Nucleic Acids Research, 1993, Vol. 21, No. 14 3281–3286

Isolation, sequencing and overexpression of the gene encoding the θ subunit of DNA polymerase III holoenzyme

Jeffrey R.Carter, Mary Ann Franden, Ruedi Aebersold[1], Doek Ryong Kim and Charles S.McHenry*

The University of Colorado Health Sciences Center, Department of Biochemistry, Biophysics and Genetics, 4200 E. Ninth Avenue, Denver, CO 80262, USA and [1]The Biomedical Research Centre and Department of Biochemistry, University of British Columbia, 2222 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3, Canada

Received February 22, 1993; Revised and Accepted May 18, 1993 GenBank accession no. L05381

ABSTRACT

The gene encoding the θ subunit of DNA polymerase III holoenzyme, designated *holE*, was isolated using a strategy in which peptide sequence was used to derive a DNA hybridization probe. Sequencing of the gene, which maps to 41.43 centisomes of the chromosome, revealed a 76-codon open reading frame predicted to produce a protein of 8,846 Da. When placed in a *tac* promoter expression vector, the open reading frame directed expression of a protein, that comigrated with authentic θ subunit from purified holoenzyme, to 6% of total soluble protein.

INTRODUCTION

DNA polymerase III holoenzyme (holoenzyme), the major replicative enzyme of *E.coli*, is a complex of at least ten different protein subunits. Holoenzyme is distinguished from other polymerases in *E.coli* by its ability to rapidly and processively synthesize products of 150 to 500 kb [1], to coordinate leading- and lagging-strand replication through an asymmetric, dimeric polymerase complex [2–5], and to interact with other proteins of the replisome [6–8]. The major processivity subunit is β, which forms a torus-shaped dimer on primed DNA [9] and clamps the polymerase to the template. The β subunit is loaded onto the primed template in an ATP-dependent reaction [10–12] catalyzed by the 5-subunit γ complex [10, 13–15] [γ, (*dnaX*; 16, 17), δ (*holA*; 18, Z.Dong, R.Onrust and M.O'Donnell, submitted), δ′ (*holB*; J.Carter, M.Franden, R.Aebersold and C.McHenry, accepted for publication; Z.Dong, R.Onrust and M.O'Donnell submitted), χ (*holC*; J.Carter, M.Franden, J.Lippincott and C.McHenry submitted; R.Crombie, H.Xiao, Z.Dong, R.Onrust, and M.O'Donnell, submitted) and ψ (*holD*; J.Carter, M.Franden, J.Lippincott and C.McHenry submitted, R.Crombie, H.Xiao, Z.Dong, R.Onrust, and M.O'Donnell, submitted)]. The catalytic polymerase subunit, α [19, 20] (*dnaE*; 21), is isolable as part of pol III core [22], which also contains ε (*dnaQ*; 23), the 3′→5′ proofreading exonuclease [24], and θ, which has no known role.

In this report, we focus on the θ subunit. This subunit copurifies with the other subunits of holoenzyme, of pol III* (holoenzyme minus the β subunit), of pol III′ (holoenzyme minus the β subunit and the γ complex) and of pol III (reviewed in references 25 and 26), and is therefore assigned as a subunit of holoenzyme. Although θ, as part of pol III, is bound to either α, ε, or both, θ is not required for α to associate with ε [17]. Reconstitution of a minimal catalytic enzyme complex capable of highly processive and rapid synthesis and of cycling to new preinitiation complexes requires only ε and α, not θ [27].

Previous efforts to elucidate the function of θ have been hampered by the lack of θ, in a form free of other holoenzyme subunits. We report here the isolation and sequencing of the structural gene for θ, and expression of the protein to about 6% of the soluble protein of the cell. This work is a critical step toward the goal of obtaining large quantities of the θ protein for biochemical and functional analysis.

MATERIALS AND METHODS

Chemicals

SDS, acrylamide, N,N′-methylene-bis-acrylamide and Coomassie brilliant blue R-250 were purchased from Bio-Rad. Tris–HCl, polyvinylpyrrolidone, dextran sulfate, bovine serum albumin, and Ficoll were purchased from Sigma. Urea was purchased from Fisher. [α-^{32}P]ATP (3000 Ci/mmol, 10 mCi/ml) was purchased from ICN. Low-molecular-weight protein standard was purchased from Pharmacia, bacteriophage λ DNA digested with *Hin*dIII was purchased from Promega. SeaKem LE agarose was purchased from FMC BioProducts. All other chemicals were reagent grade. Oligonucleotides used in this study are in Table 1.

