ Domain III-δDnaX]. In the experiment reported in Fig. 6, no free δ Domain III (δ Domain III) was detectable, and therefore a lower limit of <0.1 µg/fraction (corresponding to 20 µM in the load volume) was assumed for purposes of calculating an upper limit for the apparent \(K_D\). Because of the dilution that occurs during gel filtration, \(K_D\) values were calculated as a range varying from the concentration of components in the applied sample (500 µl) to the diluted elution volume (~2 ml).

Biotin Blot

After separation by SDS-polyacrylamide gel electrophoresis, proteins were transferred onto polyvinylidene difluoride membranes at 0.8 mA/cm² for 1.5 h in 25 mM Tris-HCl, 192 mM glycine (pH 8.3), 20% methanol and developed as described (5).

Limited Proteolysis

C (0.8) digests were carried out using chymotrypsin in 20 mM potassium phosphate (pH 7.4), 0.4 M NaCl, 5% glyceral, 0.1 mM EDTA, 5 mM DTT. At different time points, 9.4-µl aliquots from reaction mixtures were removed, mixed with 4.7 µl of stop buffer (0.18 M Tris (pH 6.8), 30% sucrose, 6% SDS, 180 mM DTT), and then immediately boiled for 5 min. Each aliquot contained 3 µg of protein. Digestion products were separated by SDS-polyacrylamide gel and stained or transferred onto a blotting membrane as described (1).

BS3 Chemical Cross-linking

BS3 stock solutions were prepared fresh before every reaction by dissolving BS3 in Buffer C. The proteins were prepared in Buffer H with 0.5 mM DTT added to holoenzyme. An aliquot of the BS3 stock solution was added to the protein complex immediately and reacted at room temperature for 40 min. Reactions were quenched for 10 min by the addition of 3× SDS-PAGE sample buffer. DnaX complexes were cross-linked at a concentration of 0.5 mg/ml with 200 µM BS3 and purified holoenzyme at a concentration of 0.2 mg/ml with 400 µM BS3.

Purification of C-terminally Tagged δ

Construction of Vector Used for Expression of C-terminally Tagged δ—The plasmid for the expression of full-length δ with a C-tagged fusion protein was constructed by PCR amplification of the portion of the template pJRC105 (33) starting with a BamHI site upstream of the holA gene and extending through the holA gene to its stop codon. Attached to the downstream primer was a sequence for an SpeI site not present in the template. The PCR product was digested with BamHI and SpeI and inserted into the cognate sites of the plasmid pAI-BCNCOI, which contains the C-tagged fusion including sequences for hexahistidine and biotinylation. The sequence of the terminus of the resulting protein is GSTGGGGVYPRGSHHHHHHLH-LKMKMVWNNCH, where the initial Gly listed is the C-terminal residue of δ. This resulting plasmid, pCOδ, was expressed in E. coli strain BLR and purified as described (38).

Lysis and Ammonium Sulfate Fractionation—Cells (125 g) were lysed as described (32). The lysis procedure was modified by increasing the concentration of lysozyme to 2 mg/g cells and adding 5 mM EDTA. C-tagged δ in the resulting supernatant (Fraction I, 465 ml) was precipitated by the addition of ammonium sulfate (0.242 g added to each ml) and collected by centrifugation (23,000 × g, 45 min, 4 °C). The precipitate was back-washed with 0.125 × fraction I volume of Buffer AS2 and centrifuged (23,000 × g, 45 min, 4 °C). The pellet (310 mg, 4.25 × 10⁶ units) was frozen at −80 °C until ready to use.

Ni²⁺-NTA Chromatography—An Ni²⁺-NTA Superflow column (Qiagen, 1 ml resin/60 mg of fraction II protein) was equilibrated with Buffer N. The fraction II pellets were thawed on ice and centrifuged at 30,000 × g for 10 min at 4 °C. The excess ammonium sulfate supernatant was removed. Pellets were dissolved in Buffer N containing 0.4 M NaCl and diluted further with the same buffer to a conductivity of Buffer N. The protein was eluted with a 10-column volume gradient of Buffer N + 10 → 100 mM imidazole. Fractions were assayed for activity, and active fractions were pooled (fraction IIIa, 18 ml, 9.8 mg, 1.4 × 10⁶ units, ~60% pure as determined by SDS-polyacrylamide gel electrophoresis (data not shown)), placed in dialysis tubing (Spectra/Por, 110-kD molecular weight cut-off), and concentrated by placing solid polyethylene glycol 8000 (Fisher Scientific) outside of the dialysis bags (4 °C) resulting in fraction IIIb (1.0 ml, 8.0 mg, 1.1 × 10⁶ units).

