JOURNAL OF BACTERIOLOGY, Nov. 2000, p. 6243–6246 Vol. 182, No. 21
0021-9193/00/$04.00+0


# Identification of the *ubiD* Gene on the *Escherichia coli* Chromosome

HAITAO ZHANG AND GEORGE T. JAVOR*

*Department of Biochemistry, Loma Linda University School of Medicine, Loma Linda, California 92354*

Received 19 June 2000/Accepted 1 August 2000

**The open reading frame at 86.7 min on the *Escherichia coli* chromosome, "*yigC*," complemented a *ubiD* mutant strain, AN66, indicating that *yigC* is the *ubiD* gene. The gene product, a 497-amino-acid-residue protein, showed extensive homology to the UPF 00096 family of proteins in the Swiss-Prot database.**

The biosynthetic steps of the ubiquinone (coenzyme $Q_8$ [$CoQ_8$]) pathway in *Escherichia coli* have been known for many years (11, 16). However, the identities of some ubiquinone biosynthetic genes remained hidden, even after the complete nucleotide sequence of *E. coli* was published. This was due partly to a lack of sequence data on ubiquinone biosynthetic genes in other organisms and partly to the multiplicity of open reading frames at the suspected locations. Thus, at 86.7 min on the *E. coli* chromosome, where the *ubiD* gene was mapped (3), there were no fewer than seven unidentified, hypothetical genes. A systematic search of these some years ago failed to locate the *ubiD* gene (R. Meganathan, personal communication). This was likely due to the fact that the some of the early gene assignments were incorrect. Recently the identities of the *ubiE* (8) and *ubiF* (7) genes were published, leaving *ubiD* the last of the known *ubi* genes to be identified.

We prepared a PUC 18 plasmid library which contained chromosomal fragments of *E. coli* K-12 strain AN256, the isogenic *ubiD*$^+$ strain of the *ubiD* mutant strain AN66. Strain AN66 (*thr-1 leuB6 ubiD410*) (3) was obtained from the *E. coli* Genetic Stock Center, New Haven, Conn., and strain AN256 (*thr-1 leuB6*) (9) was from C. F. Clarke's laboratory. Chromosomal fragments, 3 to 12 kDa in size, were obtained by partial digestion with *Sau*3AI restriction enzyme. Competent AN66 cells were electroporated in the presence of this plasmid library, and ampicillin-resistant transformants that could grow on plates which contained minimal medium (6), ampicillin (100 μg/ml), leucine and threonine (20 μg/ml each), and succinate (3 mg/ml) were isolated.

In our hands, AN66 cells spontaneously acquired the capacity to utilize succinate, at a frequency of 0.01%. Thus, when competent cultures of AN66 were transformed with the plasmid pBR322, approximately 300 transformed colonies which grew on succinate were obtained. These transformed revertant colonies (which grew faster on succinate plates than the *ubiD*$^+$ strain AN256) seriously interfered with the identification of *ubiD* gene-harboring transformants.

Therefore, several cycles of transformations were carried out to enrich the transformant population with *ubiD* gene-containing plasmids. This was done by recovering all transformed colonies, growing them together in Luria-Bertani medium with ampicillin, and extracting their plasmids. Competent cells were transformed by this preparation, and the procedure was repeated again. After two cycles, besides the 300 or so transformed revertant colonies, a strong haze was also seen on the succinate-containing selection plate.

Cells from this haze were cultured, and their plasmids were extracted. This plasmid preparation produced $1.3 \times 10^5$ transformed colonies that were able to grow on succinate as the sole carbon source. One of these was isolated and named AN66p522. Colony sizes of transformed cells on succinate plates were comparable to those of the *ubiD*$^+$ strain AN 256.

