Isolation and characterization of *Escherichia coli* mutants affected in aerobic respiration: the cloning and nucleotide sequence of *ubiG*

Identification of an *S*-adenosylmethionine-binding motif in protein, RNA, and small-molecule methyltransferases

GUANGHUI WU,¹ HUW D. WILLIAMS,^{1†} MARYAM ZAMANIAN,^{1‡} FRANK GIBSON² and ROBERT K. POOLE^{1*}

¹Microbial Physiology Research Group, King's College London, Campden Hill Road, London W8 7AH, UK

²Division of Biochemistry and Molecular Biology, John Curtin School of Medical Research, Australian National University, Canberra, ACT 2601, Australia

(Received 7 April 1992; revised 10 June 1992; accepted 26 June 1992)

We report the isolation and characterization of a mutant of *Escherichia coli* unable to grow aerobically on non-fermentable substrates, except for very slow growth on glycerol. The mutant contains cytochrome oxidases *o* and *d*, and grows anaerobically with alternative electron acceptors. Oxygen consumption rates of cell-free extracts were low relative to activities in an isogenic control strain, but were restored *in vitro* by adding ubiquinone-1 to cell-free extracts. Transformation with a cloned 2.8 kb *Clal*–*EcoRV* fragment of chromosomal DNA restored the ability of this mutant (AN2571) to grow on succinate and also restored cellular quinone levels in this strain. The plasmid also complemented a previously isolated *ubiG* mutant (AN151) for aerobic growth on succinate. The nucleotide sequence revealed a 0.7 kb portion of *gyrA*. Unidirectional nested deletions from this fragment and complementation analysis identified an open reading frame encoding a protein with a predicted molecular mass of 26.5 kDa. This gene (*ubiG*) encodes the enzyme 2-octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinone methyltransferase, which catalyses the terminal step in the biosynthesis of ubiquinone. The open reading frame is preceded by a putative Shine–Dalgarno sequence and followed by three palindromic unit sequences. Comparison of the inferred amino acid sequence of UbiG with the sequence of other *S*-adenosylmethionine (AdoMet)-dependent methyltransferases reveals a highly conserved AdoMet-binding region. The cloned 2.8 kb fragment also contains a sequence encoding the C-terminus of a protein with 42–44% identity to fungal acetyl-CoA synthetases.

Introduction

The diversity of energy-producing pathways in *Escherichia coli* has facilitated genetic studies of respiratory electron transfer and ATP synthesis (Cox & Downie,

1979, Poole & Ingledew, 1987), since mutants totally defective in aerobic respiration may be grown fermentatively or anaerobically with alternative electron acceptors. Similarly, mutations affecting either of the two aerobically-functioning terminal oxidases (cytochromes *bd* and *bo*) can be isolated, since either oxidase alone can support aerobic respiration. Mutants unable to synthesize ubiquinone, an essential intermediate in the transfer of reducing equivalents from low-potential dehydrogenases to the oxidase complexes, were isolated early on (see Gibson, 1973) as being able to grow on glucose by fermentation, but not on malate or succinate.

We set out to isolate mutants affected in the aerobic respiratory pathway terminated by the cytochrome *bo* oxidase complex. By mutagenizing a strain carrying a mutation in *cydD* (Poole *et al.*, 1989), one of four genes known to be required for expression and function of the

* Author for correspondence. Tel. 071 333 4275; fax 071 937 5396.

† Present address: Department of Biology, Imperial College of Science, Technology and Medicine, London SW7 2BB, UK.

‡ Present address: National Institute for Medical Research, Mill Hill, London NW7 1AA, UK.

Abbreviations: Cm, chloramphenicol; Tc, tetracycline; Ap, ampicillin; AdoMet, *S*-adenosylmethionine; Suc[–], not able to grow on succinate minimal medium; Suc⁺, able to grow on succinate minimal medium; PU, palindromic unit.

The nucleotide sequence data reported in this paper, beginning at nucleotide 696 in Fig. 3, have been submitted to GenBank and have been assigned the accession number M87509.

alternative *bd* oxidase complex, mutations affecting electron transfer to or within the cytochrome *bo* complex could be selected. In such a search, retention of the ability to grow anaerobically with fumarate or nitrate can be used (Au *et al.*, 1985) to eliminate mutants affected in haem B biosynthesis or oxidative phosphorylation, since these functions are required for both aerobic and anaerobic respiratory energy conservation. Using a similar approach, Au *et al.* (1985) obtained mutants defective in cytochrome *o* and thus identified a new operon (*cyo*). We report here the isolation and characterization of a mutant unable to grow aerobically on non-fermentable substrates. The mutation responsible has been identified in *ubiG*, which encodes the enzyme for the terminal step in ubiquinone biosynthesis and which was originally described by Stroobant *et al.* (1972). The cloning and complete nucleotide sequence of *ubiG* is reported, together with an analysis of the sequence, which reveals the presence of an *S*-adenosylmethionine-binding site in UbiG and of three palindromic unit sequences following the *ubiG* coding region.

Methods

Bacterial strains and plasmids. The bacterial strains used (Table 1) were derivatives of *E. coli* K12. Generalized transduction by bacteriophage P1_{vir} was as described by Miller (1972) and Silhavy *et al.* (1984). Plasmids are described in Table 2.

Growth media. LB liquid medium and LB plates were as described by Miller (1972). MG rich plates (meat infusion agar with glucose) were as described by Poole *et al.* (1989). For anaerobic growth, MG was modified by omitting glucose and adding glycerol (0.5%), nitrate (0.5%), and 1 µM-ammonium molybdate and 1 µM-potassium selenite (MGN). Succinate minimal medium (SMM) was described by Poole *et al.* (1989). In glycerol minimal medium, succinate was replaced with 0.5% glycerol (GMM). Where indicated, SMM was supplemented with 0.5% Casamino acids (SCMM). For anaerobic growth, succinate was replaced with 0.5% glycerol, the Casamino acids were omitted, and the medium was further supplemented with 1% nitrate, 0.5% fumarate, and Mo/Se as described above (NFGMM). Methionine (0.15 mM final) was added to minimal media for growth of strain AN151. Solidified medium contained 1% agar. Ampicillin (Ap) at 100 µg ml⁻¹ and chloramphenicol (Cm) at 32 µg ml⁻¹ (final concentrations) were added to media as necessary for plasmid selection.

Mutant isolation. The cytochrome *d*-deficient mutant AN2343 was grown from an overnight starter culture in SMM to a Klett value of 100 (OD₆₀₀ ~0.5). A portion (5 ml) of culture was centrifuged and the cells were resuspended in an equal volume of citrate buffer (pH 5.5; Miller, 1972). Nitrosoguanidine was added to a final concentration of 48 µg ml⁻¹ and the suspension incubated without shaking at 37 °C for 15 min, giving >60% killing. The cells were washed twice in 0.1 M-potassium phosphate buffer (pH 7.0). After outgrowth for 5 h (about 2 generation times), cells were pelleted, washed once again in phosphate buffer and diluted at 10⁻⁶ in molten SMM agar (Cox & Downie, 1979). Aliquots (4 ml) were poured onto prewarmed SMM plates and incubated at 37 °C aerobically. After 3 d, colony positions were spotted with a marker pen and each plate was overlaid with 5 ml NFGMM agar. After a further 6 d incubation in anaerobic jars, twice evacuated and

filled with H₂/CO₂, previously unmarked colonies were removed either in a plug of agar (using a sterile Pasteur pipette) or with a straight, sterile wire inserted through the overlay agar. Putative respiration-deficient mutants were purified anaerobically on NFGMM plates and scored for ability to grow aerobically on SMM plates. One of the isolates was designated AN2571. Strain AN2572 was constructed by preparing a P1_{vir} lysate on AN1495 (*recA::Tn10*) and transducing AN2571. Tc^R transductants were selected and screened for the RecA⁻ phenotype (UV sensitivity).

Isolation of *Tn10dTet* insertions close to *ubiG*. To obtain a strain carrying a single mutation, an easily selectable marker was inserted adjacent to *ubiG*. A P1_{vir} lysate was made on GE1031 (*Tn10dTet* donor) and used to transduce MG1655/pNK972 (which overproduces *Tn10* transposase) to Tet^R. A P1_{vir} lysate was then made on a pool of ~3000 Tc^R transductants and used to retransduce AN2571 to Tet^R. Transductants able to grow aerobically on SMM have *Tn10dTet* linked to the wild-type allele. A fresh lysate was made on a purified transductant and used to retransduce AN2571 to Tet^R; transductants were screened for inability to grow on SMM. HW271 was isolated as one such transductant having *Tn10dTet* linked to the mutant *ubiG* allele with a cotransduction frequency of 85%. HW272 was isolated from the same procedure as having *Tn10dTet* linked to the *ubiG*⁺ allele.