Gel Filtration—One-half of fraction IIIb (0.5 ml) was loaded onto a FPLC Superose 12 (Amersham Biosciences) column equilibrated with 20 mM potassium phosphate (pH 7.4), 0.4 M NaCl, 0.1 mM EDTA, 5% glycerol, 5 mM DTT. Two protein peaks eluted, with δ being in the second, lower molecular weight peak (fraction IV, 0.6 mg, 3.4 × 10⁷ units). Fraction IV was ~75% pure (data not shown).

Expression and Purification of δ Domain III

Construction of Vector Used for Expression of δ Domain III—The plasmid for the overexpression of δ Domain III was constructed by PCR amplification of a portion of holA using the template pJRC105 (33). The amplified domain started at the holA codon for Asn-207 and extended to the holA stop codon. Attached to the upstream primer was a sequence for a SpeI site not present in the template, and following the downstream primer was added HindIII site sequence. The PCR product was digested with SpeI and HindIII and inserted into the cognate sites of the plasmid pAI-BCNCOI, which contains the N-tagged fusion protein including hexahistidine and a biotinylation recognition sequence. The sequence of the terminus of the resulting protein is MAGCLN-DIFEAQKIEWHHHHHLVPGRSGCGGGLQNN, where the terminal Asn listed is Asn-207 of δ. The final Domain III expression plasmid was transformed into E. coli strain BL21(DE3) for growth and induction, which was performed as described previously (34).

Lysis and Ammonium Sulfate Fractionation—Cells (400 g) were lysed as described (32) except that lysozyme was increased to 2 mg/g cells, 5 mM EDTA was added, and the incubation time on ice was increased to 1.5 hours in resulting fraction I (900 ml, 9.24 g protein). δ Domain III was precipitated by the addition of ammonium sulfate (0.226 g added to each ml) and collected by centrifugation (23,000 × g, 45 min, 4 °C). The precipitate was back-washed with 0.125 × fraction I volume of Buffer AS2 and centrifuged (23,000 × g, 45 min, 4 °C). The pellets were frozen at −80 °C until ready to use.

Ni²⁺-NTA Chromatography—An Ni²⁺-NTA Superflow column (Qiagen, 1 ml resin/60 mg of fraction II protein) was equilibrated with Buffer NT. The fraction II pellets were thawed on ice and centrifuged at 35,000 × g, 10 min, 4 °C to remove excess ammonium sulfate, dissolved in Buffer NT + 0.4 M NaCl, and diluted further with the same buffer to a conductivity with Buffer NT. After loading, the column was washed with Buffer NT containing 10 mM imidazole. The protein was eluted with a 10-column volume gradient in Buffer NT + 10 → 200 mM
imidazole. The fractions were pooled based on analysis in a reduced SDS-polyacrylamide gel selecting the purest fractions of the single chromatographic peak. The pool resulted in fraction III (64 mg protein, 60 ml). Fraction III was >95% pure based on densitometric scan (data not shown).

**SDS-Polyacrylamide Electrophoresis and Immunodetection—Proteins were loaded onto a 5–17.5% gradient SDS-polyacrylamide gel (0.75 × 18 × 16 cm) and separated at 65 V overnight. The separated proteins were electrophoretically transferred to Immobilon P polyvinylidene difluoride membrane at 500 mA for 6 h and blocked in MTBS (10 mM Tris-HCl). Proteins were electrotransferred to Immobilon P polyvinylidene fluoride (PVDF) and developed with the enhanced chemiluminescence method. Immunostaining was visualized using a biotinylated secondary anti-mouse antibody (1:1000 dilution in MTBS).**

**Experimental Procedures**

James Walker and colleagues (30) performed a nearly comprehensive alignment of DnaX. From their work, we took seven conserved regions and selected conserved residues to form rules used to identify any DnaX protein and to unambiguously distinguish them from δ. DnaX proteins contain a functional Walker A and B motif and a DnaX-specific sequence conserved between these elements. Additionally, they contain a conserved RFC motif VIc, a functional RFC-like SRC sequence, and conserved sensors 1 and 2. The precise “rules” that we use for DnaX identification, derived from the Walker data, are described under “Experimental Procedures.”