The ubiquinone contents of strains AN256, AN66, and AN66p522 were determined by a method described earlier (17). Cell preparation included the growth of a 500-ml culture of strain AN256 in Luria-Bertani medium plus glucose (0.3%, wt/vol) and identical volumes of AN66 and AN66p522 in brain heart infusion broth plus glucose (0.3%, wt/vol). At an $A_{600}$ of 0.9 to 1.0, the cells were harvested, washed with distilled water, and lyophilized. We found 0.29 nmol of $CoQ_8$ per mg of dried cells of strain AN256, 0.05 nmol of $CoQ_8$ per mg of dried cells of strain AN66, and 0.73 nmol of $CoQ_8$ per mg of dried cells of strain AN66p522. Thus the chromosomal fragment on plasmid p522 fully complemented the ubiquinone deficiency of AN66.

Sequencing of this chromosomal fragment showed that it was 2,595 bp long. It started near the end of the open reading frame *yigW*_2, 146 bases downstream from the end of the *rfaH* gene (a regulatory gene of lipopolysaccharide, sex factor, and hemolysin genes, oriented in the opposite direction from *yig*W_2) and ended 252 bases into the *fre* gene. The only other open reading frame located between *rfaH* and *fre* was *yigC*. It is our contention that this 1,491-base segment, immediately upstream from the *fre* gene at 86.7 min on the *E. coli* chromosome, is the *ubiD* gene (Fig. 1).

Until recently, the *fre* gene, coding for NAD(P)H flavin oxidoreductase, was designated *ubiB* (5). However, the true

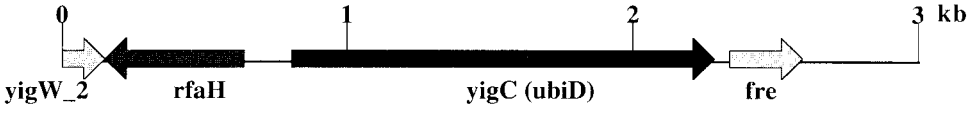


FIG. 1. Composition of the 2,595-bp-long *E. coli* chromosomal fragment in p522, showing the location of the *ubiD* gene.

* Corresponding author. Mailing address: Department of Biochemistry, Loma Linda University School of Medicine, Loma Linda, CA 92354. Phone: (909) 796-7311, ext. 48663. Fax: (909) 824-4887. E-mail: gjavor@som.llu.edu.