Bacterial strains, phages and plasmids and media

XL1Blue (Stratagene) was used for routine transformation experiments. HB101 [*supE*44 *hsdS*20 (rB$^-$mB$^-$) *recA*13 *ara*-14

* To whom correspondence should be addressed

Table 1. Oligonucleotides used in this study

PCR[a]	
	CTCGAATTCGAYCARACNGARATGGAYAARGT
	CTGCATCTAGAARRTGYTCNGGYTG
Sequencing	
	CCGCCGCCAAATCGACATTC
	GGCTCACTCTTAGTAAC
	TTCGGTCAATCTGTCAC
	AGGGAGCGTGCAGTTGA
	GTAAACGTGACAGATTG
	TGATGAGCTACCTGTTC
Reconstruction of 5′ end of gene encoding θ.[b]	
	TTTTCTGCAGCCATGGTTATTTAAGTTTGGGCTCGTA
	CCCCTGCAGAATTCAGATCTAGGAGGTTTAAAATAATGCTGAA-GAACCTGGCTAAA

[a] Bases in boldface were chosen based on peptide sequences of θ. Bases in lightface include an *Eco*RI site (underlined) and an *Xba*I site (double underlined). Nucleotide abbreviations are: R = A or G; Y = T or C; N = A, T, C, or G.
[b] The consensus ribosome-binding site is underlined and the initiation codon of the gene is double underlined.

proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1; 28] was the host strain for the overexpression experiment. Chromosomal DNA was isolated from MAF102 [18], a *lexA3*, *uvrD* [29] derivative of MG1655 [30]. MGC100 [*Δlac(I-P-O-Z-Y-A-)U169 Δlon araD139 strA supF*; 18] was the source of holoenzyme. Bacteriophage 19H3 of the Kohara mini-set collection was the generous gift of D.Berg. pBlueScript II SK+ (Stratagene) was the standard cloning vector. pHN1 is a *tac* promoter-based *dnaQ* expression plasmid (D.R.Kim and C.McHenry, unpublished data) derived from the expression plasmid pBBMD11 [31, 32]. L broth and agar [33] were used for routine cell growth. Agar medium used for λ plate lysates was L broth supplemented with glucose and MgSO4 to final concentrations of 0.2% and 10 mM, respectively. F broth (1.4% yeast extract, 0.8% peptone, 1.2% potassium phosphate pH 7.5, and 1% glucose) was used in the overexpression experiment. Ampicillin and tetracycline were used at 50 and 10 μg/ml, respectively.

DNA purification

Large-scale plasmid purification was performed using the alkaline-SDS lysis method [34] followed by two CsCl-ethidium bromide equilibrium density gradient centrifugations as described [35]. High-molecular-weight genomic DNA was isolated by lysing cells with lysozyme and SDS, phenol/chloroform extracting the sample and purifying the DNA on a CsCl ethidium bromide equilibrium density gradient as described [18]. Bacteriophage λ particles and DNA were purified using standard plate lysis and λ DNA extraction procedures [35].

Enzymes

Restriction enzymes, T4 DNA ligase and T4 polynucleotide kinase were purchased from Promega and used according to manufacturer's instructions. Lysozyme was from Worthington. Holoenzyme was prepared as described [18] to a specific activity of $5.4-7.0\times10^5$ units/mg

Agarose gel electrophoresis

Agarose gels containing 0.7−1.0% agarose dissolved in Tris borate buffer (89 mM Tris base, 2.75 mM EDTA, 89 mM boric acid) were run as described [35] for routine plasmid or phage DNA analysis. Agarose gels to be used for Southern transfer were run at 4°C. When required for restriction fragment purification, gel slices were excised without UV irradiation as described [18]. DNA was eluted from the gel using an Elutrap apparatus (Schleicher and Schuell) and then ethanol-precipitated. Alternatively, restriction fragments separated on an agarose gel made with 40 mM Tris−HCl pH 8.0, 20 mM acetic acid and 10 mM EDTA and were purified using the GeneClean II DNA purification kit (Bio 101, Inc.).

Preparation of radiolabeled DNA

The 150-bp PCR fragment was purified by gel electrophoresis and radiolabeled using the Random Primed DNA labeling kit from Boehringer Mannheim Biochemicals, following the manufacturer's instructions. Labeled DNA was separated from unincorporated nucleotide by chromatography through a G-25 (Pharmacia) gel filtration spin column equilibrated and developed with 20 mM Tris−HCl pH 8.0, 1 mM EDTA and 50 mM NaCl. Labeled DNA was heat-denatured before use in hybridization experiments.

Southern hybridization

Transfer and crosslinking of chromosomal DNA restriction fragments to a GeneScreen nylon membrane (New England Nuclear) and hybridization of radiolabeled DNA to the transferred DNA were performed as described [18]. The blot was exposed on a Molecular Dynamics phosphorimager screen for 4−24 h, and the screen was scanned on a Molecular Dynamics phosphorimager.

A blot of the Kohara [36] mini-set of λ clones was obtained from Takara Shuzo, Inc. and was used according to the supplier's instructions.