To determine the elements conserved within δ sequences that distinguish it from DnaX and δ, we performed alignments of a representative subset of δ sequences from discrete divisions of phyla of sequenced bacteria (Fig. 1). The core consensus Walker A motif is GXXGXGK/T/S (35). Examination of the aligned δ sequences indicates *Aquifex aeolicus* and *Synechocystis* could potentially have an active ATPase based on this criterion. The Walker B sequence found in some classes of ATPases is DEXX with the sequence being confined to an invariant DEAH within DnaX (30). We note that both *A. aeolicus* and *Synechocystis* do have dual acidic residues at the beginning of their Walker B motifs and could possibly bind and hydrolyze ATP. Among the remaining 27 δ sequences examined, none have both intact consensus Walker A and B motifs.

δ sequences contain apparently functional sensor 1 and 2 motifs (1) that could be used to couple ATPase activity to mechanical work. Additionally, δ sequences contain RFC motif VIc and SRC elements. To distinguish them from DnaX sequences, it is important to note that δ sequences do not contain a motif IV (30) as defined for DnaX under “Experimental Procedures” (DnaX rule 2).

**E. coli** δ contains a zinc binding element composed of four Cys residues (CXXCXXCXXC). We note that this element is shared in some organisms but is lacking from *Chlamydia trachomatis*, *Helicobacter pylori*, *Rickettsia prowazekii*, and *Synechocystis* (Fig. 1), as well as *Borrelia burgdorferi*, *Campylobacter jejuni*, *Caulobacter sp.*, *Chlamydia muridarum*, *Chlamyphila pneumoniae*, *Mesorhizobium loti*, *Thermotoga maritima*, *Thermus thermophilus*, *Ureaplasma urealyticum*, and *Xylella fastidiosa* (data not shown). When the element is present, the spacing between the first two Cys residues varies between 7 and 12, consistent with the variation observed among DnaX proteins (30). Thus, this element does not appear to be essential for δ function in general. We note that although they lack the Zn\(^{2+}\) binding element, *B. burgdorferi* and *X fastidiosa* contain one Cys residue in positions corresponding to the *E. coli* δ Zn\(^{2+}\) finger, suggesting but not proving that the ancestral sequences contained the Zn\(^{2+}\) finger and it was lost with retention of some residues. These observations have been distilled into a defined set of rules for identification of δ sequences (see “Experimental Procedures”).

### Identification of Conserved δ Sequences and Establishing Rules for the Unambiguous Identification of δ Sequences from...
FIG. 1. Sequence elements conserved between DNA polymerase III holoenzyme \( \beta \) subunits. A sequence alignment of a subset of the 34 \( \delta \) proteins identified from the finished genomes in the nr database at NCBI was performed. This subset represents \( \delta \) protein sequences from one each of the phyla of sequenced organisms in the nr database. The underlines indicate the two conserved regions used as criteria for the following analysis.
**Sequenced Organisms**—Our ability to identify δ sequences from all organisms afforded an opportunity to determine residues conserved among δ sequences. We made an alignment of all δ sequences and show a subset of diverse bacteria (Fig. 2). Six conserved stretches were identified that can be used to unambiguously identify δ sequences from candidate sequences identified from ψ-BLAST (see “Experimental Procedures”). We note that no subunits contain the Zn\(^{2+}\) binding element found in many DnaX and δ' sequences.

**Identification of Putative δ Domain III from Sequence Alignments with δ'**—The x-ray crystal structure of the δ' subunit of the *E. coli* DNA polymerase III holoenzyme has been solved, indicating a C-shaped three-domain protein (1). Kurian and colleagues (1) predicted, by sequence alignments, that the DnaX protein of the DNA polymerase III holoenzyme would have a similar structure. This prediction was supported, in part, by our ability to express a DnaX Domain III fusion protein that defines all of the stable protein-protein binding interactions in which the γ DNA subunit participates (20). Both DnaX and δ' Domain III proteins can be expressed, purified, and assembled into a complex with the remaining DnaX proteins in the absence of their respective Domains I and II (8, 20–22).

Because the ψ-BLAST search used to identify δs from all sequenced organisms also resulted in the identification of the DnaX and δ' members of the DnaX complex, we expected that δ had the same overall fold defined by the δ' prototype. If this prediction was true, an alignment of a large number of δ and δ' sequences should predict the start of Domain III of δ. This putative Domain III, if an independent folding domain, should express to high levels in *E. coli* in the absence of Domains I and II and would be expected to bind to δ' by itself and assemble into the DnaX complex. As a first step toward testing these predictions, we performed an alignment of 34 δ and 34 δ' sequences from the completely sequenced bacteria in the NCBI nonredundant (nr) data base. A subset of the aligned sequences from diverse organisms reflects the features of the comprehensive alignment (Fig. 3). Asp-208 of *E. coli* δ' is the first residue of α-helix 10 that defines the start of δ' Domain III (1). The corresponding residue in *E. coli* δ is Val-206. We constructed an expression vector that fused a biotinylation/hexahistidine sequence to the putative Domain III of δ, starting at Asn-207 as described under "Experimental Procedures."