Cloning of *ubiG*. Cloning methods, restriction endonuclease digestion, plasmid isolation, and transformations were essentially as described by Maniatis *et al.* (1982). Fragments of DNA were isolated from agarose gels using a GeneClean kit from BIO 101 Inc. Small-scale plasmid preparations were performed according to the method of Birnboim (1983). Chromosomal DNA from strain AN2540 was partially digested with *Hind*III and ligated to 5'-dephosphorylated, *Hind*III-linearized vector pHC79. The resulting concatemers were packaged into bacteriophage λ particles and used to transfect strain AN2572. Ap^R transfectants were screened for a Suc⁺ phenotype on SMM. One (of twelve) such transfectants was used to prepare cosmid DNA (pRP7) and the resulting preparation again used to transform strain AN2572 to the Suc⁺ phenotype. A clone from a single colony of transformants was selected and used to prepare cosmid DNA which was then linearized with *Cla*I. Plasmid vector pBR328 was 5'-dephosphorylated with calf intestinal alkaline phosphatase (Maniatis *et al.*, 1982) and ligated with the *Cla*I-derived fragments using DNA ligase. Strain AN2572 was grown in MGN medium, then transformed with the ligation product; transformants were selected on SCMM-Cm. One such transformant, shown to contain a 9 kb plasmid, was retained as strain RP112 containing plasmid pRP19. Plasmid pRP19 was digested with *Eco*RV and the fragments recircularized with DNA ligase. The ligation products were used to transform strain 71-18; transformants were selected on MG-Cm and screened for the presence of plasmids that gave the expected fragments on digestion with restriction endonucleases. Plasmid pMZ26 (2.8 kb insert) was able to transform AN2571 and HW271 to a Suc⁺ phenotype on SCMM-Cm.

DNA sequence determination. The *Cla*I-*Eco*RV insert in pMZ26 was cloned into the polylinker region of pBluescript II (Stratagene), and designated pRPB. Deletions in the cloned fragment were obtained using a Nested Deletion kit (Pharmacia). Exonuclease III (5 U µl⁻¹) was used without added NaCl at 37 °C, and samples were taken at 1 min intervals. Alkali-denatured double-stranded plasmid DNA (Murphy & Ward, 1989) was used directly as template. DNA sequence was determined by the method of Sanger *et al.* (1977) using reagents and Sequenase in a kit supplied by USB and [³⁵S]dATPαS (Amersham) for detection. Annealing of primer and template was achieved by incubation at 37 °C for 15 min or by boiling for 1 min, followed by immediate transfer to ice. The boiling method was preferable for overcoming sequencing problems arising from DNA secondary structure. Regions of the cloned fragment inaccessible from the deletions were sequenced using synthetic oligonucleotide primers.

Oxygen consumption measurement. Cells for oxygen consumption measurement were grown in 400 ml GMM to early stationary phase. Cells were washed and resuspended in 50 mM-phosphate buffer, and then sonicated, while chilled in an ice-salt slurry, five times, each for 15 s interspersed with 1 min intervals of cooling. The cell debris and unbroken cells were removed by centrifugation at 10000 *g* for 10 min. The supernatant fraction (cell-free extract) was stored at -70°C . Oxygen consumption was measured according to Poole (1977). Substrates were used at a final concentration of 20 mM; where indicated, ubiquinone-1 (reduced as an ethanolic solution with a few grains of NaBH_4) was added to a final concentration of 0.13 mM approximately 2 min after substrate addition when a steady rate of respiration had been achieved.

Extraction and analysis of quinones. Cells were grown in 1 litre of '56' medium (Gibson *et al.*, 1977) supplemented with 1% glucose and 0.1% Casamino acids in a 2 litre flask with shaking until the OD_{600} had reached 1.5, and were then washed with 50 mM-potassium phosphate buffer (pH 7.0) and stored for up to 2 d at -20°C . Extraction was performed as described by Kroppenstedt (1982). Thawed cells were suspended in 35 ml acetone (-20°C) and sonicated, while chilled in an ice-salt slurry, five times, each for 15 s interspersed with 1 min intervals of cooling. Cell debris was removed by filtration through Whatman no. 1 paper; the filtrate was evaporated to ~ 0.5 ml in a rotary evaporator at $\sim 35^{\circ}\text{C}$. The sample, which became turbid, was freeze-dried and the residue dissolved in acetone. Samples were applied to a silica gel F_{254} plastic-backed TLC plate (Merck Art. 5735) which was run and developed in hexane/diethyl ether (85:15, v/v; Collins, 1985). The spots, visualized by UV transillumination, were cut out and the quinones eluted with 100% ethanol. The silica powder was removed by centrifugation and spectra of the clear supernatant fractions were recorded with a Varian DMS90 spectrophotometer.

Assay of ubiquinol oxidase activity. The method was based on that of Kita *et al.* (1984). The reaction mixtures in each of two quartz cuvettes comprised 2.5 ml 60 mM-Tris/HCl (pH 7.5) to which was added 10 μl ubiquinol-1 (final concn 93 μM). This had been previously prepared by adding to 1 ml of an ethanolic solution of the quinone a few grains of NaBH_4 and storing on ice until the solution became colourless and hydrogen bubbles had dissipated. The differential rate of quinol autoxidation was recorded in a split-beam spectrophotometer at 275 nm for a few minutes before adding 10–20 μl of the cell extract to one cuvette only.

Cytochrome determinations. Cells were grown in LB medium plus 0.7% glucose in shaken Erlenmeyer flasks to give OD_{600} values (after 1 in 5 dilution), of 0.20 for *ubi* mutants and 0.55 for controls. Cells were harvested, then washed with 50 mM-Tris buffer (pH 7.5) and suspended in the same medium for spectroscopy. Spectra were obtained in a Johnson Foundation dual-wavelength scanning spectrophotometer (SDB3) as described by Williams & Poole (1987). The reference wavelength was 575 nm and the scan rate 2.86 nm s^{-1} . For cytochrome *d*, the wavelength pair 630–650 nm was measured in reduced *minus* oxidized spectra, using $\epsilon = 18.8 \text{ mM}^{-1} \text{ cm}^{-1}$. CO difference spectra were used to estimate the levels of cytochrome *b*-595 (440–470 nm, $\epsilon = 52.8 \text{ mM}^{-1} \text{ cm}^{-1}$) and cytochrome *o* (415–430 nm, $\epsilon = 145 \text{ mM}^{-1} \text{ cm}^{-1}$). We estimate the coefficient of variation for these measurements to be about 10% (A. Gil & R. K. Poole, unpublished).

Protein measurement. The procedure of Markwell *et al.* (1978) was used.

Chemicals and enzymes. T4 DNA ligase (from BRL) and all other enzymes (NBL) were used in buffers as specified by the suppliers. Alkaline phosphatase was from Boehringer. Sequencing gel solutions were from National Diagnostics. Ubiquinone-1 was a generous gift from Hoffman LaRoche Co., Basel, Switzerland.

Results

Mutant isolation and preliminary characterization

Of 13118 NTG-derived mutants of AN2343 (*cydD*) screened, 353 (2.8%) failed to grow aerobically on SMM (Suc^-) but grew anaerobically on a defined medium (NFGMM) that contained fumarate and nitrate as terminal electron acceptors. Of those mutants that on subculturing repeatedly failed to grow on succinate minimal medium (SMM), about one-third also failed to grow, or grew very slowly, on glycerol (GMM). Representatives of this class were tested for cytochrome levels and oxidase activities. The strains that did not grow on succinate but that grew well on glycerol were not studied further; many are likely to have been succinate dehydrogenase mutants. One of the selected mutants, AN2571, which failed to grow on minimal media with succinate, glycerol or lactate as sole carbon source, grew to give small colonies on a rich glucose medium (MG), and grew anaerobically with nitrate as sole terminal electron acceptor. The ubiquinol oxidase specific activity in AN2571 was only 3% of wild-type levels, and NADH oxidase, measured spectrophotometrically, was undetectable. Cytochrome *o* levels were 16% of wild-type levels. These findings are consistent with a *cyo* or *ubi* mutation. The affected gene was cloned by complementation of AN2571 using a cosmid library and subcloned into pBR328 to give pRP19 as described in Methods.

Tn10dTet insertions were made adjacent to the mutation in AN2571 as described in Methods. The mutation was cotransduced with Tn10dTet into the wild-type AN2342. About 85% of Tet^R transductants were unable to grow aerobically on SMM, indicating that the mutant phenotype is not dependent on the presence of the *cydD* mutation in AN2343 and, therefore, that the mutation is unlikely to be in *cyo*. Two Tet^R transductants, HW272 and HW271, were isolated as being able or unable, respectively, to grow aerobically on SMM. HW271 was also unable to grow on lactate as sole carbon source, but grew very slowly on GMM. Ubiquinol oxidase activities of HW271 were not significantly different from those of HW272 (results not shown). Presumably, the phenotype of AN2571 is a consequence of additional mutation(s) following NTG mutagenesis.