Production of tryptic peptides of θ

Tryptic peptides were obtained from the θ subunit essentially as described [18, 37]. Briefly, two 500-μg samples of holoenzyme were concentrated 35-fold in a colloidion bag (Schleicher and Schuell), and resolved on an SDS 7.5−17% polyacrylamide gel. One sample was transferred to a ProBlott PVDF membrane. The θ subunit was excised from the membrane and subjected to amino-terminal sequencing. The other sample was transferred to a nitrocellulose membrane. The θ subunit was removed from the blot and digested on the membrane with trypsin. Tryptic fragments were separated by reversed-phase HPLC on a C-18 narrow-bore Brownlee Aquapore Bu-300 column, and were sequenced on an Applied Biosystems 477A Protein Sequencer [37, 38].

PCR reactions

PCR reactions were performed using the Perkin Elmer GeneAmp PCR reagent kit and were incubated in a Perkin Elmer Cetus Model 480 PCR machine. Each reaction contained 1 ng of *E.coli* chromosomal DNA and 2 oligonucleotide primers each at 1 mM. Tubes containing template DNA and primers were incubated at 94°C for 7 min to denature the template, and taken to 85°C to allow addition of polymerase and dNTP's. The reactions were cycled 35 times through a 1-min incubation at 94°C, a 5-min ramp from 50°C to 65°C and a rapid return to 94°C. Reaction products were precipitated with ethanol, dissolved in 10 mM Tris−HCl (pH 8.0) and 1 mM EDTA, digested with *Eco*RI and *Xba*I, and purified by agarose gel electrophoresis.

Table 2. Peptide sequences of θ

1.	MetLeuLysAsnLeuAlaLysLeuAspGlnThrGluMetAspLysValAsnVal-AsnValAspLeuAlaAlaAla[a]
2.	XaaGlnProGluHisLeuArg[b]
3.	(Leu/Ser)ProTyrGluPro[b]

[a] The amino-terminal sequence of θ.
[b] Tryptic fragments of θ. Xaa is an unidentified residue. Residues in parentheses indicate uncertainty as to the identity of the residue.

DNA sequencing

The DNA sequence was determined by Lark Sequencing Technologies, Inc. (Houston) using the dideoxy chain termination method of Sanger et al [39]. Templates were double-strand plasmids purified by CsCl-ethidium bromide equilibrium density gradient centrifugation as described above.

Sequence database searches

GenBank DNA sequences were translated in all 6 reading frames using the TFASTA program and compared to the predicted sequence of θ using the comparison method of Pearson and Lipman [40].

RESULTS

To clone the gene encoding the θ subunit of holoenzyme required an oligonucleotide probe specific to the gene. We therefore designed oligonucleotide primers, based on partial amino acid sequence, to amplify a portion of the gene using PCR.

Peptide sequences of θ

Ten μg of θ were resolved from the other subunits of purified holoenyzme by SDS-polyacrylamide gel electrophoresis and transferred to a nitrocellulose or PVDF membrane as described [18]. The portion containing θ was cut out of each membrane. The material on the PVDF membrane was subjected to amino-terminal sequencing to yield peptide 1 (Table 2). The material on the nitrocellulose membrane was digested directly on the filter with trypsin. Tryptic fragments were resolved using a reversed-phase HPLC C-18 column, and the amino termini of two well-resolved peptides were sequenced to yield peptides 2 and 3 (Table 2).

Obtaining a probe for the gene encoding θ

Comparison of each peptide sequence to protein sequences in the data base to determine if the gene encoding θ had already been isolated revealed no matches to previously reported sequences. Therefore, we designed two oligonucleotide primers, based on peptide sequences 1 and 2, to amplify a segment of the gene. PCR using these two primers resulted in products of 75, 150 and 200 bp (data not shown). Each product was digested with *Eco*RI and *Xba*I, cloned into pBlueScript II SK + and sequenced. The 150-bp fragment was found to represent the gene encoding θ. This conclusion was based on translation of the DNA sequence (Fig. 1). Further confirmation came from the presence of DNA immediately after primer 1 sequence that encoded the carboxy-terminal 7 amino acid residues of peptide 1, and from an arginine codon before primer 2, since trypsin cleaves to the carboxy side of the basic residues arginine and lysine. Since DNA encoding peptide 3 was not found in the 150-bp fragment, we predicted that this peptide was to the carboxy side of peptide 2.