**Putative δ Domain III Is Expressed to High Levels and Is Soluble**—The recombinant candidate δ Domain III protein was expressed to high levels, consistent with it being an independently folding domain. The resulting protein was purified by Ni\(^{2+}\)-NTA chromatography and gel filtration, resulting in a protein that was ~95% pure (see "Experimental Procedures").

**C-terminally Tagged δ Is Cleaved by Chymotrypsin, Generating Fragments At or Near the Predicted Domain II–III Junction**—We expressed and purified full-length δ tagged at its C terminus with a peptide that contains a hexahistidine sequence and a sequence that is biotinylated in *vivo* (see "Experimental Procedures"). C-tagged δ was subjected to limited chymotryptic hydrolysis, and the tagged products were displayed on "bacterial blots" (Fig. 4). The predicted start of Domain III (Fig. 3) is followed five amino acids later by a Phe residue that might provide a substrate for chymotrypsin cleavage. A cleavage product of ~19 kDa is observed, consistent with the 18.6-kDa product expected with cleavage after Phe-210. The next closest potential earlier cleavage in the sequence (Trp-192) would have generated a product 2.1 kDa larger. An additional major product is observed at ~25 kDa, which could result from a cleavage at the Domain I–II junction (Fig. 4). The most likely residue cleaved in the δ sequence that would result in a tagged product of ~25 kDa is Trp-150 (25,402 Da).

**The δ' Binding Site is Contained within δ Domain III**—Purified δ Domain III, immobilized on a BIAcore chip via its N-terminal biotinylation tag, formed a complex with δ' (Fig. 5). Varying concentrations of δ' were injected over the immobilized Domain III (Fig. 5). The calculated *K*ₐ (1.7 μM, Table II) is only 2- to 3-fold weaker than the interaction of full-length δ with δ' (22). Thus, Domain III of δ, like Domain III of the related δ' subunit, is responsible for holding the two subunits together within the DnaX complex (22).

**The Site Required for δ to Assemble into the DnaX Complex Is Contained within Domain III—δ does not form a detectable stable complex with DnaX alone but forms a tight DnaXδ' complex when combined with both δ' and DnaX (23). Complex formation is much stronger than with δ' alone, suggesting a conformational change in δ' induced by δ binding or a weak δ-DnaX contact that leads to further stability of the three-protein complex. We observe that δ Domain III can form a tight three-protein complex with DnaX and δ' (Fig. 6). From gel filtration we estimate the apparent *K*ₐ of DnaX δ' with δ Domain III to be at least 5–20 nM. This binding is apparently stronger than the 114 nM value measured for the apparent *K*ₐ for interaction of γ, full-length δ, and δ' (20). Although it is possible that strain is relieved upon removal of δ Domains I and II, to confirm this possibility will require more rigorous approaches such as equilibrium sedimentation or calorimetry. We also note that the elution position of γ actually decreases upon binding δ' and δ Domain III (Fig. 6). This suggests a more hydrodynamically compact structure is formed in the γ₁ to γ₃ transition during complex formation (15).

**Domains I and II of δ Are Required for the DnaX Complex Function**—We tested whether the "DnaX" complex that lacked the δ N-terminal domains was active in assembling a functional replicase. We found that it was not, which is consistent with the requirement of Domains I and II of the related DnaX and δ' proteins for reconstitution of active DnaX complexes (data not shown). Although Domains I and II of DnaX, δ', and δ are not required for formation of stable complexes, apparently
FIG. 2. Sequence elements conserved between DNA polymerase III holoenzyme δ subunits. Sequence alignment of a subset of the 34 δ proteins identified from the nr sequence database. This subset represents δ protein from one each of the phyla of sequenced bacteria represented in the nr database. The underlines indicate the six conserved regions present in all δ proteins and used as the criteria for identifying δ. The aligned

Conserved region 1

Conserved region 2

Conserved region 3

Conserved region 4

Conserved region 5

Conserved region 6

The aligned
their interaction is essential for DNA complex to transfer a β processivity factor onto DNA, enabling processive DNA replication.