Cytochrome levels and oxygen consumption rates

In contrast to the preliminary measurements of cytochrome levels in AN2571, from which the mutant allele had not been transferred to a genetically 'clean' background, the cytochrome *o* level in HW271 was slightly higher than in the wild-type strain AN2342 (Table 3). However, both cytochrome *d* and cytochrome

Table 1. *Escherichia coli* strains

Strain	Genotype	Origin* or reference
71-18	$\Delta(lac-proAB) thi supE [F' proAB lac^qZ\Delta M15]$	Yanisch-Perron <i>et al.</i> (1985)
AN151	<i>ubiG met</i>	Leppick <i>et al.</i> (1976)
AN1495	<i>unc::mu, argH pyrE entA Str^R recA::Tn10</i>	F. Gibson
AN2342	F ⁻ (referred to as wild-type)	Poole <i>et al.</i> (1989)
AN2343	F ⁻ <i>cydD</i>	Poole <i>et al.</i> (1989)
AN2540	<i>uncC538 argH pyrE entA</i>	Cox <i>et al.</i> (1987)
AN2571	<i>cydD ubiG</i> from AN2343	This work
AN2572	F ⁻ <i>cydD recA ubiG</i>	P1 (AN1495) $\times$ AN2571, this work
GE1031	<i>supE42/Fzzf-1831::Tn10dTet</i>	G. M. Weinstock
HW271	<i>ubiG zei::Tn10dTet</i> from AN2571	This work
HW272	<i>ubiG⁺ zei::Tn10dTet</i>	This work
MG1655	F ⁻	G. M. Weinstock
RP112	AN2572/pRP19	This work

* G. M. Weinstock, University of Texas Medical School, Houston, Texas, USA.

Table 2. Plasmids

Plasmid	Description	Origin or reference
pBluescript II SK +	Ap ^R	Stratagene
pBR328	Plasmid vector Tc ^R Ap ^R Cm ^R	Bolivar <i>et al.</i> (1977)
pEC10	As pRPB, but with 1.262 kb deletion extending from <i>EcoRV</i> site	This work
PEI110	As pRPB, but with 1.009 kb deletion extending from <i>EcoRV</i> site	This work
pHC79	Cosmid cloning vector Tc ^R Ap ^R	Hohn & Collins (1980)
pMZ26	<i>ubiG⁺</i> , pRP19 digested with <i>EcoRV</i> and recircularized	This work
pND31	As pRPB, but with 1.295 kb deletion extending from <i>Clal</i> site	This work
pND42	As pRPB, but with 1.628 kb deletion extending from <i>Clal</i> site	This work
pNK972	Ap ^R , transposase expressed from <i>tac</i> promoter	Elliott & Roth (1988)
pRP7	Cosmid (pHC79) clone carrying the wild-type region corresponding to the mutation in AN2571	This work
pRP19	<i>ubiG⁺</i> , approx. 4 kb <i>Clal-Clal</i> fragment from cosmid clone pRP7 in pBR328	This work
pRPB	2.818 kb <i>Clal-EcoRV</i> insert from pMZ26 subcloned into pBluescript II	This work

Table 3. Cytochrome levels of HW271 and AN2342

	Cytochrome level [nmol (mg protein) ⁻¹]	
	AN2342	HW271
Cytochrome <i>d</i>	0.258	0.126
Cytochrome <i>b</i> -595/ haemoprotein <i>b</i> -590	0.154	0.072
Cytochrome <i>o</i>	0.093	0.132

b-595 (or haemoprotein *b*-590, which are not distinguishable in CO difference spectra of intact cells: Lorence *et al.*, 1986; Poole *et al.*, 1986) were reduced approximately twofold in HW271.

Oxygen consumption rates, a measure of total respiratory electron flux, of extracts from mutant HW271 were depressed relative to activities of the isogenic control strain HW272 (Fig. 1*a*) and of AN2342 (results not shown) after growth in GMM. However, respiratory rates with various substrates were differentially affected. Thus, oxygen consumption rates by HW271 with lactate or succinate as substrate were about 37% of the rates exhibited by HW272 whereas, with α -glycerophosphate, the respiratory activity of HW271 was 60% of that of HW272. Oxygen uptake rates were highest with lactate

and lowest with α -glycerophosphate in both mutant and wild-type. Significantly, the respiratory capacity of HW271 varied less with the different substrates than did that of HW272. The basis for continued respiration in the absence of quinone is not understood. The residual oxidase activity of the quinone mutants may be attributable to the presence of undetectable amounts of ubiquinone, or to relatively ineffective electron transfer by menaquinone, or perhaps by a ubiquinone biosynthesis precursor. Wallace & Young (1977*a*) have demonstrated that the precursor accumulated by a *ubiG* mutant did not support the oxidation of D-lactate and succinate, but they were unable to assess whether or not this precursor could function in α -glycerophosphate oxidation, because of the low activity of α -glycerophosphate dehydrogenase in these strains when grown on glucose. We do not know whether HW271 carries a leaky mutation that may produce a partially functional polypeptide. We did not attempt to measure respiration rates of intact cells because of the uncertain nature of endogenous substrates. Since cytochromes were still detectable, we tested whether respiration in HW271 was limited by ubiquinone; reconstitution *in vitro* was attempted using exogenous quinone analogues. A solution of ubiquinone-1 (~130 nM final) was added to the

Fig. 1. Oxygen uptake rates of cell-free extracts from *E. coli* HW271 (mutant; ■) and HW272 (isogenic control strain; ▨) (a) before adding ubiquinone and (b) after adding ubiquinone-1. (c) Ratio of the ubiquinone-stimulated rates to the initial rates. G-3-P, α -glycerophosphate.

reaction chamber during measurements of oxygen consumption. The results (Fig. 1b) show that the respiration rates of both mutant and control strains were increased, but the stimulative effect on the mutant was more marked than that on the control. Adding more ubiquinone-1 gave no further enhancement for either strain. The degree of enhancement also differed among the different substrates (Fig. 1c). Menaquinone (Vitamin K₁), ubiquinone-10, and ubiquinone-8 extracted from the wild-type strain had no effect on respiration (results not shown), presumably due to the solubilities of ubiquinone-8 and ubiquinone-10 in water being much lower than that of ubiquinone-1. These results are consistent with early reports (Cox *et al.*, 1970; Wallace & Young, 1977b). Long isoprenoid chain ubiquinone homologues have been used to reconstitute the respiration of pentane-extracted, lyophilized mitochondria. In those experiments, quinones were mixed with phospholipid before carrying out the measurement (Lenaz *et al.*, 1975). The partition of water-insoluble long isoprenoid chain ubiquinone homologues into the membrane phase of highly diluted membrane preparations is expected to be slight.

Fig. 2. (a) Sequencing strategy. Both strands were sequenced by either nested deletions or synthetic oligonucleotide primers. The solid and dotted arrows represent sequences determined with nested deletion clones and synthetic oligonucleotide primers respectively. (b). ORFs determined from the sequence data and overlap with the sequences reported by three other groups: I, reported by Yoshida *et al.* (1988); II, reported by Swanberg & Wang (1987); III, reported by Hussain *et al.* (1987). Numbering of nucleotides in these sequences is that used by these authors. Some of the discrepancies among these reported sequences are discussed in the text. Note that after nucleotide number 1679 in the present sequence, it is no longer similar with that reported by Hussain *et al.* (1987). The clones used for complementation experiments are also shown.

Quinone levels and complementation

Mutant HW271 complemented with pRP19, and the control strain HW272, were grown in defined medium plus glucose (Gibson *et al.*, 1977) for quinone assays. The extracts from these strains each exhibited spots on TLC plates corresponding to ubiquinone-8, the main quinone of aerobically grown *E. coli* (results not shown). Extraction of the material and spectroscopic analysis confirmed the identification. No ubiquinone was detected in HW271. The results suggest that HW271 is deficient in ubiquinone biosynthesis and that plasmid pRP19 restores quinone levels in this mutant. Menaquinone levels appeared to be low in all strains.