Mapping the gene encoding θ

E.coli chromosomal DNA was digested with *Bam*HI, *Bgl*II, *Eco*RI, *Eco*RV, *Hin*dIII, and *Pst*I used singly and in all two-enzyme combinations. The DNA was transferred to a nylon membrane and hybridized with radiolabeled 150-bp fragment. One restriction fragment from each digestion hybridized to the 150-bp fragment, indicating a single copy of the gene (data not shown). The hybridization data were used to construct a restriction map (Fig. 2) of the region of the chromosome including the gene encoding θ. Comparison of this map to the restriction map of the entire *E.coli* chromosome [36] revealed a single region of correspondence at about 40.5 minutes of the reported map. The lone discrepancy between the two restriction maps was the location of one *Bam*HI site. Hybridization of the 150-bp fragment to the Kohara mini-set collection identified only one positive clone (19H3; data not shown), which maps to 40.5 minutes [36]. Using the most recent map of the *E.coli* chromosome which correlates the physical and genetic maps [41, 42], we mapped the putative gene encoding θ to 1936−37 kb, or 41.43 centisomes.

Cloning the gene encoding θ

The θ subunit is about 10 kDa, as determined by SDS-polyacrylamide gel electrophoresis. A protein of this size would require a gene of about 250 bp. The gene-specific 150-bp fragment hybridized to a 2.3-kb *Eco*RV fragment located within a 5-kb *Bam*HI fragment (Fig. 2). Both *Bam*HI sites were greater than 250 bp from the *Eco*RV fragment, suggesting that the entire gene encoding θ was within the *Bam*HI fragment. This *Bam*HI fragment was isolated from the λ clone 19H3 and cloned into pBlueScript II SK+ to create pMAF65.

DNA sequence analysis

The limited DNA sequence obtained during analysis of the PCR fragments was used to design two primers, oriented in opposite directions, to begin DNA sequencing. DNA sequence obtained using these primers was used to design additional sequencing primers, and this strategy was repeated to determine double-strand DNA sequence for 641 bp that included the putative gene encoding θ (Genbank accession number L05381) The initiation codon was identified by determining the DNA that encoded peptide 1, the 23-residue amino terminus of θ. Three observations confirmated this identification: 1) peptides 2 and 3 were encoded by DNA in the same reading frame as peptide 1; 2) the distance between the DNA encoding peptides 1 and 2 was consistent with the 150-bp fragment obtained in the PCR experiments; 3) as predicted from sequencing of the 150-bp fragment, peptide 3 was encoded by DNA downstream of DNA encoding peptide 2. The termination codon was identified as the first in-frame stop codon after the DNA encoding peptide 1. Translation of the putative gene predicted that θ contains 76 amino acid residues and is a protein of 8,846 Da, consistent with the estimate of 10,000 Da obtained by SDS-polyacrylamide gel electrophoresis.

The initiation codon of putative gene encoding θ begins 5 bp after a sequence (AGGAGAT) that differs from the consensus Shine−Dalgarno ribosome-binding site [43] by only 1 base. We identified sequences with similarities to consensus −35 (TTGCCA)and −10 (TAACTT) promoter sequences which, if active, would produce a transcript with only 6 residues before the ribosome-binding site. The existence of termination codons in all three reading frames within the 160 bp preceding the

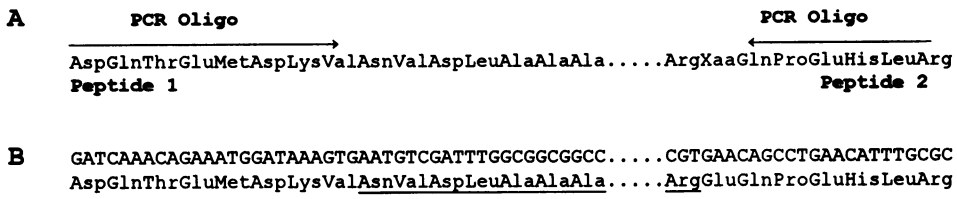

Figure 1. DNA sequence of a PCR product that is part of the gene encoding θ. **(A)** Oligonucleotides for PCR were designed based on regions of peptide sequences 1 and 2. Xaa indicates an unidentified residue in peptide sequence 2. **(B)** Three PCR products were cloned and sequenced, and the sequence of each product was translated. A partial DNA sequence and translation of the 150-bp product is shown. Amino acid residues expected in a fragment of the gene encoding θ are underlined.

initiation codon indicated that translation of θ probably is not coupled with that of an upstream gene. The amino terminus of a potential open reading frame of at least 101 bp begins 100 bp downstream of the putative gene encoding θ.

Search of the GenBank version 67 to identify regions of similarity between the protein sequence of the putative gene encoding θ and that of other proteins indicated that residues 27 to 55 of θ are 34% identical and 48% similar to a region of the primosomal replication factor Y (also called n'; 44, 45) (Fig. 3). Replication factor Y is a subunit of the primosome that binds specific sites on ssDNA templates [46, 47] to initiate primosome assembly [46], and is a DNA-dependent ATPase and helicase [48, 49].

