Identification of Two *Escherichia coli* Pseudouridine Synthases That Show Multisite Specificity for 23S RNA†

Lixuan Huang,‡,§ Jung Ku,§ Manee Pookanjanatavip,|| Xiangrong Gu, Dorothy Wang, Patricia J. Greene, and Daniel V. Santi*

Departments of Biochemistry and Biophysics and of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0448

Received May 1, 1998; Revised Manuscript Received September 3, 1998

ABSTRACT: Several putative *Escherichia coli* pseudouridine (Ψ) synthases have been identified by iterative searching of genomic databases for ORFs homologous to known Ψ synthases [Gustafsson et al. (1996) *Nucleic Acids Res. 24*, 3756−3762]. Of these, *yceC* and *yfiI* were proposed to encode Ψ synthases which modify 23S rRNA. In the present work, *yceC* and *yfiI* were cloned and overexpressed in *E. coli*, and the encoded enzymes, YceC and YfiI, were purified to homogeneity. Both proteins converted Urd residues of rRNA to Ψ, thus confirming their identities as Ψ synthases. However, in in vitro experiments both enzymes extensively modified Urd residues of both 23S rRNA and 16S rRNA. Gene-disruption of *yceC* resulted in the absence of Ψ modification at positions U955, 2504, and 2580 of 23S RNA, thus identifying these sites as in vivo targets for YceC. Likewise, *yfiI* disruption resulted in the absence of Ψ modification at positions U1911, 1917, and possibly 1915 of 23S RNA. Disruption of *yceC* did not affect the growth under the conditions tested, whereas *yfiI*-disrupted cells showed a dramatic decrease in growth rate. Since YceC and YfiI hypermodify RNA in vitro, factors in addition to ribonucleotide sequence must contribute to the in vivo specificity of these enzymes.

Pseudouridine (Ψ),[1] the carbon 1′−carbon 5 glycoside isomer of uridine, is the most abundant modified nucleoside in RNA *(1−3)*. A total of 17 Ψs have been identified in *Escherichia coli* tRNA and rRNA, with 7 Ψs in tRNAs, 1 in 16S rRNA, and 9 in 23S rRNA.

Ψ residues are formed by posttranscriptional modification of specific uridine residues in RNA, and several *E. coli* Ψ synthases have been identified. TruA and TruB modify tRNAs, RsuA forms Ψ516 of 16S rRNA, and RluA forms Ψ746 in 23S rRNA and Ψ32 in tRNA *(4−7)*. The enzymes that catalyze formation of the remaining eight Ψs of 23S rRNA (at positions 955, 1911, 1915, 1917, 2457, 2504, 2580, and 2605) have not yet been identified.[2] Moreover, as with most other RNA modifications, little is known about the biological role of Ψ residues in RNA.

We have undertaken a program directed at identifying the enzymes responsible for the modification of nucleotides in *E. coli* RNAs *(8)*. Our approach is to (i) utilize sequences of known modification enzymes to identify homologous sequences in genomic databases, (ii) clone and express the corresponding gene products, (iii) determine whether the expressed enzymes possess the predicted activities, and (iv) if so, determine the site(s) of RNA modification. This approach should enable systematic studies of the functions of specific modified residues of RNA.

In a previous report, sequences of the four known Ψ synthases from *E. coli* were used as probes for iterative searching of the genomic database for homologous ORFs *(8)*. Using the sequence of RluA, the 23S rRNA Ψ746 synthase, as a probe, two unidentified *E. coli* ORFs, *yceC* and *yfiI*, had the most conserved homologues in *H. influenzae* and *M. genitalium*. They were thus postulated to catalyze the formation of the most conserved Ψs in 23S rRNA which occur in domains IV and V.

In this report, we describe the cloning and expression of the ORFs *yceC* and *yfiI* from *E. coli* and the purification and characterization of the encoded proteins. We show that YceC and YfiI are Ψ synthases that hypermodify both 16S and 23S rRNA in vitro, with about 1 in every 10−20 Urd residues converted to Ψ. Gene disruption experiments permitted us to identify the in vivo target sites for YceC and YfiI, and studies of gene-disrupted cells revealed a dramatic decrease in growth rate of the *yfiI*-disrupted cells.

† This work was supported by U.S. Public Health Service Grant GM 51232 from the National Institutes of Health.

* Correspondence should be addressed to this author.

‡ Present address: DuPont Company CR&D, Experimental Station, P.O. Box 80402, Wilmington, DE 19880.

§ These two authors contributed equally to this study.

|| Present address: Department of Chemistry, Chiang Mai University, Chiang Mai, Thailand 50200.

[1] Abbreviations: *yceC*, open reading frame of *E. coli* genomic DNA (Genbank accession number P23851); YceC, protein encoded by *yceC*; *yfiI*, open reading frame of *E. coli* genomic DNA (Genbank accession number P33643); YfiI, protein encoded by *yfiI*; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; DTT, dithiothreitol; Ψ, pseudouridine; 2d-TLC, two-dimensional thin-layer chromatography; Amp, ampicillin; Kan, kanamycin; Cm, chloramphenicol; superscript r or s, resistant or sensitive; CMC, 1- cyclohexyl-3-(2-morpholinoethyl)-carbodiimide metho-*p*-toluenesulfonate; TS, temperature sensitive.

[2] After this work was submitted, Conrad et al. (*23*) reported that YceC (renamed RluC) catalyzes the in vivo conversion of U955, 2504, and 2580 of 23S RNA to Ψ.