δ Domain III binds δ and DNA

their proteins are from the same bacteria described in Fig. 1. The amino acids indicated by white letters on a black background are identical, and those with black letters on a gray background are similar. Also identified, but not shown in the figure, were the following sequences (with GenBankTM accession numbers): Bacillus halodurans, BAB05056; Buchnera, P57520; Borrelia burgdorferi, P70156; Campylobacter jejuni, E81411; Caulobacter crescentus, AAK25713; Chlamydia muridarum, H81715; Chlamydia pneumoniae, D72086; Escherichia coli O157, BAB34101; Hemophilus influenzae, P43747; Helicobacter pylori, D82904; Lactococcus lactis, D72086; Mycobacterium leprae, Q50362; Neisseria meningitides, NP03741; Neisseria gonorrhoeae, NP03741; Pasteurella multocida, AAK03300; Pseudomonas aeruginosa, G83148; Thermotoga maritima, H72407; Treponema pallidum, D82260; Vibrio cholerae, D82590; Xylella fastidiosa, D82590.
13254

**TABLE II**

Kinetic and equilibrium constants of δ Domain III–δ interaction determined by surface plasmon resonance

| δ concentration (μM) | $k_{on}$ (μM$^{-1}$ s$^{-1}$) | $k_{off}$ (s$^{-1}$) | $K_D$ (μM) |
|----------------------|--------------------------|----------------|-----------|
| 0.5                  | $4.4 \times 10^4$       | $6.8 \times 10^{-2}$ | 1.6       |
| 2                    | $3.3 \times 10^4$       | $7.9 \times 10^{-2}$ | 2.4       |
| 5                    | $6.1 \times 10^4$       | $6.2 \times 10^{-2}$ | 1.0       |
| Average              | $4.6 \pm 1.4 \times 10^4$ | $7.0 \pm 0.9 \times 10^{-2}$ | $1.7 \pm 0.7$ |

$K_D$ was determined by using the measured rate constants ($K_D = k_{on}/k_{off}$). The $K_D$ for the δ–δ’ interaction determined using the BIAcore was 0.86 ± 0.07 μM (22).

**FIG. 6. δ Domain III assembles into a DnaX,δδ’ complex.** The indicated proteins were mixed and gel-filtered by Superose 6 FPLC, and selected fractions were electrophoresed on SDS-polyacrylamide gels as described under “Experimental Procedures.” Column fractions (Fr) are identified above the Coomassie Blue-stained gels. When δ’ was run alone (data not shown), the peak of its elution was fraction 44.

**FIG. 7. δ Cross-links to DnaX.** γ complex (2 μg, lane 1), τ complex (2 μg, lane 2), and holoenzyme (14 μg, lane 3) were cross-linked with BS3 as described under “Experimental Procedures,” immunoblotted with an anti-δ monoclonal antibody, and developed with the ECL chemiluminescence method. The migration position of molecular size markers is designated in the right margin. Cross-links are apparent in the holoenzyme lane (Holo) for the position of the γδ (88 kDa), τδ (117 kDa), and higher order δ-containing cross-links. The ~109-kDa band likely results from cross-linking of δ to additional components, perhaps in a complex containing three proteins, and was not used for interpretation of the position of components within the DnaX complex. Cross-linked proteins run slightly larger than standards of the same molecular weight on SDS gels because they are unable to unfold into a random SDS-coated structure.

Detection of a δ-DnaX Contact by Chemical Cross-linking—We know that Domain III of DnaX is responsible for DnaX oligomerization and for binding to δ (20, 21). Similarly, Domain III of δ’ is responsible for contact with DnaX and δ (22, 23). The preceding results indicate δ Domain III binds δ’ and participates in the apparently cooperative assembly of DnaX complex. This could occur by binding to δ’, altering its conformation so that it interacts more strongly with DnaX, or by making a direct contact with DnaX, contributing binding energy to complex formation directly. To determine whether δ directly contacts DnaX (or whether it has a surface lysine within 11.4 Å of a surface lysine on DnaX), we performed cross-linking studies with BS3, a reagent previously used to show the γ subunit within native holoenzyme binds the δ’ and ψ subunits exclusively. Our results indicate that δ cross-links to the γ subunit within the γ complex (Fig. 7, lane 1), the τ subunit within τ complex (Fig. 7, lane 2), and to both γ and τ within native DNA polymerase III holoenzyme that contains a $\gamma_2\psi$ DnaX core (Fig. 7, lane 3). This result indicates a probable δ-DnaX contact and formation of a pentameric core of the essential subunits of the DnaX complex, all contacting one another principally through the interactions of their respective domains.