Codon usage

If θ is stoichiometric with other holoenzyme subunits, it would be expected to exist at about 10−20 copies per cell. A proposed mechanism for maintaining low cellular protein levels is the frequent use of 8 codons (AUA, UCG, CCU, CCC, ACG, CAA, AAU, AGG) whose cognate tRNA's are poorly represented in the cell [50]. To address whether the use of rare codons might modulate expression of θ, we compared the codon usage of the previously sequenced gene to that of other holoenzyme genes and of 25 highly expressed *E.coli* genes, 10 of which (ribosomal protein genes) are very highly expressed. The putative gene encoding θ has 10.5% rare codons, i.e., 2.5- to 6-fold higher than the percentage in highly expressed genes, and higher than the percentage in several other holoenzyme-subunit genes (Table 3).

Overexpression of the gene encoding θ

To test the hypothesis that we had isolated the structural gene for θ, we cloned the gene downstream of the strong *tac* promoter and analyzed whether the gene expresses a protein and if the protein comigrates with the θ subunit from purified holoenzyme. The gene was amplified by PCR from pMAF65 using two primers. One primer included recognition sites for *Pst*I, *Eco*RI and *Bgl*II, a consensus ribosome-binding site, the first 7 codons of the gene (including a single-base deviation from the wild-type gene sequence that converted the rare asparagine codon AAT to the more common codon AAC). The second primer was complementary to the last 7 codons of the gene, including the putative termination codon, and contained an *Nco*I site. The PCR product obtained using these primers was cloned into the *Sma*I site of pBlueScript II SK+ to create pHN100 and sequenced to ensure that no mistakes were introduced during amplification of the gene. The PCR product was removed from pHN100 with *Eco*RI and *Pst*I and inserted into pHN1, a *dnaQ*-overexpressing plasmid, in place of *dnaQ* to create pNH101, the *tac* promoter-based θ-expression plasmid (Fig. 4). pHN101 was introduced into HB101 to create HN101, which was tested for the ability to produce θ upon IPTG induction (Fig. 5). The negative control strain HB101 and the positive control strain HN1 (HB101 containing pHN1) were included in this experiment. A protein that comigrates with the θ subunit of purified holoenzyme was produced only by HN101 and only when the strain was induced with IPTG (Fig. 5, compare lanes 5 and 6). HB101 treated with IPTG failed to produce this protein (Fig. 5, lane 2), and treatment of HN1 with IPTG induces production of only the ε subunit (Fig. 5, lane 4). These results strengthen the conclusion that we have isolated the gene encoding θ. Densitometry of lanes 5 and 6 of Fig. 5 indicated that this protein was about 6% of the soluble protein.

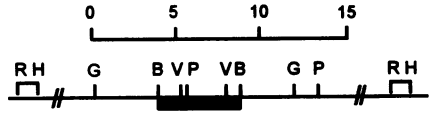

Figure 2. Restriction map of the region of the chromosome including the gene encoding θ. MAF102 chromosomal DNA was transferred onto a GeneScreen nylon membrane and hybridized with the radiolabeled 150-bp PCR fragment. A restriction map was constructed based on the sizes of chromosomal restriction fragments that hybridized to the probe. Kilobases are indicated above the restriction map. Paired diagonal lines indicate discontinuities. Restriction site abbreviations are: B, *Bam*HI; G, *Bgl*II; R, *Eco*RI; V, *Eco*RV; H, *Hin*dIII; P, *Pst*I. The 5-kb *Bam*HI fragment (indicated by the filled rectangle) was isolated from λ clone 19H3 and cloned into pBlueScript II SK+ to create pMAF65. The *Eco*RI and *Hin*dIII restriction fragments were too large to precisely localize these restriction sites.

```
27  PheLysGluArgTyrAsnMetProValIle
    • •  • • • •   • •
265 PheArgGluArgPheAsnAlaProValGlu

37  AlaGluAlaValGluArgGluGlnProGlu
                                •
275 ValLeuHisSerGlyLeuAsnAspSerGlu

47  HisLeuArgSer...TrpPheArgGluArg
      •   •   •   •   •
285 ArgLeu...SerAlaTrpLeuLysAlaLys
```

Figure 3. Comparison of regions of the θ subunit and factor Y of the primosome. Residues of θ are in boldface and residues of factor Y are in lightface. Residues are numbered relative to the first codon of the respective proteins. Identities and similarities between residues of the two proteins are indicated by dots between residues. Periods indicate breaks in amino acid sequences required to align the sequences.

Table 3. Percentage of rare codons[a] in the three possible reading frames of *E.coli* genes

Genes[b]	Percent Rare Codons
gene	
encoding θ	10.5
dnaN	4.1
dnaE	4.9
holB	6.3
dnaX	6.8
holC	6.8
dnaQ	7.0
holA	8.7
dnaG	11.3
25 genes[d]	4.2
10 genes[d]	1.7

[a] Rare codons include AUA (Ile), UCG (Ser), CCU and CCC (Pro), ACG (Thr), CAA (Gln), AAU (Asn), and AGG (Arg) [50]
[b] Sequence data were derived from the following sources: *dnaN* [14], *dnaE* [55], *holB* (Carter et al., accepted), *dnaX* [16], *holC* (Carter et al., submitted), *dnaQ* [56], *holA* [18], *dnaG* [18].
[c] coding frame.
[d] 25 nonregulatory genes and 10 ribosomal protein genes [50].