Identification of the affected gene in mutant HW271 as *ubiG*

Conventional genetic mapping procedures failed to give satisfactory results. However, preliminary DNA sequence analysis of the insert in pMZ26 (a 2.8 kb *Cla*I–

HindIII
 10 30 50 70 90
 ATCGATAAGCTTCTTCAATACCGCGACGACCGTTAATGATTGCCGCCGTGCGGAAGTCCGGCCCCGGATGTTTCCATCAGCCCTTCAATGCTGATGTC
 TAGCTATTTCGAAGAAGTTATGGCGCTGCTGGCAATTACTAACGGCGCGCAGCCCTTCAGGCCGGGGCCCTACACAAGTAGTCGGGAAGTTACGACTACAG
 110 130 150 170 190
 TTCATCATCAATATACGCCAGACAACCGTTGATGACTTCGGTCAGGTGTTGGGGGGGATGTTGGTTGCCATACCTACGGCGATACCGGAAGAACCGGTC
 AAGTAGTATTATGCGGTCTGTTGGCACTACTGAAGGCAGTCCAACACGCCGCCCTACAACCAACGGTATGGATGCCGCTATGGCCCTCTTGGCAAG
 210 230 250 270 290
 ACCAGCAGGTAGGAATTTTGGTTGGCATGACGTCCCGAATTTTTCCGTGCCGTATAGTTATCAACGAAATCGACCGTCTCTTTTTCGAGATCGGCCA
 TGGTCGTCCAATCCTTAAACCAACCGTACTGCAGGCCCTTAAAAAAGGCACGGCAGTATCAATAGTTGCTTTAGCTGGCAGAAAAAGCTCTAGCCGGT
 310 330 350 370 390
 TCAGTTCACTGGGCAATTTTCGCCAGACGGATTTCCGTATAACGCATTGCCGCCGAGAGTCCCGTCGATAGAACCGAAGTTACCTGACCGTCTACCG
 AGTCAAGTACCGGTTAAAAAGCGGTCTGCCTAAAGGCATATTCGCTAACGGCGCGCTCTCAGCGGCAGCTATCTTGCTTCAATGGAGTGGCAGATGGTC
 410 430 450 470 490
 CATATAACGCAGCGAAGTGGCTGCCCATGCCGACGATCGTGTATAGACCGCCGAGTACCATTGGGATGTTATTTACCGATTACGTACCAACGACA
 GTATATTGGCTGCTCTTACCGACGCGGTACGCTGCTAGCACAATATCTGGCGCTCAGTGGTACCCCTACCATAAATGGCTAATGCAGTGGTTGCTGT
 510 530 550 570 590
 CGGGCAGATTTTTTATAGGCTTTGTTCCAGTCATTGCCATAGTACGTTTATGGCGTAAAGTACGCGACGGTGTACCGGCTTCAGGCCATCTCGGACATCTG
 GCCCGTCTAAAAAATATCCGAAACAAAGTCAAGTAAACGGATCATGCAAGTACCGCATTTTCATGCGCTGCCACATGGCCGAAGTCCGGTAGAGCTGTAGAC
 610 630 650 670 690
 GCAGCGCAGGCCCAACAAATGACCGACATCGCATAATCCAGATAGGAGCTCTTCAGCTCTTCTCAATGTTGACCGGTGTAATTTCTCTCGAAGGTGCGT
 CGTCGCGTGCCGGTGTACTGGCTGTAGCGTATTAGGTCTATCTCGAGAAGTCGAGAAGGAGTTACAACCTGGCCACATTAAGAGAGACGGTTCACGCGA
 710 730 750 770 790
 CATCTAACCGCTATCCCTCTACTGTATCCCGGATTCAAAGGTTCGCAAAATATAACACAGCCGCGCAGTTTGAGGTAAACCTATACGCTTTATTCACATCC
 GTAGATTGGCGATAGGGAGATGACATAGGGCTAAGTTTCCAGCGTTTAATATTGTGTGCGGCGCTCAAACCTCCATTGGATATGCGAAATAAGTGTAGG
 810 830 850 870 890
 AATGCCGTATATACTCGTTTGTCTTGCCAATTACGGAGTGAAGTGCATGAATGCCGAAAAATCGCCGGTAAACCATTAACGTAGACACCAAGAGATC
 TTACGGACTATATAGCAAAACAGAACGGTTAATGCCTCATCTTCACGGTTACTTACGGCTTTTACGGGCCATTGGTATTGCATCTGGTGCTTCTCTAG
 910 930 950 970 990
 GCTAAATTTGAAGCCGTGCGCTCCCGCTGGTGGGATCTGGAAGGTGAGTTCAAACCGCTGCACCGCATTAAACCGCTGCGGTCTGGGCTATATTGCCGAGC
 CGATTTAACTTCGGCAGCGGAGGGCGACCCCTAGACCTTCCACTCAAGTTTGGCGACGTGGCGTAATTGGGCGACGACCCGATATAACGGCTCG
 A K F E A V A S R W W D L E G E F K P L H R I N P L R L G Y I A E R
 1010 1030 1050 1070 1090
 GTGCTGGCGGTTTATTGGCAAAAAGGTGCTCGATGTCGGTTTGGCGCGCGCATTCTGGCCGAGAGTATGGCGCGCGAAGGCGCGCAGGTGACCGGTCT
 CACGACCGCCAAATAAACCGTTTTCACGAGCTACAGCCAACACCGCGCGCTAAGACCGGCTCTCATACCGCGCGCTTCGCGCTGCCACTGGCCAGA
 A G G L F G K K V L D V G C G G G I L A E S M A R E G A T V T G L
 1110 1130 1150 1170 1190
 GGATATGGGCTTTGAGCCATTGCAGGTGGCAAACTGCACGCACTGGAAGCGGCATTCAAGTGGATTACGTGCAGGAAACCGTGGAAAGACACGCGCA
 CCTATACCCGAAACTCGGTAACGTCCACCGTTTTCAGCTGCGTGACCTTTCGCGTAAGTCCACCTAATGCACGTCTCTTGGCACCTTCTCGTGCGCGCT
 D M G F E P L Q V A K L H A L E S G I Q V D Y V Q E T V E E H A A
 1210 1230 1250 1270 1290
 AAACATGCCGGCAGTATGATGTTGACCTGCATGGAGATGCTGGAGCAGTCCCGATCCGCAAGTCAAGTGGTGCAGGCTGTGCGCAACTGGTGAAC
 TTTGTACGGCCGCTCATACTACCACTGGAGCTACCTTACGACCTCGTGAGGGGTAGGCGTCAGTCACAGTCTCGGACACGCGTTGACCACTTTG
 K H A G Q Y D V V T C M E M L E H V P D P Q S V V R A C A Q L V K P
 1310 1330 1350 1370 1390
 CAGCGCGCATGTCTTTTCTCGACACTTAACCGCAACGGCAAGTCATGGCTGATGGCGGTGGTGGTGGCGGAATATATTTTGGCGATGGTCCCAAGG
 GTCCGCCGTACAGAAAAAGAGCTGTGAATTGGCGTTGCCGTTACGATACCGACCTACCGCCACCAACACGCTTATATAAAACGCGTACCAAGGTTTCC
 G G D V F F S T L N R N G K S W L M A V V G A E Y I L R M V P K G
 1410 1430 1450 1470 1490
 CACGCATGATGTGAAGAAGTTTATTAACCGGCAAGTGTGGGTGGGTGGATCAGACCAAGTTTAAAGAGCGGCATATCACTGGGTGCATTACAAC
 GTGCGTACTACACTTCTTCAATAATTTGGCCGTCTTAACGACCCGACCACTAGTCTGGTCAAAATTTCTCGCGTATAGTACCCGACGTAATGTTG
 T H D V K K F I K P A E L L G W V D Q T S L K E R H I T G L H Y N