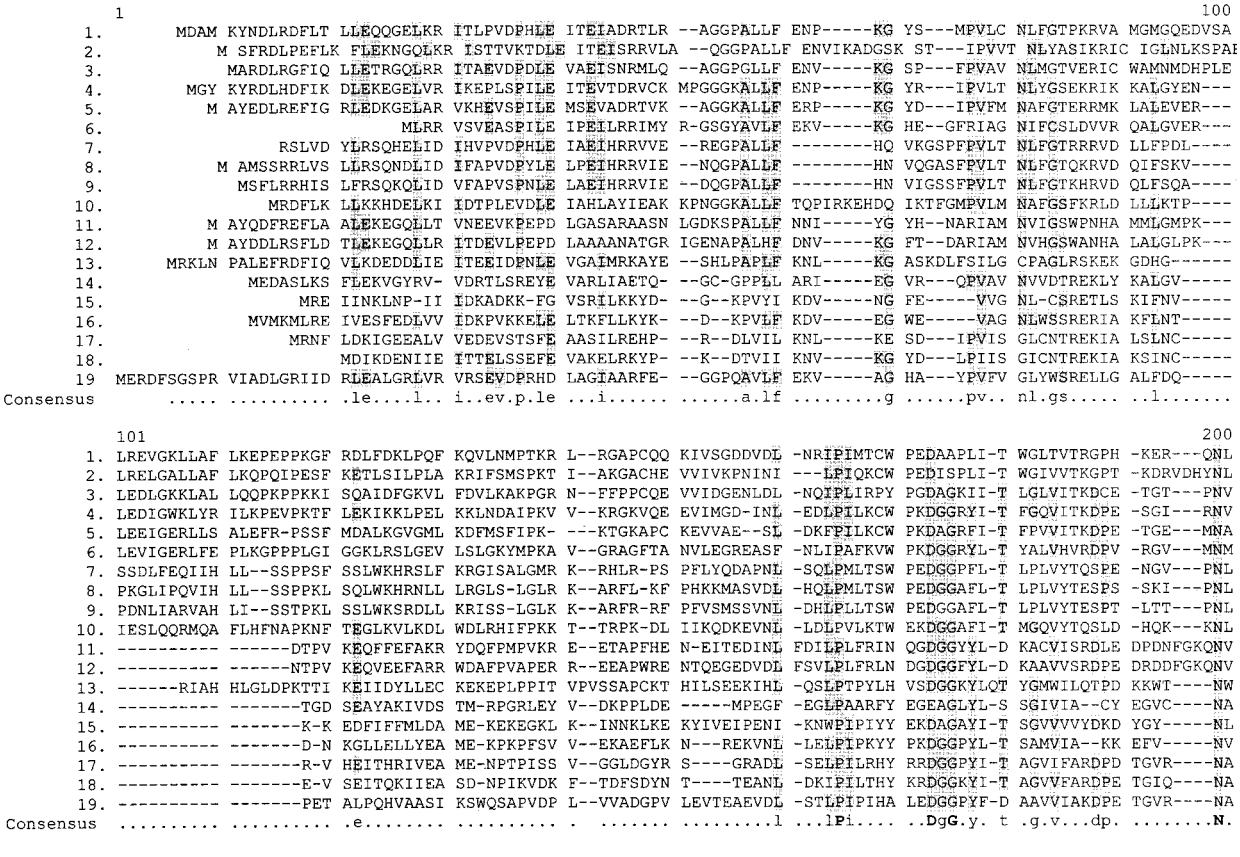

FIG. 2. Homology between the *ubiD* gene product and 18 members of the UPF 00096 family of proteins, as arranged by the Multaline program. The identity of each protein is shown in Table 1. Lowercase letters designate consensus amino acids in 50 to 90% of the sequences, and boldface capital letters indicate 90 to 100% homology.

FIG. 2—*Continued.*

*ubiB* gene is now shown to be the former open reading frame *yigR*, approximately 6 kb downstream from its previous location (12).

The *yigC* segment and its upstream region were isolated by PCR from the AN256 chromosome (primers used were as follows: in the forward direction, 5′-GATCATCGGTGCCAG GCAATTCACAGCC-3′, in the reverse direction, 5′-TCAGG CGCTTTTACCGTTGTTAAAA-3′). It was cloned into a pNoTA/T7 shuttle vector (manufactured by 5 Prime → 3 Prime Inc.), and this construct was transformed into AN66 cells. This plasmid, designated p613, complemented the *ubiD* mutant trait of AN66 cells to the same extent as the larger plasmid, p522.

Based on its nucleotide sequence, the product of *ubiD* gene is a 497-amino-acid protein, its molecular mass is 55,603.7 Da, and its theoretical isoelectric point is 5.31. The *ubiD* gene product is one of two enzymes (3-octaprenyl-4-hydroxybenzoate carboxy-lyase) which catalyze the decarboxylation of 3-octaprenyl-4-hydroxy benzoate to 2-octaprenylphenol. Earlier work with this enzyme suggested that it is a membrane-associated protein, although during cell fractionation much activity was found in the cytoplasmic fraction (9). Analysis of its amino acid sequence for transmembrane helices indicated zero (13), one (positions 215 to 235) (K. Hofmann and W. Stoffel, Biol. Chem. Hoppe-Seyler **347:**166, abstr. MF C-35), or two (positions 226 to 232 and 334 to 340) (4) such regions, depending on which program was used. This enzyme's molecular mass by gel filtration measurement was reported to be approximately 340,000 Da (9). This suggests that it is a hexameric protein in vivo.