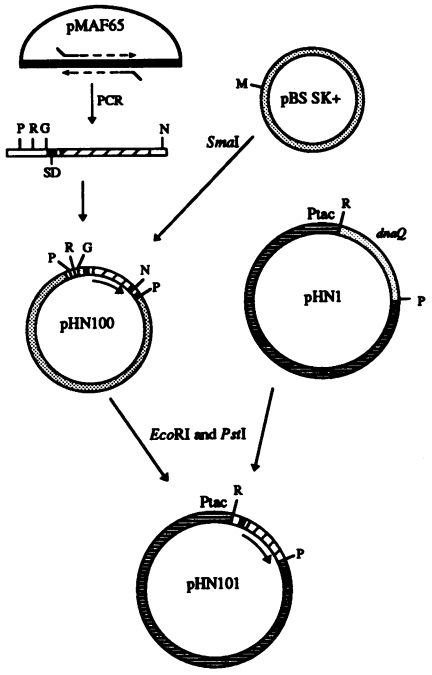

Figure 4. Construction of a plasmid to overexpress the θ subunit. pMAF65 is the original clone containing the θ subunit. The gene encoding θ was amplified from pMAF65 by PCR using primers complementary to the 5′ and 3′ ends of the gene. The PCR product was inserted into the *SmaI* site of pBlueScript II SK+ to create pHN100. The gene was removed from pHN100 and inserted into the *tac* expression plasmid pHN1 in place of *dnaQ* to create pHN101, the θ overexpression plasmid. Restriction enzyme recognition site abbreviations are: G, *BglII*; M, *SmaI*; N. *NcoI*; P, *PstI*; R, *EcoRI*. 'SD' indicates the Shine–Dalgarno consensus site. The curved arrow indicates the gene encoding θ. Dashed arrows indicate DNA that was amplified by PCR. The plasmids are not drawn to scale.

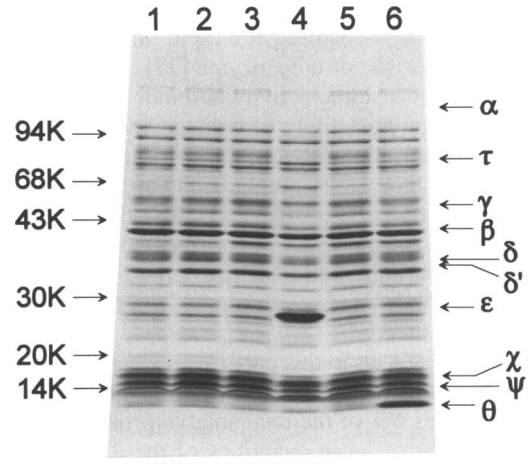

Figure 5. Overexpression of the θ subunit. Cells were grown in L broth, induced with IPTG and lysed as described in 'Materials and Methods'. Protein from equal cell masses were separated by SDS-polyacrylamide gel electrophoresis and stained with Coomassie Brilliant Blue. Lanes contain: 1, HB101 non-induced; 2, HB101 induced; 3, NH1 (HB101 containing pHN1), non-induced; 4, HN1 (HB101 containing pHN1), induced; 5, HN101 non-induced; 6, HN101 induced. Molecular mass markers (kDa) and holoenyzme subunits are indicated to the left and the right, respectively.

DISCUSSION

Using a reverse genetic approach, we have isolated the gene encoding the θ subunit of holoenzyme. This conclusion is based on three criteria. (i) A 150-bp DNA fragment was amplified by PCR using two primers designed from peptide sequences 1 and 2 of θ. This PCR product encoded 7 consecutive residues from peptide 1 that were not used in design of the primers, and a basic residue expected to be to the carboxy side of peptide 2. (ii) The PCR fragment, used as a gene-specific probe, allowed isolation of a 76-codon open reading frame that encoded peptides 1, 2, and all of peptide 3, which was encoded by DNA outside of the 150-bp probe. (iii) The 76-codon open reading frame directed expression of a protein with a predicted molecular mass of 8,846 Da that comigrates with θ found in purified holoenzyme. This open reading frame has also been isolated by the O'Donnell laboratory (P.S.Studwell-Vaughan and M.O'Donnell, submitted). Following the suggestion of Ken Marians (Sloan Kettering), we and the O'Donnell laboratory call the gene encoding θ *holE*. In more recent studies (D.R.Kim and C.McHenry, preliminary results) *holE* has been co-expressed with *dnaE* (α subunit) and *dnaQ* (ϵ subunit) from a single plasmid. The *holE* product (θ) complexes with α and ϵ to reconstitute pol III core ($\alpha\epsilon\theta$), and complexes with ϵ to form an $\epsilon\theta$ complex which, unlike ϵ alone, is soluble.