10.1021/bi981002n CCC: $15.00
Published on Web 10/22/1998

MATERIALS AND METHODS

Plasmids pCW1 and pWK1 used for in vitro transcription of *E. coli* 23S rRNA and *E. coli* 16S rRNA, respectively, were gifts from J. Ofengand *(9)*. p67YF0, used for preparation of yeast $tRNA^{Phe}$, was a gift from O. C. Uhlenbeck (Department of Chemistry and Biochemistry, University of Colorado, Boulder) *(10)*. pCG50, which encodes stem 90 of 23S rRNA (nt 2495−2582), was a gift from H. Noller (UC Santa Cruz). pUC4K containing the Kan^r gene was purchased from Pharmacia. pMAK705 *(11)* containing a TS origin of replication was provided by S. R. Kushner (University of Georgia, Athens, GA). [5-^{3}H]UTP (20.1 Ci/mmol), was purchased from Moravek Biochemicals. [α-^{35}S]dATP (1000 Ci/mmol), [α-^{32}P]UTP (3000 Ci/mmol) and [γ-^{32}P]ATP (3000 Ci/mmol) were from Amersham, and CMC was from Aldrich. Oligonucleotide synthesis and DNA sequencing were performed by the UCSF Biomolecular Resource Center.

Cloning yceC from E. coli Genomic DNA. A PCR reaction was performed using *E. coli* MC1061 genomic DNA (20 ng) as the template and two primers (0.3 μM each): 5′-GGGAATTCCATATGAAAACAGAGACTCCATCC-3′, which introduced a *Nde*I site (underlined) and hybridized to nt 1−21 of *yceC*; and 5′-CGGGATCCTTAGCGCGCGTTACGCAGC-3′, which introduced a *Bam*HI site (underlined) and hybridized to nt 942−960 including the stop codon. The 1 kb PCR product was digested with *Nde*I and *Bam*HI, ligated with *Nde*I/*Bam*HI-digested vector pET-15b to form the N-terminal $(His)_6$ fusion (pLH-YceC), and transformed into *E. coli* DH5α. DNA from Amp^r transformants was analyzed by restriction digestion and DNA sequencing.

Expression and Purification of YceC. pLH-YceC was transformed into *E. coli* BL21(DE3)(pLysS) cells. An overnight culture was used to inoculate 250 mL of LB (50 mg/L Amp). The culture was grown until the cell density reached A_{600} ∼0.8 when IPTG was added to 1 mM final concentration. The induced culture was grown for 3 h at 37 °C, and cells were harvested by centrifugation at 4000*g* for 15 min at 4 °C. The cell pellet was resuspended in 10 mL of buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris·Cl, pH 7.9), lysed by passing through a French pressure cell at 16 000 psi, and centrifuged at 36000*g* for 25 min. YceC-$(His)_6$ was purified using His-Bind column chromatography as described in the pET system manual (fifth edition, Novagen, Inc.). Cleavage of YceC-$(His)_6$ with thrombin to remove the His tag was performed as described in the Novagen manual. Unless otherwise specified, YceC-$(His)_6$ was used for the studies described here.

Cloning yfiI from E. coli Genomic DNA. PCR was performed using *E. coli* MC1061 genomic DNA (40 ng) as the template and two primers (1 μM each): 5′-CATGCCATGGCACAACGAGTACAGCT-3′, which introduced a *Nco*I site (underlined) and hybridized to nt 1−21 of *yfiI* gene; and 5′-CAACCGCTCGAGTCATAACCAGTCCACTTCATCCTT-3′, which introduced a *Xho*I site (underlined) and hybridized to nt 958−984 including the stop codon. The 1 kb PCR product was digested with *Nco*I and *Xho*I, ligated with *Nco*I/*Xho*I-digested pET-15b, and transformed into *E. coli* DH5α to give pMP-YfiI. DNA from Amp^r transformants was analyzed by restriction digestion and DNA sequencing.

Expression and Purification of YfiI. pMP-YfiI was transformed into *E. coli* BL21(DE3) cells. A 3 mL overnight culture of pMP-YfiI/BL21(DE3) was used to inoculate 1 L of LB (100 mg/L Amp). The culture was grown until the cell density reached A_{600} ∼0.8 when IPTG was added to 1 mM final concentration. The induced culture was grown for 3 h at 37 °C; cells were harvested by centrifugation at 4000*g* for 15 min at 4 °C and stored at −80 °C.

Frozen cell pellets from 500 mL of culture were thawed and suspended in 15 mL of buffer A (10 mM Tris-HCl, pH 8.0, 0.5 mM EDTA, 10% glycerol, 0.5 mM DTT) containing 100 mg/mL PMSF and 2 mg/mL leupeptin. Cells were lysed by two passages through a French pressure cell. The supernatant was loaded on a Q-Sepharose column (10 mL bed volume) equilibrated with buffer A, and protein was eluted with a 50 mL gradient of 0−800 mM KCl in buffer A. Fractions containing YfiI were identified using SDS−PAGE (Phast gel-Pharmacia) where YfiI appears as a prominent protein band with a molecular mass of 37 kDa. Fractions were pooled, concentrated, and desalted using a Centriprep concentrator from Amicon. The concentrated protein was loaded on a hydroxyapatite column (10 mL bed volume) equilibrated with buffer B (10 mM KP_i, pH 6.8, 0.5 mM EDTA, 10% glycerol, and 0.5 mM DTT), and eluted with a 50 mL gradient of 10 mM to 1 M KP_i, pH 6.8, 0.5 mM EDTA, 10% glycerol, and 0.5 mM DTT. Fractions with high concentrations of YfiI as determined by SDS−PAGE were pooled, and loaded onto a phosphocellulose column (5 mL bed volume) equilibrated with buffer B. Proteins were eluted with a 25 mL gradient of 0−800 mM KCl in buffer B, and fractions containing homogeneous YfiI were pooled, concentrated, and desalted using a Centriprep concentrator.