We isolated the *ubiD* gene from strain AN66 by PCR and sequenced it, for the purpose of locating the site of mutation. The long gene was sequenced in overlapping segments, and the last fragment was sequenced in the reverse direction as well. (The following primers were used: 1, 5′-ATGGACGCCATG AAATATAACGATT-3′; 2, 5′-GCGTGGCGATGGGCATG GGGCAGG-3′; 3, 5′-GCATTCCCATTATGACCTGCTGG CCGG-3′; 4, 5′-GGTGCCGATCCCGCCACGATTCTCGG-

TABLE 1. Proteins of UPF0096 family

| Organism | Name of protein | Primary accession no. |
|---|---|---|
| *Escherichia coli* | YigC (UbiD) | |
| *Rickettsia prowazekii* | Y821_RICPR | Q9ZCD6 |
| *Synechocystis* sp. (strain PCC6803) | Y936_SYNY3 | P72861 |
| *Aquifex aeolicus* | Y612_AQUAE | 067542 |
| *Archeoglobus fulgidus* | Y209_ARCFU | 030030 |
| *Aeropyrum pernix* | YF71_ARCFU | Q9YBM7 |
| *Chlamydia trachomatis* | Y085_CHLTR | 084087 |
| *Chlamydia psitacci* | Y66K_CHLPS | 034023 |
| *Chlamydia pneumoniae* | Y328_CHLPN | Q9Z8L0 |
| *Helicobacter pylori* | Y396_HELPY | 025157 |
| *Bacillus subtilis* | YCLC_BACSU | P94405 |
| *Streptomyces* sp. (strain D7) | VDCC_STRD7 | Q9X697 |
| *Saccharomyces cerevisiae* | YD39_YEAST | Q03034 |
| *Aeropyrum pernix* | YK78_AERPE | Q9YA60 |
| *Methanococcus janaschii* | YB33_METJA | Q58533 |
| *Pyrococcus horiochii* | Y963_PYRHO | 058701 |
| *Methanobacterium thermoautotrophicum* | YD94_METTH | P41655 |
| *Methanobrevibacter smithii* | YPUE_METSM | P22349 |
| *Rhodospirillum rubrum* | YCOM_RHORU | P72315 |

3′; and 5, 5-GGGCGTCCGCCAGATGAGCCCGCGGCGG TG-3′ [all forward direction] and 5′-TCAGGCGCTTTTACC GTTGTTAAAA-3′ [reverse direction].) Comparison of the results with the published nucleotide sequence showed a single deviation. Codon 452, GGG (glycine), appeared in the mutant as AGG (arginine). A homology study (see below) shows that this glycine is a consensus residue.

A comparison (14) between the amino acid sequences of the *ubiX* and *yigC* (*ubiD*) gene products showed no significant similarities. A Blast search of the entire protein database of 525,243 sequences yielded 86 hits of similarities with *yigC* (*ubiD*). One of these hits was a 29% sequence identity and 48% similarity with a 4-hydroxybenzoate decarboxylase from *Clostridium hydroxybenzoicum*.

In the Swiss-Prot database *yigC* is listed as a member of the family UPF (uncharacterized protein family) 00096, with a taxonomic range of archaea, eubacteria, and eukaryota. Table 1 lists these proteins, and Fig. 2 shows regions of homology among them, as assigned by the Multaline program (2). Clearly not all of these proteins can be functional homologues of *ubiD*. Two proteins are from the same organism (*Aeropyrum pernix*). The gram-positive organisms *Bacillus subtilis* and *Streptomyces* sp. synthesize menaquinones, and the cyanobacterium *Synechocystis* sp. makes plastoquinone instead of ubiquinone (1). *Helicobacter pylori* also utilizes menaquinone instead of ubiquinone (10). Archaebacteria also employ a variety of ubiquinone analogues (15). However, due to the extensive homology between these proteins, it is reasonable to expect that they all function in some membrane-associated decarboxylation process.

We thank R. Meganathan for communicating unpublished results to us. We are grateful to Catherine F. Clarke for strain AN256 and also for sending us the manuscript of reference 12 before it was published.