**DISCUSSION**

Although the δ subunit of the DNA polymerase III holoenzyme is not well conserved, using the power of the ψ-BLAST program, a subtle pattern emerges that permits the identification of δ from all organisms. ψ-BLAST establishes a more powerful search pattern with each iteration, as additional members of a protein family are included. This subtle pattern is contained within the more highly conserved DnaX and δ’ family members, permitting their identification in the same search along with the other members of the AAA+ ATPases, which have been demonstrated to be in a common family. This led to the prediction that δ has a three-domain fold in common with δ’ and DnaX. An alignment of δ and δ’ sequences led to a prediction that Domain III of E. coli δ began at Val-206, the residue that aligns with Asp-208, the first residue of helix 10 of the δ’ Domain III.

The validity of our prediction of the existence of a δ Domain III was initially supported by our ability to express and purify a soluble fusion protein that starts with E. coli δ residue Asn-207. Additionally, proteolysis indicated cleavage, most likely at Phe-210, only five residues after the predicted start of δ Domain III. The crystal structure of the DnaX complex indicates a DnaX–δ as well as a δδ’ contact in a pentameric ring (26). Although the putative δ Domain III was based on sequence alignment of a large number of δ and δ’ subunits, the just published crystal structure of δ as part of a $\gamma_2\delta\delta'$ complex indicates that our estimate was approximately correct, with Asn-207 three residues behind the end of the last helix (α9) of δ Domain II. Thus, the expressed δ Domain III contains the entire δ Domain III preceded by the Domain II–III linker and the three final residues of Domain II. An additional proteolytic clip of C-tagged δ generated a product of ~25 kDa, most consistent with cleavage after Trp-150. This residue is located five residues from the start of α-helix 6 of δ, suggesting that both interdomain regions of δ are exposed, at least in the absence of other subunits.

δ Domain III bound nearly (within 2–3-fold) as strongly to δ’ as full-length δ, indicating that the energetically important contacts between δ and δ’ are confined to Domain III. Furthermore, δ Domain III assembles with other members of the DnaX complex, preserving the intracomplex interactions of the full-length subunit, a property shared with both DnaX Domain III and δ’ Domain III (20–23). Thus, the DnaX complex is held together by Domain III interactions between the five essential protomers, DnaX, and δδ’. The crystal structure of the DnaX complex reveals extensive contacts in all three domains of most
of the components. Apparently the Domain I and II contacts occur as a consequence of the proteins being held in close proximity by Domain III interactions and do not significantly stabilize the DnaX complex. Thus, the free energy of DnaX, δ', and δ Domains I and II solvation must be close to that gained by intersubunit contacts. These weak contacts are likely essential for DnaX complex function, as proposed by an elegant model based on the structure (26). Deletion of Domains I and II from any of the DnaX complex components results in an inactive complex (this work and Refs. 4, 5, 20, 22, and 23).

We note that N-terminally tagged δ Domain III can interact with δ' but cannot effectively assemble into a DnaX complex when immobilized by its N terminus on the surface of a BIAcore chip. N-tagged δ Domain III can assemble into a DnaX complex when free in solution, and therefore the chip surface must be creating steric hindrance that interferes selectively with binding of N-tagged δ Domain III with DnaX. Examination of the DnaX complex crystal structure indicates a contact between the face of δ Domain III that contains the δ Domains II and III linker (the site where the tag is located) and DnaX.

We previously observed that tags placed on the C terminus of δ' and derivatives interfered with assembly of δ' into DnaX complexes but not its interaction with δ (22). The DnaX complex structure indicates that the most C-terminal residue in the structure of δ' (Leu-334) contacts Domain III of γ#2. C-terminal δ' Leu-334 is not near the δ contact, which is consistent with C-terminal δ tags not interferring.