Construction of Stem 90h of 23S RNA. Two complementary oligonucleotides (5′-GAAGATCTTTGGAGTACTCTACGAAAGTAGAGCTGATGAGGCCG-3′ and 5′-CGGGATCCAGCTTGGAGTTTCGGCCTTTCGGCCTCATCAGCTCTAC-3′, with *Bgl*II and *Bam*HI sites underlined, respectively) were synthesized, each containing part of a hammerhead ribozyme sequence *(12)*. The two oligonucleotides (450 pmol each) were annealed in 60 μL of T4 DNA polymerase buffer (50 mM NaCl, 10 mM Tris·Cl, pH 7.9, 10 mM $MgCl_2$, 1 mM DTT) at 95 °C for 3 min and cooled slowly to room temperature. Polymerization was performed in 200 μL of T4 DNA polymerase buffer containing 33 μM of each dNTP and 6 units of T4 DNA polymerase. The reaction was incubated at 25 °C for 15 min, and quenched by adding EDTA (pH 8.0) to 10 mM and heating at 75 °C for 10 min. The polymerization product was purified by acrylamide gel electrophoresis. The gel purified duplex DNA was digested by *Bgl*II and *Bam*HI, and ligated into *Bam*HI-linearized pCG50 to give pCG50H. The resulting plasmid was analyzed by restriction analysis and DNA sequencing.

Synthesis of RNA Substrates. Full-length *E. coli* 23S rRNA, 16S rRNA, and yeast $tRNA^{Phe}$ were prepared by runoff in vitro transcription from restricted pCW1, pWK1, and p67YF0 DNAs, respectively. Stem 90h RNA was prepared by runoff in vitro transcription of *Bam*HI-linearized pCG50H.

The in vitro transcription reactions contained the appropriate DNA template, 4 mM each NTP, 20 mM $MgCl_2$, and 0.1 mg/mL T7 RNA polymerase *(13)*. [5′-^{32}P]UMP-RNA

and [5-^{3}H]Ura-RNA were prepared in similar reactions containing 0.1 mM [5′-α-^{32}P]UTP (1−4.0 Ci/mmol) or [5-^{3}H]UTP (1.0 Ci/mmol) instead of UTP, respectively. The synthesized RNAs were purified using Qiagen columns according to manufacturer's instructions (Qiagen Inc.).

Tritium Release Assays. Tritium release assays were carried out as described *(14, 15)*.

Nucleotide Analysis. Reactions (20 μL) containing 0.25 μM [5′-^{32}P]UMP-tRNA (8×10^4 cpm), 0.25 μM [5′-^{32}P]-UMP-16S rRNA (1.3×10^6 cpm), or 0.25 μM [5′-^{32}P]UMP-23S rRNA (2.7×10^6 cpm) and 1 or 2 μM ΨS in assay buffer (20 mM Tris-HCl, pH 8.0, 100 mM NH_4Cl, and 2 mM EDTA) were incubated at 37 °C for 3 or 5 h. The reaction mixtures were diluted with 80 μL of water, extracted with an equal volume of phenol, and precipitated with 3 volumes of ethanol. After the precipitates were washed with 80% ethanol and dried, the pellets were dissolved in 10 μL of 30 mM NaOAc, pH 5.3, and digested to completion with 1 μL of nuclease P1 (1 μg/μL) at 37 °C for 2 h. The [5′-^{32}P]NMPs produced were subjected to 2d-TLC on cellulose plates, using the following mobile phases: (1) isobutyric acid−concentrated NH_4OH-H_2O (66/1/33, v/v/v); (2) 2-propanol−concentrated $HCl-H_2O$ (70/15/15, v/v/v) *(16)*. The spots corresponding to pU and pΨ were detected by autoradiography, excised and extracted with H_2O overnight, and counted in 6 mL of Bio-Safe II (Research Products International Corp.).

Ψ *Modification in Stem 90h of 23S RNA.* Sequence analysis of RNA was performed as described *(17)*. Stem 90h RNA (13 μg), prepared from pCG50H and previously treated with YceC, was dissolved in 1 μL of deionized H_2O, mixed with 9 μL of formamide, and heated for 20 s or 1 min in a sealed glass capillary tube in boiling water. The mixture was rapidly cooled on ice, transferred to an Eppendorf tube, precipitated with ethanol, and dissolved in 2 μL of ddH_2O. The formamide-treated RNA was phosphorylated by incubating at 37 °C for 30 min in a 10 μL reaction containing 50 mM Tris·Cl (pH 7.5), 15 mM $MgCl_2$, 15 mM 2-mercaptoethanol, 20 μCi of [γ-^{32}P]ATP (3000 Ci/mmol), and 2.5 units of T4 polynucleotide kinase.

The ^{32}P-labeled RNA fragments were separated on denaturing PAGE and visualized by autoradiography. Several bands on the gel were excised and extracted in 0.4 mL of MG buffer (0.5 M NH_4OAc, 10 mM $MgCl_2$, 0.2 mM EDTA) overnight. Each eluate was transferred to an Eppendorf tube, and 10 μg of carrier tRNA was added followed by 1 mL of ethanol to precipitate RNA. The precipitate was dried and dissolved in 5 μL of 20 mM NH_4OAc (pH 5.3) containing 1 μg of P1 nuclease. The reaction mixture was incubated for 3 h at 37 °C. Two microliters of the digested RNA was analyzed by one-dimensional TLC on cellulose plates using *tert*-butyl alcohol−concentrated $HCl-H_2O$ (70/15/15 v/v/v) as eluant. The chromatograms were visualized by autoradiography.

YceC Gene Disruption. PCR was performed using *E. coli* MC1061 genomic DNA as the template (100 ng) and two primers (0.5 μM each): 5′-AACGCGTCGACTAATTCG-GTACGGTCGTCGCCT-3′, which introduced a *Sal*I site (underlined) and hybridized 1004−1025 nt upstream from the *yceC* start codon; and 5′-CACACGAGCTCCACACT-GACTACGGATTCGGCG-3′, which hybridized 993−1015 nt downstream from the *yceC* stop codon. The 3 kb PCR product was digested with *Sal*I to cleave at the 5′ primer and a natural site 589 nt 3′ to the *yceC* stop codon; the resulting 2.6 kb fragment was ligated with *Sal*I-digested pUC19 to yield pDW3. pDW3 was digested with *Mfe*I 85 nt downstream from the *yceC* start codon, and ligated with the 1.2 kb *Eco*RI fragment of pUC4K containing the Kan^r gene. The resulting plasmid (pDW2) was digested with *Sph*I and *Xba*I, and the 3.9 kb fragment was gel-purified and ligated to *Sph*I−*Xba*I digested pMAK705 to give pDW1.