## REFERENCES

1. **Collins, M. D., and D. Jones.** 1981. Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implications. J. Bacteriol. **45:**316–354.
2. **Corpet, F.** 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. **16:**10881–10890.
3. **Cox, G. B., I. G. Young, L. M. McCann, and F. Gibson.** 1969. Biosynthesis of ubiquinone in *Escherichia coli* K-12, location of genes affecting the metabolism of 3-octaprenyl-4-hydroxybenzoic acid and 2-octaprenyl phenol. J. Bacteriol. **99:**450–458.
4. **Cserzo, M., E. Wallin, I. Simon, G. von Heijne, and A. Elofsson.** 1997. Prediction of transmembrane alpha helices in procaryotic membrane proteins: the Dense Alignment Surface method. Protein Eng. **10:**673–676.
5. **Daniels, D. L., G. Plunkett III, V. Burland, and F. R. Blattner.** 1992. Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes. Science **257:**771–778.
6. **Davis, B. D., and E. S. Mingioli.** 1950. Mutants of *Escherichia coli* requiring methionine or vitamin $B_{12}$. J. Bacteriol. **60:**17–28.
7. **Kwon, O., A. Kotsakis, and R. Meganathan.** 2000. Ubiquinone (coenzyme Q) biosynthesis in *Escherichia coli*: identification of the *ubi*F gene. FEMS Microbiol. Lett. **186:**157–161.
8. **Lee, P. T., A. Y. Hsu, H. T. Ha, and C. F. Clarke.** 1997. A C-methyltransferase involved in both ubiquinone and menaquinone biosynthesis: isolation and identification of the *Escherichia coli ubiE* gene. J. Bacteriol. **179:**1748–1754.
9. **Leppik, R. A., I. G. Young, and F. Gibson.** 1976. Membrane-associated reactions in ubiquinone biosynthesis in *Escherichia coli*. Biochim. Biophys. Acta **436:**800–810.
10. **Marcelli, S. W., H. T. Chang, T. Chapman, P. A. Chalk, R. J. Miles, and R. K. Poole.** 1996. The respiratory chain of *Helicobacter pylori*: identification of cytochromes and the effects of oxygen on cytochrome and menaquinone levels. FEMS Microbiol. Lett. **138:**59–64.
11. **Meganathan, R.** 1996. Biosynthesis of the isoprenoid quinone menaquinone (vitamin $K_2$) and ubiquinone (coenzyme Q), p. 642–656. *In* F. C. Neidhardt et al. (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press Washington, D.C.
12. **Poon, W. W., D. E. Davis, H. T. Ha, T. Jonassen, P. N. Rather, and C. F. Clarke.** 2000. Identification of *Escherichia coli ubiB*, a gene required for the first monooxygenase step in ubiquinone biosynthesis. J. Bacteriol. **182:**5139–5146.
13. **Sonnhammer, E. L. L., G. von Heine, and A. Krogh.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences, p. 175–182. *In* J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen (ed.), Proceedings of the Sixth International Conference on Intelligent systems for molecular biology. AAAI Press, Menlo Park, Calif.
14. **Tatusova, T. A., and T. L. Madden.** 1999. Blast 2 sequences—a new tool for comparing protein and nucleotide sequences. FEMS Microbiol. Lett. **174:**247–250.
15. **Thurl, S., W. Whitke, I. Buhrow, and W. Schafer.** 1986. Quinones from archaebacteria. II. Different types of quinones from sulphur-dependent archaebacteria. Biol. Chem. Hoppe-Seyler **367:**191–197.
16. **Young, I. G., P. Strobant, P. C. G. Macdonald, and F. Gibson.** 1973. Pathway for ubiquinone biosynthesis in *Escherichia coli* K-12: gene-enzyme relationships and intermediates. J. Bacteriol. **114:**42–52.
17. **Zeng, H., I. Snavely, P. Zamorano, and G. T. Javor.** 1998. Low ubiquinone content in *Escherichia coli* causes thiol hypersensitivity. J. Bacteriol. **180:**3681–3685.