Analysis of *holE* DNA sequence revealed a level of poorly used codons higher than that seen in any other holoenzyme-subunit gene. It has been suggested that a high level of poorly used codons contributes to a low level of gene expression (reviewed in 51). Since the percent of rare codons in *holE* is very high, we suspected that the presumed naturally low level of chromosomal *holE* expression could be due in part to the relatively high number of poorly used codons. Although the *holE* overexpression plasmid contained a single base change to eliminate one rare codon, it is not likely that elimination of only one rare codon allowed expression of θ to 6% of total cell protein. Indeed, even though the gene was expressed by the strong *tac* promoter on a high copy number plasmid, expression of θ to this level is more consistent with the hypothesis that codon usage bias does not influence the level of gene expression [52].

The role of θ in DNA replication is not understood. There is no evidence that θ is required for rapid and processive DNA synthesis characteristic of holoenzyme [27], and the stimulation of the $3' \rightarrow 5'$ proofreading activity is minor [53]. Instead of an exclusive effect on an activity of holoenzyme, θ might provide allosteric communication between holoenzyme and an additional component of the DNA replication apparatus, e.g, the primosome, to ensure the coordinated functioning of components of the replication apparatus. Use of purified holoenzyme subunits in the coupled leading- and lagging-strand synthesis system developed by Marians and coworkers [1, 54] might serve to test this hypothesis directly.

The results presented in this paper provide a foundation upon which to investigate the activity of θ. Availability of a strain that overproduces θ as 6% of the total soluble protein of *E.coli* will allow purification of large quantities of the θ subunit, and, in turn, *in vitro* biochemical experiments aimed at elucidating the role of θ.

ACKNOWLEDGEMENTS

We thank Matthew Olson for preparing the holoenzyme used in Fig. 5, Janine Mills for synthesis of oligonucleotides, Julie Lippincott and Edward Bures for protein sequencing, and Michael O'Donnell and colleagues for exchanging DNA sequence of *holE* with us prior to publication. This work was supported by a grant from the Medical Research Council of Canada (R.A.), and by National Institutes of Health grant GM36255 and American Cancer Society grant NP-819 to C.S.M. R.A. is a scholar of the Medical Research Council of Canada.