Plasmid pDW1 was transformed into JM101, and gene disruption was performed as described *(11)*. Briefly, following transformation and growth at 30 °C, cultures were plated at 44 °C on medium containing 20 μg/mL Cm plus 25 μg/mL Kan. Co-integrates were resolved by three successive subcultures (1/1000 dilution) at 30 °C and then plating on media containing 25 μg/mL Kan. Colonies growing at 44 °C in Kan and having a Kan^r/Cm^s phenotype were selected. PCR analysis was performed using primers which hybridized to the 5′ and 3′ ends of *yceC*.

YfiI Gene Disruption. PCR was performed using *E. coli* MC1061 genomic DNA (100 ng) as the template and two primers: 5′-GGGGTACC**TCTAGA**GTCTTCCCGTCTTG-GTGAAT-3′, which introduced a *Kpn*I site (underlined) and an *Xba*I site (boldface) and hybridized 1109−1128 nt upstream from the *yfiI* start codon; and 5′-GAAGATCTCT-GCGCAAACCGTGCCGG-3′, which introduced a *Bgl*II site (underlined) and hybridized 283−300 nt downstream from the *yfiI* stop codon. The 2.4 kb PCR product was gel-purified, digested with *Kpn*I and *Bgl*II, and cloned into *Kpn*I−*Bam*HI-digested pUC19 to yield pDW6. Plasmid pDW6 was digested with *Bam*HI, a restriction site located 116 nt downstream from the *yfiI* start codon, and ligated with the 1.2 kb fragment containing the Kan^r gene from a *Bam*HI digest of pUC4K. This plasmid (pDW5) was digested with *Xba*I, and the 3.7 kb fragment was gel-purified and cloned into *Xba*I-digested pMAK705 to yield pDW4.

Plasmid pDW4 was transformed into JM101, and gene disruption was performed as described above with the following modifications. After successive subculturing at 30 °C to effect resolution, cells were plated at 30 °C on media containing Kan, and replicate colonies were screened on Cm or Kan media at 44 °C. Plasmid DNA was prepared from master plate colonies whose replicates grew on Kan but grew very poorly on Cm at 44 °C. These plasmids were digested with *Xba*I to distinguish wild-type from Kan-disrupted *yfiI*. Colonies with wild-type *yfiI* on the plasmid were restreaked at 44 °C on Kan and screened for a Kan^r/Cm^s phenotype. PCR analysis utilizing primers which hybridize to the 5′ and 3′ ends of *yfiI* was performed.

Measurement of Growth Rates. Single colonies of wild-type JM101, *yceC*-disrupted JM101, or *yfiI*-disrupted JM101 were grown to stationary phase in LB or M9 medium containing 25 μg/mL Kan, and then diluted 100-fold into a 50 mL culture of fresh medium. The cultures were grown at 37 °C, and the OD_{600} was measured at time points.

Analysis of Ψs in RNA. Ribosomal RNAs were isolated from the wild-type and *yceC*- or *yfiI*-disrupted strains as described *(18)*. The presence of Ψ at specific positions was determined as described *(19)* except [α-^{35}S]dATP was used instead of [α-^{32}P]dATP.

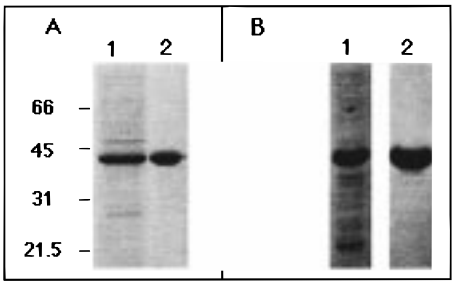

FIGURE 1: SDS−PAGE analysis of YceC and YfiI. Panel A: lane 1, crude extract of YceC; lane 2, YceC purified using the Ni^{2+}-containing affinity column. Panel B: lane 1, crude extract of YfiI; lane 2, YfiI purified by conventional chromatography.

RESULTS

Cloning and Overexpression of yceC. *yceC* was amplified from *E. coli* genomic DNA using PCR. The 5′ primer introduced a *Nde*I site overlapping the ATG start codon, and the 3′ primer introduced a *Bam*HI site following the TAA stop codon. The *Nde*I−*Bam*HI-restricted PCR fragment containing *yceC* was ligated into the *Nde*I−*Bam*HI-restricted vector pET-15b to yield pLH-YceC. DNA sequence analysis confirmed that the cloned sequence was correct. pLH-YceC contains *yceC* with an in-frame N-terminal $(His)_6$ leader peptide sequence under the control of a T7 promoter. IPTG induction of pLH-YceC in BL21(DE3)(pLysS) cells yielded YceC at about 70% of the total soluble cellular proteins (lane 1, Figure 1A). YceC-$(His)_6$, purified to homogeneity on a Ni^{2+} affinity column, migrated as a 37 kDa protein on SDS−PAGE (lane 2, Figure 1A), in agreement with the predicted molecular mass of 38 kDa. The His tag could be completely removed by thrombin cleavage for 2 h at 20 °C. The two forms of YceC were analyzed by tritium release assay. No difference was observed in the rate or extent of reaction (data not shown). Therefore, studies were performed using the enzyme with the His tag.