Examination of the crystal structure of δ in the context of the DnaX complex provides a topological orientation of the regions of conservation used initially to generate rules for the unambiguous identification of δ sequences (Fig. 2). Conserved region 1 is located in β-strand 2 and the loop that follows. Conserved region 2 includes α-helix 3 and the loop following that contains the highly conserved Phe-74 residue that inserts into β during the β dimer opening reaction (25). Conserved region 3 contains α-helix 4 and β-strand 5 within δ Domain I. A conserved basic residue in the loop just after the C terminus of β5 is found on the face of opposite Phe-74 that contacts β. It is solvent-exposed but may participate in an important contact in an alternative conformation found in another kinetic step of the β loading reaction. Conserved region 4 is found within δ Domain II. It encompasses β-strand 7 and α-helices 7 and 8. Near the end of helix 8 is found a conserved basic residue (Arg-187) in E. coli followed by a conserved Leu. Arg-187 (in E. coli δ) was found on the surface of δ that contacts γ#3 near residue 173. Conserved region 5 includes α-helices 10–13 within Domain III. Conserved Phe-215 is located at the start of helix 10. A conserved acidic residue (Glu-242) is located immediately before helix 12. Both Phe-215 and Glu-242 are found in the interior cavity of the DnaX complex structure exposed to solvent. They may make important contacts in an alternative conformation. Conserved region 6 encompasses the final two helices (15–16), which compose the distal end of δ. The residues conserved within helices 15–16 are primarily on the interior surface of these two helices, suggesting they are conserved from intra- rather than intermolecular interactions. An exception is an acidic residue (Asp-315) found in a loop just between these two helices, which is located, solvent-exposed, just under the pore created by assembly of the five subunits of the DnaX, δδδδδ complex.

The DnaX complex in native holoenzyme has the composition τγδδδ' ψψ(16). Using chemical cross-linking, we have previously demonstrated that δ' and ψ are exclusively bound to γ in native holoenzyme despite their ability to associate with either the τ or γ DnaX proteins in complexes assembled in vitro (11). Using cross-linking, we observe a contact between δ and both τ and γ in native holoenzyme. We propose a subunit arrangement (Fig. 8) with τ in place of γ#2 and γ#3 in native holoenzyme. In this arrangement, δ could be cross-linked to the adjacent DnaX #3 subunit and δ' to the adjacent DnaX (γ#1) subunit.

B33 cross-links proteins by reaction with solvent-exposed lysine residues. Among residues in the DnaX complex structure, only Lys-161 of the adjacent DnaX#3 subunit was close enough to cross-link to any of the lysine residues in δ (δ Lys-313 is 15 Å from DnaX#3 Lys-161). The proximity to allow cross-linking was estimated liberally, permitting up to 5 Å of thermal motion from the positions established in the crystal structure. All other lysines in the DnaX complex are at least 24 Å away, too far for the 11.4 Å B33 reagent to cross-link. Examination of δ indicated that only Lys-85 and Lys-130 were close enough to two exposed DnaX#1 lysines to cross-link (Lys-161 and Lys-51 of DnaX#1 with distances of 14.6 and 12.7 Å, respectively).

Because native holoenzyme presumably has the two τ#7 adjacent, like in DNA polymerase III (6), permitting holoenzyme to dimerize, the DnaX#2 position is likely occupied by τ (Fig. 8). This leaves the problem of how δ could cross-link to a distal γ at 231 (Fig. 8) in addition to the adjacent γ#3 unresolved. The lysine-rich KAKK sequence near the C terminus of γ was removed to permit crystallization and thus was not part of the DnaX complex structure. Examining the structure of the DnaX complex from the top view (above the associated Domain III) indicates multiple lysines on the top surface of δ, unlike the top Domain III face of the other subunits. The C-terminal residue of γ that is part of the structure (Pro-368) is displaced over the adjacent subunit (γ#3 over γ#2, γ#2 over γ#1, and γ#1 over δ'). This creates an asymmetric structure where the C terminus of γ#1 is actually closest to δ lysines (Lys-334 and Lys-261) on the top surface of δ. Lys-334 and

---

3 M.-S. Song, unpublished observation.
likely second (δ-τ) cross-link observed for δ would be between δ Lys-261 or Lys-334 and a basic C-terminal residue of DnaX. The sequence between Pro-368 and KAKK (residues 426–429 of mature DnaX) is more than sufficient to bring these residues into the proximity of δ Domain III.

The arrangement proposed in Fig. 8 is speculative but is most consistent with the cross-linking and structural data. Further experimentation will be required to determine decisively the location of the two τ subunits within the three DnaX components of the DnaX complex. Another important issue, not resolved by this work, is whether Domain III of DnaX, δ, and δ’ just function to hold the complex together, with all of the interesting aspects of the reaction taking place by interaction of Domains I and II of the various subunits, or whether important communication occurs through the Domain III link. Domain III also provides a potential conduit for information flowing from Domain IV of τ and its link to the DnaB helicase and DNA primase, potentially permitting coordination of Okazaki fragment cycling with primer synthesis as proposed previously (36).