-35
 -10
 RBS
 M N A E K S P V N H N V D H E E I
 ubiG

1510 1530 1550 1570 1590
 CCGATCACTAACTTTTAACTCGGCCCCGGCGTGGATGTGAATATATGCTGCACACGCAGAAATAGTGAGGTGATGTTTGGCCGGCCCAATGCCTG
 GGCTAGTGATTATGAAAAATTTAGCCGGGGCCGCACCTACACTTGATATACGACGTGTGCGTCTTATTCACTCCAACACCAACGGCGCGGTTACGGAC
 P I T N T F K L G P G V D V N Y M L H T Q N K *
 1610 1630 1650 1670 1690
 ATGCGACGCTTGCCGCGTCTTATCAGGCTTACAAATGCTCCCGTAGGCGGATAAGGCGTTACGCCGATCCGGCAATCAATGCCTGATCGGACGCTG
 TACGCTGCGAACGGCGCAGAATAGTCCGGATGTTTACGAGGGGATCCGGCTATTCCGCAATGCCGCTAGGCGGTTAGTACGGACTACGCTGCGAC
 1710 1730 1750 1770 1790
 TCGCGTCTTATCAGGCTTACAAACCGTTACCGACTCGCATCGGGCAATTTGTGGTTACGATGGCATCGCGATAGCCTGCTTCTCTCAAGCAGCTTCTCG
 AGCGCAGAAATAGTCCGGATGTTTGGCAATGGCTGAGCGTAGCCGTTAACACCAATGCTACCGTAGCGCTATCGGACGAAGAGAAGTTCTGTCGAAGAGC
 * S P M A I A Q K E E L K E
 1810 1830 1850 1870 1890
 ACTACGCCAGGATCGGCAAGCGTCGAGGTATCGCCAGGTTGCTGGTATCGCCCGCCGAATTTTGGCGAGAATACGGCGCATAATTTTGGCGGAGCGGG
 TGATGCGGTCTAGCCGTTCCGAGCTCCATAGCGGTCCAACGACCATAGCGGGCGGCGTTAAACCGCTCTTATGCCGCTATTAAACAGGCTCGCCC
 V V G P D A L T S T D G L N S T D G A A I K R L I R R M I K G S R T
 1910 1930 1950 1970 1990
 TTTTAGGCAAGGAGTCGGTCCAGTGCAGCAGCTGTGGCTGCGCAGCGGGCAATCTCTTACGCACCCAGTTGCGGACTTCTGCGTACAGTTCTGGTGA
 AAAATCGTCCCTCAGCCAGGTACGTCGTGACGACGCGAGCGGTGCGCCGGTTAGAGAAATGCGTGGGTCAACGCTGAAGACGCATGCAAGACCACT
 K P L S D T W H L V D P T A L P G I E K R V W N R V E A Y L E P S
 2010 2030 2050 2070 2090
 CGGTTCTCTCCCGTGATTAAGCGTGACGTAGGCGTAGATCGCCTGACCTTTAATATTGTGCGGAATACCTACTACGGCGGCTTCGGCAATCTTCGGATGC
 GCCAAGGAGGGGCACTAATTCGCACTGCATCCGCATCTAGCGGACTGGAATTTATAACACGCTTATGGATGATGCCGCCAAGCCGTTAGAAGCCTACG
 P E E G H N L T V Y A Y I A Q G K I N H P I G V V A A E A I K P H
 2110 2130 2150 2170 2190
 GCCACCAGCGCGGACTCAATCTCTGCCGTCCCGAGACGGTGACCGGAGACGTTTACGACGCTCGTCCACACGCCCGGTTATCCAGTAATAGCCATCTTCAT
 CGGTGGTGCGGGCTGAGTTAGAGACGGCAGGGGTCTGCCACTGGCCTCTGCAAGTCGTGCAGCAGGTGTGCGGGCAATAGGTGATTATCGGTAGAAGTA
 A V L A S E I E A T G L R H G S V N L V D D V R G T I W Y Y G D E D
 2210 2230 2250 2270 2290
 CGCGACGCGCGCGCTCGCGCTGAAATACATATTTTGAAGGTGGAGAAGTAGGTCTGTTCAAAACGTTCTGATCGCCAAACAGCGTACGCGCTGACC
 GCGCTGCGCGCGGACGCGGCTTTATGTATAAAACTTCCACCTCTTCATCCAGCAAGTTTTCGAAGCACTAGCGGTTTGTGCGATGCGCGGACTGG
 R R A G D G S F Y M N K F T S F Y T Q E F R E H D G F L T R A Q G
 2310 2330 2350 2370 2390
 CGGCCAGGAGTCGGTGATTACCAAGGCTACCTTCGGTGGCCCCCTCCAGCGGGTTACCTTCGTTATCGACAGCGCGGTTGACGCCGAAGAACGGACGT
 CGCGTCTCTCAGCCACTAATGGTCCGATGGAAGCCACCGGGGAGGTGCGCCAAATGGAAGCAATAGCTGGTGCAGGCAACGTGCGGCTTCTGCCTGCA
 P W S D T I V L S G E T A G E L P N G E N D V L A P Q V G F F P R
 BglI
 2410 2430 2450 2470 2490
 GTTGCGCAACCGGCTTTTTCAGCTCGGTAGCGCCAGGACGGGTTGATCATGAAACCGCGGTTTCGGTCTGCCACCAGGTATCGACACCGGACATTTCT
 CAACGGCTTGCCGAAAGTCAGGCTCGCGTCCGTCGCCCACTAGTACTTTGGCGCCAAAGCCAGACGGTGGTCCATAGCTGGTGGCTGTAAAGA
 T A S G A K L E T A G P L P T I M F G G T E T Q W W T D V V P C K E
 2510 2530 2550 2570 2590
 CGTTGCGGATTTTTCAGTACCACTCCACGCTTCCGGTTAATTGGCTCGCCACGGAACCGAGAATGCGCAGCGACGAACGGTCGGTCCGCTTCGAT
 GCAACGGCTAAAAAAGGTGATGGTGAGGTCGGAAGGCCAATTAACCGAGCGGTTGCTTGGCTCTTACGCGTGGTGTGCGAGCCACGGAAGCTA
 N G I K K W Y W E W A E P N I P E G V S G L I R L S S R D T G E I
 2610 2630 2650 2670 2690
 CGCTTTATCGCCTTCGCGCATCAGCGCGGATCGCGTGGGTGCGGTATAGAGAATATTGACCTGATGCTTGTCCACCACCTGCGCCATACGGGACGGC
 GCGAAATAGCGGAAGCGGTAGTCGCGCGCTAGCGGACCCACGCCATATCTTTATAACTGGACTACGAACAGGTGGTGGACGCGGTATGCCGCTCCG
 A K D G E A M L A R I A T P A T Y L I N V Q H K D V V Q A M R A P
 2710 2730 2750 2770 2790
 GTCGGCCAGTTGGGTACGCCCTTCAACATCAGCGTGGTCGCACCGCAGCCAGCGGCGGTACAGCAAGTAACTGTGTCGGTCAACCGAGCCACATCGGC
 CAGCCGGTCAACCATGCGGAGTTTGTAGTCGACACGCGTGGGCGTGGTGGCGGCGATGTCGTTTATGACACAGGCGAGTGGGTCGGGTGATGCCG
 T P W N P V G E F M L T T A G C G A P R V A L L Q T R D G L G C R R
 2810
 GGTGCACAGTAGATATC
 CCACGTGGTCATCTATAG
 H V L L Y
 EcoRV
 orfX ←

Fig. 3. Complete nucleotide sequence and its encoded protein sequence. The sequence from 478 to 696 was taken from Hussain *et al.* (1987).

EcoRV fragment from pRP19) showed that the *ClaI* end was identical to part of the *gyrA* gene, which maps at 48 min on the *E. coli* linkage map adjacent to *ubiG* (Bachmann, 1990). Confirmation that the gene cloned in pRP19 and pMZ26 is *ubiG* was obtained by transforming the *ubiG* mutant AN151 (Leppick *et al.*, 1976) with these plasmids and finding that it was complemented for aerobic growth on SMM. To determine the physical map location of *ubiG*, HW271 was coinfectd with λ CI857 and one of seven phage clones, 370–376, from the ordered Kohara miniset library (Kohara, 1990). Only clone 376 (E13A5) gave rise to colonies on SMM medium. This indicates that *ubiG* is located in the 7 kb region of this clone which does not overlap with clone 375, that is at approximately 2348–2358 kb on the physical map and close to 48 min on the linkage map.

Sequence of the 2.8 kb insert in pMZ26

The insert from pMZ26 was cloned into the DNA sequencing vector pBluescript II, and designated pRPB. Fig. 2(a) shows the strategy for sequencing this 2.818 kb *ClaI*–*EcoRV* fragment. Nucleotide sequences were obtained using both strands as templates and were in agreement. Fig. 2(b) shows the extent to which the complete nucleotide sequence of the *ClaI*–*EcoRV* fragment (Fig. 3) includes regions studied by Hussain *et al.* (1987), Swanberg & Wang (1987) and Yoshida *et al.* (1988). These workers did not identify the ORF that we assign to *ubiG* (see below). The *ClaI* end of the sequence (nucleotide numbers 6–478) is identical with the reported *gyrA* gene (Swanberg & Wang, 1987; Yoshida *et al.*, 1988), but identity starts from a *HindIII* site (nucleotide 6) adjacent to the *ClaI* site. The *ClaI* site is not part of the reported *gyrA* sequence and may be a cloning artefact. The sequences reported by Hussain *et al.* (1987) and Swanberg & Wang (1987) also disagree with each other at a number of sites; for example, at nucleotide number 1029 and 1045, both Gs are reported as A by Swanberg & Wang (1987), and the A at nucleotide 1087 is missing in their sequence. Our sequence (Fig. 3) agrees with that of Hussain *et al.* (1987).

Several potential long ORFs, shown in Fig. 2(b), were found by searching for initiation and termination codons in the sequence shown in Fig. 3. To identify the complementing region of the insert, the *ubiG* mutants AN151 and HW271 were transformed with plasmids deleted for several regions (see Fig. 2b); transformants were selected anaerobically on complex medium plus ampicillin and screened aerobically on LB plates and on SMM. Transformation of mutants AN151 and HW271 with pEI110 restored aerobic growth on LB to levels exhibited by the wild-type strains and by the same mutants transformed with pMZ26. This result suggests

Fig. 4. The palindromic units at the end of the *ubiG* gene. For explanation of box, see text.

that the long ORF extending almost to nucleotide 1679 in Fig. 2(b), and labelled there as *ubiG*, is the gene that complements these mutants. Plasmid pEC10 is deleted such that the protein product lacks the last four amino acid residues of the presumptive UbiG protein; it gave partial complementation of aerobic growth of these mutants on LB. The other two deletion clones, pND31 and pND42, did not complement either mutant on LB.

There is a potential-ribosome binding site GGAG (Shine & Dalgarno, 1974) beginning at nucleotide number 835, 11 bp upstream of the ATG codon at nucleotide number 850 (Fig. 3). However, unequivocal determination of the translation initiation site must await determination of the N-terminal sequence of the protein. The potential –35 and –10 regions were also determined, but the actual transcription initiation site must be determined by primer extension experiments or S1 mapping.