Cloning and Overexpression of yfiI. *yfiI* was amplified from *E. coli* genomic DNA using PCR. The 5′ primer contained a *Nco*I site overlapping the ATG start codon, and the 3′ primer contained a *Xho*I site following the TGA stop codon. The *Nco*I−*Xho*I-restricted PCR fragment containing *yfiI* was ligated into the *Nco*I−*Xho*I-restricted vector pET-15b to yield pMP-YfiI. DNA sequence analysis confirmed that the cloned sequence was correct. IPTG induction of pMP-YfiI in BL21(DE3) cells yielded expression of YfiI at about 50% of the total soluble cellular proteins (lane 1, Figure 1B). YfiI was purified to homogeneity by conventional chromatography. The homogeneous protein migrated as a 37 kDa protein on SDS−PAGE, in agreement with the predicted molecular mass (lane 2, Figure 1B).

Ψ *Formation Catalyzed by YceC and YfiI in Vitro.* Both 23S rRNA and 16S rRNA are good substrates for YceC and YfiI. When [5′-^{32}P]UMP-23S rRNA or [5′-^{32}P]UMP-16S rRNA was treated with YceC or YfiI and digested with nuclease P1, [5′-^{32}P]ΨMP was observed by 2d-TLC. After incubation of 0.25 μM RNA with 2 μM Ψ synthase at 37 °C for 3 h, Ψ formation reached a plateau. For YceC, the ratios between ΨMP and UMP reached 1:18 for 23S rRNA and 1:22 for 16S rRNA. For YfiI, the ratios between ΨMP and UMP were about 1:20 for 23S rRNA and 1:8 for 16S rRNA. Ψ formation was not observed when [5′-^{32}P]UMP-tRNA was treated with YceC or YfiI under similar conditions. Autoradiographs of the analyses of YceC- and YfiI-modified 23S rRNAs are shown in Figure 2 and Figure 3, respectively.

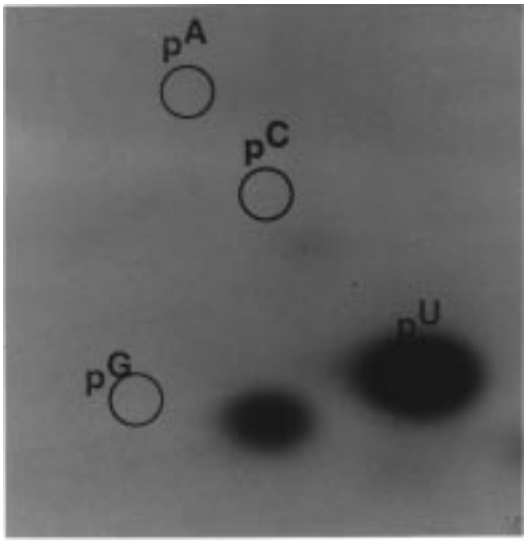

FIGURE 2: Ψ formation by YceC incubated with 23S RNA for 5 h. TLC analysis of P1 nuclease digested pseudouridylated [5′-^{32}P]-Ura RNA. The modified, digested RNA samples were loaded onto a cellulose plate, and developed using the following mobile phases: (a) isobutyric acid−concentrated NH_4OH−H_2O (66/1/33, v/v/v); (b) *tert*-butyl alcohol−concentrated HCl−H_2O (70/15/15, v/v/v). Two spots corresponding to the positions of ΨMP and UMP were detected.

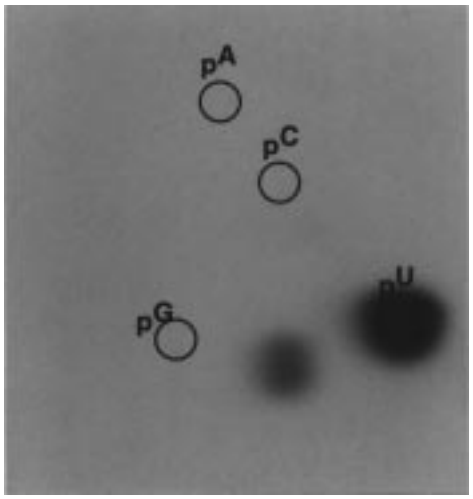

FIGURE 3: Ψ formation by YfiI incubated with 23S RNA for 5 h. Analysis as described in the legend to Figure 2.

U2504 and U2580 Are in Vitro Targets for YceC. U2504 and U2580 were identified as two of the in vitro target sites for YceC by nucleotide sequencing. The in vitro transcription product of *Bam*HI-linearized pCG50H contains stem 90h RNA followed by a 56 nt hammerhead ribozyme sequence which undergoes self-cleavage to give a homogeneous 3′ terminus *(12)*. The resulting 101 nt consists of stem 90h RNA (nt 2495−2582) followed by GAUCUUUGGAGUA-3′, which is the remaining segment of the cleaved hammerhead ribozyme. Stem 90h RNA (7.3 nmol) was incubated with 8 nmol of YceC in 500 μL of TNE buffer at 37 °C for 2 h; an additional 8 nmol of YceC was added, and