REFERENCES

1. Gueutlich, B., Onrust, R., Sali, A., O’Donnell, M., and Kuriyan, J. (1997) Cell 91, 335–345
2. Kim, S., Dallmann, H. G., McHenry, C. S., and Marians, K. J. (1996) Cell 84, 643–650
3. Kim, S., Dallmann, H. G., McHenry, C. S., and Marians, K. J. (1996) J. Biol. Chem. 271, 21406–21412
4. Gao, D., and McHenry, C. S. (2001) J. Biol. Chem. 276, 4441–4446
5. Gao, D., and McHenry, C. S. (2001) J. Biol. Chem. 276, 4435–4440
6. McHenry, C. S. (1982) J. Biol. Chem. 257, 2657–2663
7. Studwell-Vaughan, P. S., and O’Donnell, M. (1991) J. Biol. Chem. 266, 19833–19841
8. Song, M. S., Pham, P. T., Olson, M., Carter, J. R., Franden, M. A., Schaap, R. M., and McHenry, C. S. (2001) J. Biol. Chem. 276, 35165–35175
9. McHenry, C. S., Oberfelder, R., Johanson, K., Tomasiewicz, H., and Franden, M. A. (1987) in DNA Replication and Recombination (Kelly, T. J., and McMacken, R., eds) pp. 47–62, Alan R Liss, Inc., New York
10. Maki, S., and Kornberg, A. (1988) J. Biol. Chem. 263, 6555–6560
11. Glover, B. P., and McHenry, C. S. (2000) J. Biol. Chem. 275, 3017–3020
12. Olsón, M. W., Dallmann, H. G., and McHenry, C. S. (1995) J. Biol. Chem. 270, 28570–28577
13. Kolman, Z., Yushakov, A., Andjelkovic, J., and O’Donnell, M. (1998) EMBO J. 17, 2436–2449
14. Glover, B. P., and McHenry, C. S. (1998) J. Biol. Chem. 273, 23476–23484
15. Pritchard, A. E., Dallmann, H. G., Glover, B. P., and McHenry, C. S. (2000) EMBO J. 19, 6536–6545
16. Glover, B. P., Pritchard, A. E., and McHenry, C. S. (2002) J. Bacteriol., in press
17. McHenry, C. S., and Johanson, K. O. (1984) in Proteins Involved in DNA Replication (Hubscher, U., and Spadari, S., eds) pp. 315–319, Plenum Press
18. Johanson, K. O., and McHenry, C. S. (1984) J. Biol. Chem. 259, 4589–4595
19. Glover, B. P., and McHenry, C. S. (2001) Cell 105, 925–934
20. Gao, D., and McHenry, C. S. (2001) J. Biol. Chem. 276, 4447–4453
21. Glover, B. P., Pritchard, A. E., and McHenry, C. S. (2001) J. Biol. Chem. 276, 35842–35846
22. Song, M. S., Dallmann, H. G., and McHenry, C. S. (2001) J. Biol. Chem. 276, 40668–40679
23. Song, M. S., and McHenry, C. S. (2001) J. Biol. Chem. 276, 48709–48715
24. Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. (1999) Genome Res. 9, 27–43
25. Jeruzalmi, D., Yurieva, O., Zhao, Y., Young, M., Stewart, J., Hingorani, M., O’Donnell, M., and Kuriyan, J. (2001) Cell 106, 417–428
26. Jeruzalmi, D., O’Donnell, M., and Kuriyan, J. (2001) Cell 106, 429–441
27. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
28. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
29. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
30. Walker, J. R., Hervas, C., Ross, J. D., Blinkova, A., Walbridge, M. J., Pumarega, E. J., Park, M. O., and Neely, H. R. (2000) J. Bacteriol. 182, 6106–6113
31. Dallmann, H. G., Thimmig, R. L., and McHenry, C. S. (1995) J. Biol. Chem. 270, 29555–29562
32. Cull, M. G., and McHenry, C. S. (1995) Methods Enzymol. 263, 22–35
33. Carpenter, J. R., Franden, M. A., Aebersold, R., and McHenry, C. S. (1992) J. Bacteriol. 174, 7013–7025
34. Kim, D. B., and McHenry, C. S. (1996) J. Biol. Chem. 271, 20690–20698
35. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) EMBO J. 1, 945–951
36. Li, X., and Marians, K. J. (2000) J. Biol. Chem. 275, 34757–34765
37. Kunau, W. H., Beyer, A., Franken, T., Gette, K., Marzinc, M., Sadowsky, J., Skalez-Rorowski, A., and Wiebel, F. F. (1993) Biochimie (Paris) 75, 209–224