After the termination codon, three PU (palindromic unit) sequences are found (Fig. 4). There is no typical *rho*-independent terminator. However, between the second and third PU, there is a sequence (boxed in Fig. 4) which matches exactly the one found at the 3' terminus of some *rho*-dependent terminators (CAATCAA; Lau *et al.*, 1984). This sequence does not suggest a *rho*-dependent terminator, because the distance between the termination codon and the first PU sequence is only 19 bp (Fig. 3). The *rho*-dependent termination needs a sequence of at least 70–80 bp without any secondary structure (Bear & Peabody, 1988). Some PU sequences can terminate transcription but most do not (Gilson *et al.*, 1987). It is not yet clear whether they are terminators in this case, but two factor-independent terminators were identified in this region at nucleotide numbers 1641 and 1752 using a computer-based search for terminators (Brendel & Trifonov, 1984).

Sequence comparison of the UbiG protein with other proteins

The nucleotide-derived amino acid sequence (Fig. 3) was used to search the OWL database by using the Prosrch

(a) UbiG 59	K V L D V G C G G G I L A E S M A R E G
	Δ Δ Δ O O O Δ Δ Δ
	β β β β β T T T T α α α α α α α
(b) DHHBMT 129	S V L D V G C G G G I L S E S L A R L K
	Δ Δ Δ O O O Δ Δ Δ
	β β β β β T T T T T β β α α α α α
(c) PIMTSBOVIN 80	K A L D V G S G S G I L T A C F A R M V
	Δ Δ Δ O O O Δ Δ Δ
	α α β β β C C C T T T β β β β β β β
(d) ERMA SCORDI 35	P I I E I G P G S G A L T H P M A H L G
	Δ Δ Δ O O O Δ Δ Δ
	β β β β β C C C T ? C β β β α α α α
(e) MLSB SSTRSA 32	T V Y E I G T G K G H L T T K L A K I S
	Δ Δ Δ O O O Δ Δ Δ
	β β β β β C C C T T C β β α α α α α

Fig. 5. Alignments of sequences around the putative AdoMet binding site. Predicted secondary structures are also indicated; β , β sheet, α , α helix, T, turn; C, coil. Δ indicates a conserved hydrophobic amino acid; Λ indicates a conserved charged or polar amino acid. $\bigcirc$, invariant glycine. Proteins were chosen from a multiple sequence alignment (not shown) to represent different groups of methyltransferase. (a) UbiG (this work); (b) 3,4-dihydroxy-5-hexaprenylbenzoate (DHHB) methyltransferase (product of yeast *COQ3* gene); (c) protein- β -aspartate methyltransferase, bovine; (d) rRNA adenine- N^6 methyltransferase, *Corynebacterium diphtheriae*; (e) rRNA adenine- N^6 methyltransferase, *Streptococcus sanguis*. Sequences (c) and (d) were obtained from Swissprot, sequence (b) is from Clarke *et al.* (1991). The numbers to the left of the sequences indicate amino acid residue numbers in each protein. The three conserved Gs in the AdoMet binding sequence are boxed.

program through the DAP (Distributed Array Processor; Collins & Coulson, 1988). No protein that showed significant overall similarity was detected, but a number of proteins exhibited some local similarities with the UbiG protein. Thus, the segment of UbiG from amino acid residues 59 to 77 is similar to a sequence found in some methyltransferases and other enzymes that use *S*-adenosylmethionine (AdoMet) as substrate. Several sequences that are apparently conserved in RNA and DNA methylases have already been noted (e.g. Lauster *et al.*, 1989; Ingrosso *et al.*, 1989). One of these (Region I of Ingrosso *et al.*, 1989) corresponds precisely to the sequence VLE/DXGXGXG found in EryG (a methyltransferase that catalyses the last step in the biosynthesis of erythromycin in *Saccharopolyspora erythraea*) and ErmE (an rRNA methyltransferase in the same organism) (Haydock *et al.*, 1991). These authors scrutinized a number of recently sequenced AdoMet-dependent methyltransferases and concluded that this glycine-rich motif is a common element in enzymes of this class.

Based on the above database search results, we used the 'profilesearch' program (within the 'UWGCG' package, version 7.0) to search the SwissProt database

(release 19.0) and PIR database (release 29.0). A motif was identified similar to, but more general than, that described above, which occurs in some, but not all, methyltransferases. The sequence is $\Delta\Delta D/EAGXGXGX\Delta\Delta\Delta\Delta\Delta$, where Δ is a hydrophobic amino acid, and Λ is a charged, or occasionally a polar, amino acid. The first G is replaced by A in glycine methyltransferase, a conservative replacement. Examples are given in Fig. 5. Secondary structure prediction (Chou & Fasman, 1978) indicates that the region preceding the GXGXG sequence tends to adopt a β sheet structure. The GXGXG region itself is (not surprisingly) strongly predicted as a turn. Predictions for the regions following this sequence are contradictory, in spite of some sequence similarities. The whole motif defined above may, however, be indicative of a common tertiary structural motif. Structural differences between these proteins are expected of course, since although they use the same methyl group donor (AdoMet), the methyl group acceptors are different. Outside the region of this motif there is little or no evidence of tertiary structure similarity on the basis of sequence comparisons. The methyltransferases that share this motif include rRNA adenine- N^6 methyltransferases, protein- β -aspartate methyltransferases, protein- γ -glutamate *O*-methyltransferases, L-isoaspartyl/D-aspartyl protein methyltransferases and some DNA methyltransferases, although most DNA methyltransferases do not share this motif.

Another protein encoded by the *ClaI*-*EcoRV* fragment

The protein encoded by the long 'orfX', part of which extends from the *EcoRV* end of the insert in pMZ26, was also sought in the OWL database. The 290 amino acids of the C-terminus of this protein are very similar to acetyl-CoA synthetase from *Aspergillus nidulans* and *Neurospora crassa* (Connerton *et al.*, 1990) with 44% and 42% identity respectively. Examination of the 48 min region of the *E. coli* linkage map (Bachmann, 1990) suggests that this protein might belong to the *ato* operon, which is involved in short-chain fatty acid degradation (Jenkins & Nunn, 1987). The gene for acetyl-CoA synthetase in *E. coli* is reported to map at 40 min (Bachmann, 1990).

Discussion

E. coli, together with other Gram-negative facultative bacteria, contains isoprenoid quinones of both the benzene and naphthalene series. The major benzoquinone is ubiquinone-8 (with the figure referring to the number of prenyl units present in the side chain), whilst the naphthoquinones are either menaquinone or demethylmenaquinone. The biosynthetic pathways of ubi-

Fig. 6. The 48 min region of the *E. coli* physical map (Kohara *et al.*, 1987). The *Pst*I fragment is probably that which Gibert *et al.* (1988) used to construct a *ubiG-lacZ* fusion. The dotted arrow shows the direction of transcription proposed by Gibert *et al.* (1988), and the shaded region is the sequence that differs from that reported by Hussain *et al.* (1987).

quinone and menaquinones have been studied in detail (for reviews see Gibson, 1973; Bentley & Meganathan, 1987). Both quinones comprise a nucleus derived from chorismate, a prenyl side chain derived from prenyl pyrophosphate, and a nuclear methyl group derived from AdoMet (except in the case of demethylmenaquinone). The final step in the biosynthesis of ubiquinone is the methylation of 2-octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinone, a reaction catalysed by the product of the *ubiG* gene (Leppick *et al.*, 1976). The gene described in this paper is identified as *ubiG* by complementation experiments and by the map position, adjacent to *gyrA*, provided by the sequence data.

Previous attempts to isolate ubiquinone-deficient mutants have exploited the fact that such mutants are unable to grow on succinate, have low oxidase activities, and grow slowly on glucose aerobically, yet normally on glucose anaerobically (Wallace & Young, 1977*a, b*; Cox & Downie, 1979). In contrast to mutants defective in cytochrome *d* (*cyd*) or *o* (*cyo*) terminal oxidases, quinone mutants should also fail to grow anaerobically on certain alternative electron acceptors. Recently, Wissenbach *et al.* (1990) have shown that menaquinone is specifically involved in reduction of fumarate, that menaquinone or demethylmenaquinone can serve as mediators in reduction of trimethylamine-*N*-oxide, but that nitrate respiration can utilize either menaquinone or ubiquinone. This result is consistent with the phenotype of the *ubiG* mutant described here, which grows anaerobically on either fumarate or nitrate.

The enzyme encoded by *ubiG* has been purified by gel filtration (Leppik *et al.*, 1976), and shown to have a native molecular mass of about 50 kDa. The subunit composition has not been reported, but the deduced amino acid sequence of UbiG indicates a molecular mass of about 26 kDa, suggesting that it may represent one subunit of a homodimer.