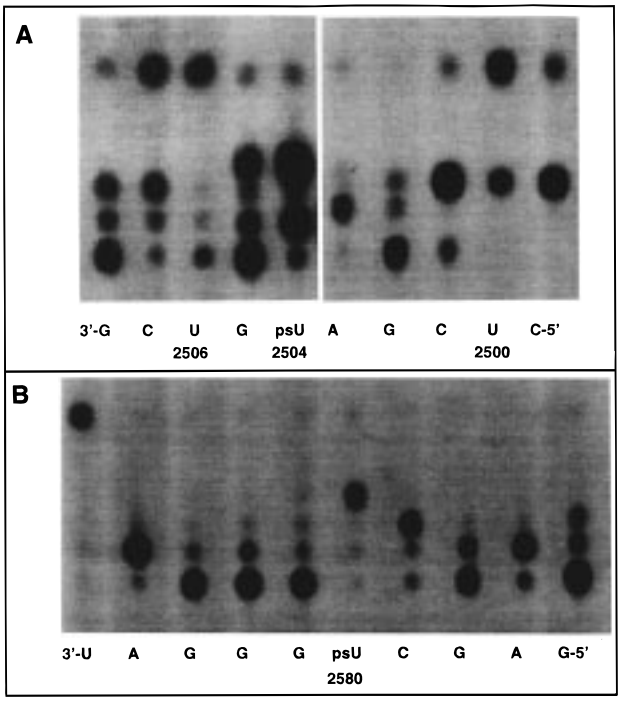

FIGURE 4: Direct nucleotide sequencing analysis of pseudouridylated stem 90h. The purified stem 90h RNA was subjected to limited formamide cleavage and phosphorylation using [γ-^{32}P]ATP. The RNA fragments were separated on a denaturing polyacrylamide gel and digested with P1 nuclease. The digested RNA was then loaded onto a cellulose plate, developed using *tert*-butyl alcohol−concentrated HCl−H_2O (70/15/15, v/v/v) as the mobile phase, and visualized by autoradiography. Panel A: TLC analysis of the nucleotide sequence around nt 2504, and identification of Ψ2504. Panel B: TLC analysis of the nucleotide sequence around nt 2580, and identification of Ψ2580. The 5′GAU3′ in this panel are from the hammerhead sequence.

the reaction mixture was incubated for an additional 1 h at 37 °C. The reaction mixture was extracted with phenol/$CHCl_3$ and precipitated with ethanol, and the RNA was purified by 12% urea−PAGE. Thirteen micrograms of YceC-treated stem 90h RNA was subjected to limited formamide cleavage for 20 s or 1 min at 100 °C, and kinased with [γ-^{32}P]ATP to label the 5′-OH groups of the fragments. The reaction mixture was loaded onto 8% urea−PAGE to separate long ^{32}P-labeled RNA fragments, or onto 20% urea−PAGE to separate the short ^{32}P-labeled RNA fragments. Approximately 10 bands were excised from the RNA ladder in each of the size ranges corresponding to the estimated positions of the RNA fragments containing the naturally modified Urd residues (Ψ2504 and Ψ2580). For Ψ2504, the bands that were close to the 5′ end of stem 90h (at the top of the RNA ladder on the 8% gel) were excised. For Ψ2580, the bands that were close to the 3′ end of stem 90h (at the bottom of the RNA ladder on the 20% gel) were excised. The excised RNA bands were extracted, ^{32}P-labeled, and digested by P1 nuclease, and the nucleotides were analyzed by TLC. Ψ was identified in the sequences 5′-CUCGAΨGUCG-3′ (Figure 4A) and 5′-GAGCΨGGGAU-3′ (Figure 4B). These are the unique sequences in stem 90h which correspond to Ψ2504 and Ψ2580, respectively, in 23S RNA. Therefore, U2504 and U2580 of 23S rRNA are both in vitro targets for YceC in vitro while the Urd residues adjacent to Ψ2504 are not.

Generation of YceC- and YfiI-Deficient Strains. *E. coli* strains lacking YceC and YfiI activities were prepared by replacement of the wild-type genes with corresponding ORFs interrupted by a Kan^r gene. Cassettes containing the interrupted target gene plus flanking chromosomal sequences were generated in pUC19 and transferred into the TS plasmid pMAK705. Integration of TS plasmids into the chromosome was selected by growth at 44 °C in Cm, and resolution of the co-integrates was achieved by growth at 30 °C. The resulting cultures contain cells with the Kan-disrupted gene in the TS plasmid plus the wild-type chromosomal gene, and cells with the wild-type gene in the TS plasmid plus the Kan-disrupted gene in the chromosome; the plasmid is cured from these cells by growth at 44 °C.

For *yceC*, the desired Kan^r/Cm^s strain was easily detected after plating the resolved liquid culture at 44 °C. PCR analysis using primers which hybridize at the 5′ and 3′ ends of the gene showed a distinct band at 2.2 kb, the molecular mass expected for the Kan-disrupted gene, and the absence of the 1 kb band observed for the wild-type gene. For *yfiI*, plating cultures at 44 °C yielded only Kan^r/Cm^r colonies with no detectable cytoplasmic plasmid, indicating the presence of co-integrates. Therefore, colonies from the resolved cultures were replicated on a master plate containing Kan and incubated at 30 °C, and separately on plates containing Kan or Cm and incubated at 44 °C. Plasmid DNA prepared from colonies from the master plate whose replicates grew well on Kan at 44 °C but poorly on Cm was analyzed. A master colony containing pMAK705 with wild-type *yfiI* on the plasmid was identified as a resolved strain with the desired Kan-disrupted chromosomal *yfiI*. Streaking this colony on Kan at 44 °C yielded plasmid-cured colonies with the desired Kan^r/Cm^s phenotype. The Kan^r/Cm^s colonies were small, indicating that disruption of *yfiI* reduced the growth rate, and that in liquid culture the Kan^r/Cm^r co-integrates containing a wild-type gene would have dominated the culture. The small colony variant was used for the studies reported below.

PCR using primers which hybridize at the ends of *yfiI* showed the 2.2 kb band expected for the Kan-disrupted gene and a band at 1 kb. Sequence analysis of the 1 kb band revealed a deletion of most of the Kan^r cassette and a frameshift of the *yfiI* reading frame. The event leading to this aberrant recombination was not identified.

Growth of YceC- and YfiI-Deficient Cells. The growth rate of the *yceC* mutant strain was compared to the parent strain in both rich (LB) and minimal (M9) media. The absence of wild-type YceC did not affect the growth rate under the conditions tested. The doubling time for JM101 was 25 min, compared with 24 min for the *yceC*-disrupted strain in LB, and 81 and 73 min, respectively, in M-9. The absence of wild-type YfiI drastically reduced the growth rate, with a measured doubling time of 77 min in LB.