REFERENCES

1. Wu,C.A., Zechner,E.L. and Marians,K.J. (1992) *J. Biol. Chem.*, **267**, 4030–4044.
2. Hawker,J.R.,Jr. and McHenry,C.S. (1987) *J. Biol. Chem.*, **262**, 12722–12727.
3. Johanson,K.O. and McHenry,C.S. (1984) *J. Biol. Chem.*, **259**, 4589–4595.
4. McHenry,C.S. and Johanson,K.O. (1984) In Hubscher,U. and Spadari,S. (eds), Proteins Involved in DNA Replication Plenum Publishing Corp., New York, N.Y., pp. 315–319.
5. Wu,C.A., Zechner,E.L., Hughes,A.J.J., Franden,M.A., McHenry,C.S. and Marians,K.J. (1992) *J. Biol. Chem.*, **267**, 4064–4073.
6. Zechner,E.L., Wu,C.A. and Marians,K.J. (1992) *J. Biol. Chem.*, **267**, 4054–4063.
7. Wu,C.A., Zechner,E.L., Reems,J.A., McHenry,C.S. and Marians,K.J. (1992) *J. Biol. Chem.*, **267**, 4074–4083.
8. Kornberg,A. and Baker,T.A. (1992) DNA Replication, 2nd edn, W. H. Freeman and Co., New York, NY.
9. Kong,X.P., Onrust,R., O'Donnell,M. and Kuriyan,J. (1992) *Cell*, **69**, 425–437.
10. O'Donnell,M.E. (1987) *J. Biol. Chem.*, **262**, 16558–16565.
11. Wickner,W. and Kornberg,A. (1973) *Proc. Natl. Acad. Sci. USA*, **70**, 3679–3683.
12. Wickner,S. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 3511–3515.
13. Maki,S. and Kornberg,A. (1988) *J. Biol. Chem.*, **263**, 6555–6560.
14. McHenry,C. and Kornberg,A. (1977) *J. Biol. Chem.*, **252**, 6478–6484.
15. McHenry,C.S., Oberfelder,R., Johanson,K., Tomasiewicz,H. and Franden,M.A. (1986) In Kelly,T. and McMacken,R. (eds), Mechanisms of DNA Replication and Recombination Alan R. Liss, Inc., New York, N.Y., pp. 47–61.
16. Flower,A.M. and McHenry,C.S. (1986) *Nucleic Acids Res.*, **14**, 8091–8101.
17. Stukenberg,P.T., Studwell-Vaughan,P.S. and O'Donnell,M. (1991) *J. Biol. Chem.*, **266**, 11328–11334.
18. Carter,J.R., Franden,M.A., Aebersold,R. and McHenry,C.S. (1992) *J. Bacteriol.*, **174**, 7013–7025.
19. Maki,H., Horiuchi,T. and Kornberg,A. (1985) *J. Biol. Chem.*, **260**, 12982–12986.
20. Maki,H. and Kornberg,A. (1985) *J. Biol. Chem.*, **260**, 12987–12992.
21. Welch,M.M. and McHenry,C.S. (1982) *J. Bacteriol.*, **152**, 351–356.
22. McHenry,C.S. and Crow,W. (1979) *J. Biol. Chem.*, **254**, 1748–1753.
23. Scheuermann,R., Tam,S., Burgers,P.M.J., Lu,C. and Echols,H. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7085–7089.
24. DiFrancesco,R., Bhatnagar,S.K., Brown,A. and Bessman,M.J. (1984) *J. Biol. Chem.*, **259**, 5567–5573.
25. McHenry,C.S. (1988).*Ann. Rev. Biochem.*, **57**, 519–550.
26. McHenry,C.S. (1991) *J. Biol. Chem.*, **266**, 19127–19130.
27. Studwell,P.S. and O'Donnell,M. (1990) *J. Biol. Chem.*, **265**, 1171–1178.
28. Boyer,H.W. and Roulland-Dussoix,D. (1969) *J. Mol. Biol.*, **41**, 459–472.
29. Washburn,B.K. and Kushner,S. (1991) *J. Bacteriol.*, **178**, 2569–2575.
30. Guyer,M.G., Reed,R.R., Steitz,J.A. and Low,K.B. (1980) Cold Spring Harbor Symp. Quant. Biol., **45**, 135–140.
31. McHenry,C.S., Griep,M.A., Tomasiewicz,H. and Bradley,M. (1990) In Molecular Mechanisms in DNA Replication and Recombination Alan R. Liss, Inc., New York, N.Y., pp. 115–126.
32. Fürste,J.P., Pansegrau,W., Blöcker,H., Scholz,P., Bagdasarian,M. and Lanka,E. (1986) *Gene*, **48**, 119–131.
33. Miller,J.H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513–1523.
35. Sambrook,J., Frisch,E.F. and Maniatis,T. (1989) Molecular cloning: a laboratory manual, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
36. Kohara,Y., Akiyama,K. and Isono,K. (1987) *Cell*, **50**, 495–508.
37. Aebersold,R.H., Leavitt,J., Saavedra,R.A., Hood,L.E. and Kent,S.B.H. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6970–6974.
38. Hewick,R.M., Hunkapiller,M.W., Hood,L.E. and Dreyer,W.J. (1981) *J. Biol. Chem.*, **256**, 7990–7997.
39. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
40. Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2444–2448.
41. Rudd,K.E. (1992) In Miller,J.H. (ed), A Short Course in Bacterial Genetics Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 2.3–2.43.
42. Rudd,K.E., Miller,W., Ostell,J. and Benson,D.A. (1990) *Nucleic Acids Res.*, **18**, 313–321.
43. Shine,J. and Dalgarno,L. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 1342–1346.
44. Lee,E.H., Masai,H., Allen,J.,G.C. and Kornberg,A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4620–4624.
45. Nurse,P., DiGate,R.J., Zavitz,K.H. and Marians,K.J. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4615–4619.
46. Shlomai,J. and Kornberg,A. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 799–803.
47. Zipursky,S.L. and Marians,K.J. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 6521–6525.
48. Lasken,R.S. and Kornberg,A. (1988) *J. Biol. Chem.*, **263**, 5512–5518.
49. Lee,M.S. and Marians,K.J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8345–8349.
50. Konigsberg,W. and Godson,G.N. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 687–691.
51. de Boer,H.A. and Kastelein,R.A. (1986) In Reznikoff,W. and Gold,L. (eds), Maximizing Gene Expression Butterworths, Stoneham, MA, pp. 225–285.
52. Andersson,S.G.E. and Kurland,C.G. (1990) *Microbiol. Rev.*, **54**, 198–210.
53. Maki,H. and Kornberg,A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4389–4392.
54. Mok,M. and Marians,K.J. (1987) *J. Biol. Chem.*, **262**, 16644–16654.
55. Tomasiewicz,H.G. and McHenry,C.S. (1987) *J. Bacteriol.*, **170**, 5735–5744.
56. Maki,H., Horiuchi,T. and Sekiguchi,M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7137–7141.