The region that encodes the UbiG protein was sequenced by Hussain *et al.* (1987) as part of their study

of the *gyrA* gene, but their sequence data were not available in a databank until recently. They identified an ORF called *pufX*, (protein of unknown function) upstream of the *gyrA* sequence, which we have now identified as *ubiG*. The region from nucleotide number 6 to nucleotide number 478 and from nucleotide number 696 to 1679 of our sequence is identical to their sequence (Fig. 2*b*). We did not sequence the region between nucleotide number 478 and nucleotide number 696. However, after nucleotide number 1679, the two sequences are no longer identical. No known restriction site has been located in this region, so it is unlikely to be an artificial rearrangement. Mapping the genes around *gyrA* has always been problematic. Bachmann (1987) suggested that perhaps one (or more) of the strains which had been used in several laboratories in the mapping of this region carry an inversion of the chromosome between *udk* and *gyrA*. It is interesting to note that the sequence of Hussain *et al.* (1987) has only two PU sequences but ours has three. The PU sequences are highly conserved; this suggests that they may be sites where recombination of the chromosome between homologous region has occurred (see Gilson *et al.*, 1987).

A database search has revealed a short stretch of the UbiG polypeptide which is highly similar to a sequence that occurs in some, but not all, methyltransferases. Within this family of methyltransferases, the sequence motif $\Delta\Delta D/E\Delta G X G X G X \Delta X X X \Delta \Delta \Delta$ is highly conserved, while other parts of the proteins show very little resemblance to each other. This common motif may reflect a common tertiary structural element which corresponds to the common nature of the reactions catalysed, and of the methyl group donor. This region corresponds to Region I identified by Ingrosso *et al.* (1989), but we are unable to locate in UbiG sequences that correspond to regions II and III in that paper. Since regions II and III are present in only a few methyltransferases, and are not conserved in recently sequenced AdoMet-dependent methyltransferases (see Haydock *et*

al., 1991), their general significance must be questioned. The AdoMet binding sequence described in this paper is not shared by most DNA methylases (Reich & Everett, 1990; Wenzel *et al.*, 1991), in which an alternative motif has been proposed.

Recently, Clarke *et al.* (1991) sequenced the *COQ3* gene coding for the 3,4-dihydroxy-5-hexaprenylbenzoate methyltransferase involved in ubiquinone biosynthesis in *Saccharomyces cerevisiae*. The predicted amino acid sequence of the gene product showed 40% identity with the predicted protein product of the open reading frame on the *E. coli* genome identified as *ubiG* in the present work. The *COQ3* gene product codes for one of the early reactions in ubiquinone biosynthesis in the yeast whereas *ubiG* codes for the terminal methylation in *E. coli*. It seems likely, therefore, that the other transmethylation enzymes in the ubiquinone biosynthetic pathways in *E. coli* and *S. cerevisiae* will also show sequence homology.

Gibert *et al.* (1988) used a *Bam*HI site on a cloned *Pst*I–*Pst*I fragment to make a *ubiG*–*lacZ* fusion. This fragment is known to contain the *nrdAB* gene and part of *gyrA*: *ubiG* is between them (Platz & Sjöberg, 1980). These genes are known to be located at the 48 min region of the *E. coli* chromosome, and the restriction sites of this fragment match the *E. coli* physical map of Kohara *et al.* (1987, Fig. 6). Presumably, the *Bam*HI site in the middle of this fragment is that which Gibert *et al.* (1988) used to construct the *ubiG*–*lacZ* fusion. They concluded that *ubiG* is transcribed towards the *gyrA* gene, and subject to catabolite repression. Our sequence analysis results show an opposite transcriptional direction, as indicated in Fig. 6. No Cap binding site (Ebright *et al.*, 1987), as reported by Gibert *et al.* (1988), has been identified and no *Bam*HI site exists in the fragment on which we are working. The *Bgl*II site on the insert of pMZ26 is close to the *Eco*RV end (Fig. 3). The results suggest a chromosomal rearrangement of our fragment on the left-hand side of the *Bgl*II site which is close to *ubiG* (Fig. 6). One possible explanation is that the *ubiG* mutant (AN86) which Gibert *et al.* (1988) used, and AN151 used in the present study, each carry a mutation in one of two closely linked genes, the product of both being required for synthesis of 2-octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinone methyltransferase. In this case the UbiG protein might consist of two different subunits. However, recent complementation experiments with AN86 (G. Wu, unpublished results) using plasmid pMZ26 described in this paper gave results identical with AN151 described here.

We thank G. B. Cox for stimulating discussions and encouragement, I. G. Young for gifts of strains, and B. Sutton and P. Cunningham for advice on sequence analyses. This work was supported by grants to R.K.P. from SERC, the Royal Society and a Visiting Fellowship at ANU. G.W. acknowledges with thanks the award of a K.C. Wong Scholarship.

References

- AU, D. C.-T., LORENCE, R. M. & GENNIS, R. B. (1985). Isolation and characterization of an *Escherichia coli* mutant lacking the cytochrome *o* terminal oxidase. *Journal of Bacteriology* **161**, 123–127.
- BACHMANN, B. J. (1987). Linkage map of *Escherichia coli* K-12, edition 7. In *Escherichia coli and Salmonella typhimurium, Cellular and Molecular Biology*, pp. 807–876. Edited by F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter & H. E. Umbarger. Washington, DC: American Society for Microbiology.
- BACHMANN, B. J. (1990). Linkage map of *Escherichia coli* K-12, Edition 8. *Microbiological Reviews* **54**, 130–197.
- BEAR, D. G. & PEABODY, D. S. (1988). The *E. coli* Rho protein: an ATPase that terminates transcription. *Trends in Biochemical Sciences* **13**, 343–347.
- BENTLEY R. & MEGANATHAN, R. (1987). Biosynthesis of the isoprenoid quinones ubiquinone and menaquinone. In *Escherichia coli and Salmonella typhimurium, Cellular and Molecular Biology*, pp. 512–520. Edited by F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter & H. E. Umbarger. Washington, DC: American Society for Microbiology.
- BIRNBOIM, H. C. (1983). A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods in Enzymology* **100**, 243–255.
- BOLIVAR, F., RODRIGUEZ, R. L., GREENE, P. J., BETLACH, M. C., HEYNECKER, H. L., BOYER, H. W., CROSA, J. H. & FALKOW, S. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**, 95–113.
- BRENDEL, V. & TRIFONOV, E. N. (1984). A computer algorithm for testing potential prokaryotic terminators. *Nucleic Acids Research* **12**, 4411–4427.
- CHOU, P. Y. & FASMAN, G. D. (1978). Empirical predictions of protein conformation. *Annual Review of Biochemistry* **47**, 251–276.
- CLARKE, C. F., WILLIAMS, W. & TERUYA, J. H. (1991). Ubiquinone biosynthesis in *Saccharomyces cerevisiae*. Isolation and sequence of *COQ3*, the 3,4-dihydroxy-5-hexaprenylbenzoate methyltransferase gene. *Journal of Biological Chemistry* **266**, 16636–16644.
- COLLINS, J. F. & COULSON, A. F. W. (1988). Significance of protein similarities. *Methods in Enzymology* **183**, 474–487.
- COLLINS, M. D. (1985). Analysis of isoprenoid quinones. *Methods in Microbiology* **18**, 329–366.
- CONNERTON, I. F., FINCHAM J. R. S., SANDEMAN, R. A. & HYNES, M. J. (1990). Comparison and cross-species expression of the acetyl-coA synthetase gene of the ascomycete fungi, *Aspergillus nidulans* and *Neurospora crassa*. *Molecular Microbiology* **4**, 451–460.
- COX, G. B. & DOWNIE, J. A. (1979). Isolation and characterization of mutants of *Escherichia coli* K-12 affected in oxidative phosphorylation or quinone biosynthesis. *Methods in Enzymology* **56**, 106–117.
- COX, G. B., HATCH, L., WEBB, D., FIMMEL, A. L., LIN, Z.-H., SENIOR, A. E. & GIBSON, F. (1987). Amino acid substitutions in the ϵ -subunit of the F_1F_0 -ATPase of *Escherichia coli*. *Biochimica et Biophysica Acta* **890**, 195–204.
- COX, G. B., NEWTON, N. A., GIBSON, F., SNOSWELL, A. M. & HAMILTON, J. A. (1970). The function of ubiquinone in *Escherichia coli*. *Biochemical Journal* **117**, 551–562.
- EBRIGHT, R. H., KOLB, A., BUC, H., KUNKEL, T. A., KRAKOW, J. S. & BECKWITH, A. (1987). Role of glutamic acid-181 in DNA-sequence recognition by the catabolite gene activator protein (CAP) of *Escherichia coli*: altered DNA-sequence-recognition properties of [Val¹⁸¹]CAP and [Leu¹⁸¹]CAP. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 6083–6087.
- ELLIOTT, T. & ROTH J. R. (1988). Characterization of Tn10dCam: a transposition defective Tn10 specifying chloramphenicol resistance. *Molecular and General Genetics* **213**, 332–338.
- GIBERT, I., LLAGOSTERA, M. & BARBE, J. (1988). Regulation of *ubiG* gene expression in *Escherichia coli*. *Journal of Bacteriology* **170**, 1346–1349.
- GIBSON, F. (1973). Chemical and genetic studies on the biosynthesis of ubiquinone by *Escherichia coli*. *Biochemical Society Transactions* **1**, 317–326.
- GIBSON, F., COX, G. B., DOWNIE, J. A. & RADIK, J. (1977). A mutation affecting a second component of the F_0 portion of the magnesium