Identification of in Vivo Targets for YceC and YfiI. 23S rRNA was isolated from the *yceC*- and the *yfiI*-disrupted strains, and the regions known to contain Ψ in RNA from wild-type organisms were analyzed for the presence or absence of Ψ (Figure 5). The *yceC*-disrupted strain lacks Ψ at positions 955, 2504, and 2580, and the *yfiI*-disrupted strain lacks Ψ at positions 1911 and 1917. It may also lack Ψ at 1915; in several experiments, a low intensity band was visible below position 1915 in the RNA from the *yfiI*$^-$ strain,

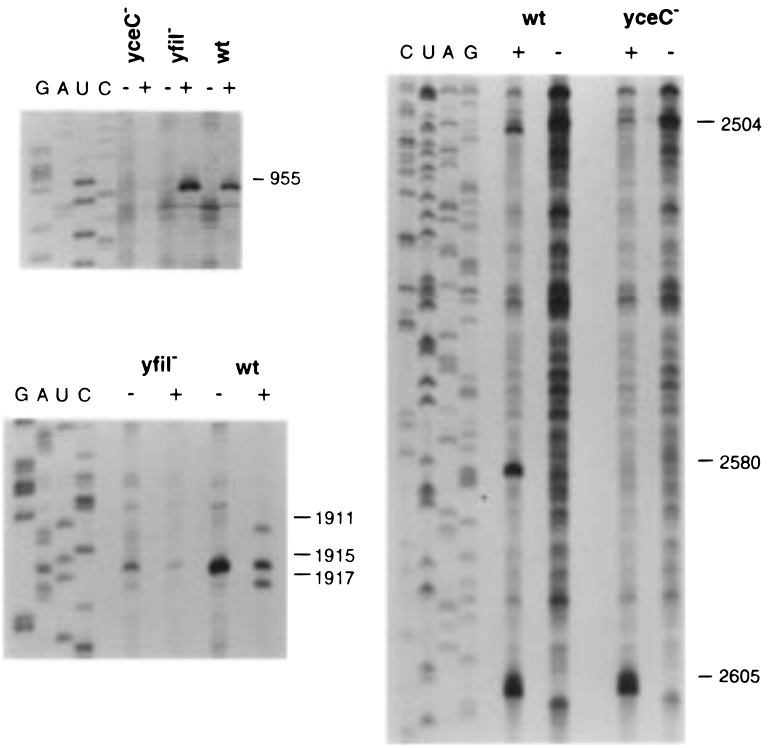

FIGURE 5: Ψ sequence analysis of 23S RNA. Ribosomal RNA was isolated from wild-type (wt) and *yceC*- or *yfiI*-disrupted strains and sequenced for the presence of Ψ as described *(19)*. The naturally occurring Ψ sites are indicated by lines adjacent to the modified Us. The CMC-induced stops occur one base below these positions. RNA for GAUC sequencing lanes was from an in vitro transcript of 23S RNA. Lanes with RNA not treated with CMC are labeled (−), and lanes with CMC-treated RNA are labeled (+).

but it was not sufficiently prominent to unequivocally distinguish it from background.[3]

DISCUSSION

Based on sequence homology analysis, the *E. coli* ORFs *yceC* and *yfiI* were proposed to encode Ψ synthases that modify 23S rRNA *(8)*. If correct, YceC and YfiI would join RluA and RsuA as verified Ψ synthases of rRNA.[2] We cloned *yfiI* and *yceC* from *E. coli* genomic DNA by PCR and expressed them in *E. coli* as about 50% and 70%, respectively, of the total soluble protein of crude extracts. Both proteins were purified to apparent homogeneity.

We tested unmodified in vitro synthesized 16S rRNA, 23S rRNA, and yeast $tRNA^{Phe}$ as substrates for YceC and YfiI. As expected, the enzymes did not catalyze Ψ formation in $tRNA^{Phe}$. With 23S rRNA as substrate, both YceC and YfiI catalyzed Ψ formation, confirming that they are indeed Ψ synthases. However, rather than converting the expected 1 or a few of the 591 Urd residues of 23S rRNA to Ψ, both enzymes catalyzed extensive modification. Under the conditions used, YceC converted an average of 1 of 18 of the Urd residues to Ψ, and YfiI converted 1 of 20 Urd residues. Moreover, both enzymes extensively modified 16S RNA which normally contains only a single Ψ residue at position 716. YceC converted ~1 of every 22 of the Urd residues of 16S RNA to Ψ, whereas YfiI converted ~1 of 8 Urd residues. Although it is not known whether ~5% of the sites are completely modified or a larger number is partially modified, it is clear that the numerous residues in both 16S and 23S rRNA modified by YceC and YfiI result from a relaxed specificity of these Ψ synthases in vitro.

To determine whether one of these Ψ synthases showed any specificity for Urd residues of 23S rRNA, we examined the YceC-catalyzed modification of a fragment of 23S rRNA containing known targets of Ψ synthases (U2504 and U2580), as well as U residues that are not normally modified. We prepared unmodified RNA corresponding to stem 90 (nt 2495−2582), which contains both Ψ2504 and Ψ2580, treated it with YceC, and subjected the suspected regions of the product to nucleotide sequence analysis. Ψ was identified at both position 2504 and position 2580, but not at position 2500 or 2506. Thus, although YceC modifies more than the expected number of Urd residues in vitro, it does not randomly modify all Urd residues of 23S rRNA. Nevertheless, the in vitro target sites for YceC and YfiI greatly exceed the in vivo sites, and it is clear that other factors are important for the in vivo specificity of these enzymes (viz., higher elements of RNA structure, associated proteins, or yet undiscovered factors).

By gene disruption experiments, we created strains of *E. coli* deficient in either YceC or YfiI, and analyzed the 23S RNA for Ψ at the sites that contain the modified nucleoside in the wild-type organism. Using this approach, we associated formation of Ψ955, 2504, and 2580 with YceC, and formation of Ψ1911 and 1917 with YfiI; YfiI may also modify U1915, but a low intensity band on the sequencing gel made this identification ambiguous.