- ion-stimulated adenosine triphosphatase of *Escherichia coli* K12. The *uncC424* allele. *Biochemical Journal* **164**, 193–198.
- GILSON, E., CLEMENT, J. M., PERRIN, D. & HOFNUNG, M. (1987). Palindromic units: a case of highly repetitive DNA sequences in bacteria. *Trends in Genetics* **3**, 226–230.
- HAYDOCK, S. F., DOWSON, J. A., DHILLON, N., ROBERTS, G. A., CORTES, J. & LEADLEY, P. F. (1991). Cloning and sequence analysis of genes involved in erythromycin biosynthesis in *Saccharopolyspora erythraea*: sequence similarities between EryG and a family of S-adenosylmethionine-dependent methyltransferases. *Molecular and General Genetics* **230**, 120–128.
- HOHN, B. & COLLINS, J. (1980). A small cosmid for efficient cloning of large DNA fragments. *Gene* **11**, 291–298.
- HUSSAIN, K., ELLIOTT, E. J. & SALMOND, G. P. C. (1987). The *parD*⁻ mutant of *Escherichia coli* also carries a *gyrA*_{am} mutation. The complete sequence of *gyrA*. *Molecular Microbiology* **1**, 259–273.
- INGROSSO, D., FOWLER, A. V., BLEIBAUM, J. & CLARKE, S. (1989). Sequence of the D-aspartyl/L-isoaspartyl protein methyltransferase from human erythrocytes. Common sequence motifs for protein, DNA, RNA, and small molecule S-adenosylmethionine-dependent methyltransferases. *Journal of Biological Chemistry* **264**, 20131–20139.
- JENKINS, L. S. & NUNN, W. D. (1987). Genetic and molecular characterization of the genes involved in short-chain fatty acid degradation in *Escherichia coli*: the *ato* system. *Journal of Bacteriology* **169**, 42–52.
- KITA, K., KONISHI, K. & ANRAKU, Y. (1984). Terminal oxidases of *Escherichia coli* aerobic respiratory chain. I. Purification and properties of cytochrome *b*_{562-o} complex from cells in the early exponential phase of aerobic growth. *Journal of Biological Chemistry* **259**, 3368–3374.
- KOHARA, Y. (1990). Correlation between the physical and genetic maps of the *Escherichia coli* K-12 chromosome. In *The Bacterial Chromosome*, pp. 29–42. Edited by K. Drlica & M. Riley. Washington, DC: American Society for Microbiology.
- KOHARA, Y., AKIYAMA, K. & ISONO, K. (1987). The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell*, **50**, 495–508.
- KROPFENSTEDT, R. M. (1982). Anwendung chromatographischer HP-Verfahren (HPTLC und HPLC) in der Bakterien-Taxonomie. *GIT Labor-Medizin* **5**, 266–275.
- LAU, L. F., ROBERTS, J. W., WU, R., GEORGES, F. & NARANG, S. A. (1984). A potential stem loop structure and the sequence CAAUCAA in the transcript are insufficient to signal ρ -dependent transcription termination at λ IR1. *Nucleic Acids Research* **12**, 1287–1299.
- LAUSTER, R., TRAUTNER, T. A. & NOYER-WEIDNER, M. (1989). Cytosine-specific type II DNA methyltransferases. A conserved enzyme core with variable target-recognizing domains. *Journal of Molecular Biology* **206**, 305–312.
- LENAZ, G., PASQUALI, P., BERTOLI, E., PARENTI-CASTELLI, G. & FOLKERS, K. (1975). The inhibition of NADH oxidase by the lower homologs of coenzyme Q. *Archives of Biochemistry and Biophysics* **169**, 217–226.
- LEPPICK, R. A., STROOBANT, P., SHINEBERG, B., YOUNG, I. G. & GIBSON, F. (1976). Membrane-associated reactions in ubiquinone biosynthesis. 2-Octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinone methyltransferase. *Biochimica et Biophysica Acta* **428**, 146–156.
- LORENCE, R. M., KOLAND, J. G. & GENNIS, R. B. (1986). Coulometric and spectroscopic analysis of the purified cytochrome *d* complex of *Escherichia coli*: evidence for the identification of 'cytochrome *a*₁' as cytochrome *b*₅₉₅. *Biochemistry* **25**, 2314–2321.
- MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning – A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- MARKWELL, M. A. K., HAAS, S. M., BIEBER, L. L. & TOLBERT, N. E. (1978). A modification of the Lowry procedure to simplify protein determination in membrane and lipoprotein samples. *Analytical Biochemistry* **87**, 206–210.
- MILLER, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- MURPHY, G. & WARD, E. S. (1989). Sequencing of double-stranded DNA. In *Nucleic Acids Sequencing. A Practical Approach*, pp. 99–115. Edited by C. J. Howe & E. S. Ward. Oxford: IRL Press.
- PLATZ, A. & SJOBERG, B.-M. (1980). Construction and characterization of hybrid plasmids containing the *Escherichia coli* *nrd* region. *Journal of Bacteriology* **143**, 561–568.
- POOLE, R. K. (1977). The influence of growth substrate and capacity for oxidative phosphorylation on respiratory oscillations in synchronous cultures of *Escherichia coli* K12. *Journal of General Microbiology* **99**, 369–377.
- POOLE, R. K. & INGLEDEW, W. J. (1987). Pathway of electrons to oxygen. In *Escherichia coli and Salmonella typhimurium, Cellular and Molecular Biology*, pp. 170–200. Edited by F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter & H. E. Umbarger. Washington, DC: American Society for Microbiology.
- POOLE, R. K., BAINES, B. S. & APPLEBY, C. A. (1986). Haemoprotein b-590 (*Escherichia coli*), a reducible catalase and peroxidase: evidence for its close relationship to hydroperoxidase I and a 'cytochrome *a*₁b' preparation. *Journal of General Microbiology* **132**, 1525–1539.
- POOLE, R. K., WILLIAMS, H. D., DOWNIE, J. A. & GIBSON, F. (1989). Mutations affecting the cytochrome *d*-containing oxidase complex of *Escherichia coli* K12: identification and mapping of a fourth locus, *cydD*. *Journal of General Microbiology* **135**, 1865–1874.
- REICH, N. O. & EVERETT, E. A. (1990). Identification of peptides involved in S-adenosylmethionine binding in the *EcoRI* DNA methylase. Photoaffinity labelling with 8-azido-S-adenosylmethionine. *Journal of Biological Chemistry* **265**, 8929–8934.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.
- SHINE, J. & DALGARNO, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America* **71**, 1342–1346.
- SILHAVY, T. J., BERMAN, M. L. & ENQUIST, L. W. (1984). *Experiments with Gene Fusions*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- STROOBANT, P., YOUNG, I. G. & GIBSON, F. (1972). Mutants of *Escherichia coli* K-12 blocked in the final reaction of ubiquinone biosynthesis: characterization and genetic analysis. *Journal of Bacteriology* **109**, 134–139.
- SWANBERG, S. L. & WANG, J. C. (1987). Cloning and sequencing of the *Escherichia coli* *gyrA* gene coding for the A subunit of DNA gyrase. *Journal of Molecular Biology* **197**, 729–736.
- WALLACE, B. J. & YOUNG, I. G. (1977a). Aerobic respiration mutants of *Escherichia coli* accumulating quinone analogues of ubiquinone. *Biochimica et Biophysica Acta* **461**, 75–83.
- WALLACE, B. J. & YOUNG, I. G. (1977b). Role of quinones in electron transport to oxygen and nitrate in *Escherichia coli*. Studies with a *ubiA*⁻ *menA*⁻ double quinone mutant. *Biochimica et Biophysica Acta* **461**, 84–100.
- WENZEL, C., MOULARD, M., LOBNER-OLESEN, A. & GUSCHLBAUER, W. (1991). Crosslinking of Dam methyltransferase with S-adenosylmethionine. *FEBS Letters* **280**, 147–151.
- WILLIAMS, H. D. & POOLE, R. K. (1987). The cytochromes of *Acetobacter pasteurianus* NCIB 6428. Evidence of a role for a cytochrome *a*₁-like haemoprotein in electron transfer to cytochrome oxidase *d*. *Journal of General Microbiology* **133**, 2461–2472.
- WISSENBAACH, U., KROGER, A. & UNDEN, G. (1990). The specific functions of menaquinone and demethylmenaquinone in anaerobic respiration with fumarate, dimethylsulfoxide, trimethylamine N-oxide and nitrate by *Escherichia coli*. *Archives of Microbiology* **154**, 60–66.
- YANISCH-PERRON, C., VIEIRA, J. & MESSING, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.
- YOSHIDA, H., KOJIMA, T., YAMAGISHI, J. & NAKAMURA, S. (1988). Quinolone-resistant mutations of the *gyrA* gene of *Escherichia coli*. *Molecular and General Genetics* **211**, 1–7.