The multisite specificities of YceC and YfiI are not without precedent. In *E. coli*, Ψ synthase I acts on U38, 39, and 40 of *E. coli* tRNA *(20)*, and in *S. cerevisiae* Pus I catalyzes Ψ

[3] Note added in proof: Using a 19 base oligonucleotide substrate, we have shown that Urd residues corresponding to U1911, U1915, and 1917 are all modified in vitro by YfiI.

formation at U34 and U36 of intron-containing precursor $tRNA^{Ileu}$, and U27/28 of several yeast tRNAs *(21)*. It is interesting that the multisite Ψ synthases modify residues in relatively close proximity, which may be a revealing clue regarding their mechanisms of specificity.

Disruption of *yceC* did not affect cell growth under the conditions tested, while disruption of *yfiI* resulted in a 3-fold decrease in growth rate in rich medium. The target Ψs of YfiI are of particular interest since they represent the most conserved Ψs in RNA, and are found in a small stem loop (the 1916 loop) in domain IV of 23S RNA proposed to be important for interactions with 16S RNA and in decoding (*23*). Most important, YfiI targets represent RNA modifications required for normal growth, and as such portend utility in understanding an important biological function of Ψ.

In summary, we cloned two *E. coli* ORFs, *yceC* and *yfiI*, from *E. coli* genomic DNA, overexpressed them in *E. coli*, and purified the recombinant proteins to homogeneity. We demonstrated that both proteins catalyze conversion of Urd residues of rRNA to Ψ. However, both enzymes show relaxed specificity in vitro in that they modify numerous Urd residues in both 23S and 16S rRNA. Since the in vitro target sites greatly exceed the possible in vivo sites, other factors must be important for the in vivo specificity of these enzymes.

Gene disruption of *yceC* and *yfiI* followed by identification of the Ψs absent in 23S RNA revealed in vivo modification sites for YceC to be at positions 955, 2504, and 2580, and for YfiI to be at positions 1911 and 1917, and possibly 1915. Together with the Ψs formed by RluA and RsuA, these enzymes account for all but two or three Ψ modifications in *E. coli* rRNA. Although the Ψs formed by YceC do not show a discernible phenotype, the Ψs formed by YfiI represent RNA modifications that are required for optimal growth. An understanding of the molecular effects of Ψ deletion at YfiI target sites in 23S RNA could finally lead to an understanding of one or several functional roles of Ψ. Such studies are in progress.

ACKNOWLEDGMENT

We thank Claes Gustafsson and Ralph Reid for helpful discussions, and we thank Ming Yu for technical assistance.

REFERENCES

1. Bjork, G. R., Ericson, J. U., Gustafsson, C. E. D., Hagervall, T. G., Jonsson, Y. H., and Wikstrom, P. M. (1987) *Annu. Rev. Biochem 56*, 263−287.
2. Bakin, A., Lane, B. G., and Ofengand, J. (1994) *Biochemistry 33*, 13475−13483.
3. Patton, J. R. (1993) *Biochem. J. 290*, 595−600.
4. Arps, P. J., Marvel, C. C., Rubin, B. C., Tolan, D. A., Penhoet, E. E., and Winkler, M. E. (1985) *Nucleic Acids Res. 13*, 5297−5315.
5. Nurse, K., Wrzesinski, J., Bakin, A., Lane, B. G., and Ofengand, J. (1995) *RNA 1*, 102−112.
6. Wrzesinski, J., Bakin, A., Nurse, K., Lane, B. G., and Ofengand, J. (1995) *Biochemistry 34*, 8904−8913.
7. Wrzesinski, J., Nurse, K., Bakin, A., Lane, B. G., and Ofengand, J. (1995) *RNA 1*, 437−448.
8. Gustafsson, C., Reid, R., Greene, P. J., and Santi, D. V. (1996) *Nucleic Acids Res 24*, 3756−3762.
9. Weitzmann, C., Cunningham, P., and Ofengand, J. (1990) *Nucleic Acids Res. 18*, 3515−3520.
10. Sampson, J. R., and Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1033−1037.
11. Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P., and Kushner, S. R. (1989) *J. Bacteriol. 171*, 4617−4622.
12. Grosshans, C. A., and Cech, T. R. (1991) *Nucleic Acids Res. 19*, 3875−3880.
13. Chu, W., and Horowitz, J. (1989) *Nucleic Acids Res. 17*, 7241−7252.
14. Huang, L. X., Pookanjanatavip, M., Gu, X. G., and Santi, D. V. (1998) *Biochemistry 37*, 344−351.
15. Gu, X. R., Yu, M., Ivanetich, K. M., and Santi, D. V. (1998) *Biochemistry 37*, 339−343.
16. Silberklang, M., Gillum, A. M., and RajBhandary, U. L. (1979) *Methods Enzymol. 59*, 58−109.
17. Kuchino, Y., Mizushima, H., and Nishimura, S. (1990) in *Transfer RNAs and other soluble RNAs* (Cherayil, J. D., Ed.) pp 69−82, CRC Press, Inc., Boca Raton, FL.
18. Moazed, D., Stern, S., and Noller, H. F. (1986) *J. Mol. Biol. 187*, 399−416.
19. Bakin, A., and Ofengand, J. (1993) *Biochemistry 32*, 9754−9762.
20. Kammen, H. O., Marvel, C. C., Hardy, L., and Penhoet, E. E. (1988) *J. Biol. Chem. 263*, 2255−2263.
21. Motorin, Y., Keith, G., Simon, C., Foiret, D., Simos, G., Hurt, E., and Grosjean, H. (1998) *RNA−Publ. RNA Soc. 4*, 856−869.
22. Ofengand, J., and Bakin, A. (1997) *J. Mol. Biol. 266*, 246−268.
23. Conrad et al. (1998) *J. Biol. Chem. 273*, 18562−18566.

BI981002N