Characterization and Nucleotide Sequence of pqqE and pqqF in Methylobacterium extorquens AM1

AMY L. SPRINGER, ROOPA RAMAMOORTHI,† AND MARY E. LIDSTROM*

Environmental Engineering Science, California Institute of Technology, W. M. Keck Laboratories 138-78, Pasadena, California 91125

Received 16 November 1995/Accepted 24 January 1996

Methylobacterium extorquens AM1 pqqEF are genes required for synthesis of pyrroloquinoline quinone (PQQ). The nucleotide sequence of these genes indicates PqqE belongs to an endopeptidase family, including PqqF of Klebsiella pneumoniae, and M. extorquens AM1 PqqF has low identity with the same endopeptidase family. M. extorquens AM1 pqqE complemented a K. pneumoniae pqqF mutant.

Pyrroloquinoline quinone (PQQ) is a prosthetic group required by several bacterial dehydrogenases, including methanol dehydrogenase (MDH) of gram-negative methylotrophs and some glucose dehydrogenases (3). PQQ is derived from two amino acids, tyrosine and glutamic acid (12, 24), but the pathway for its biosynthesis is unknown (5, 18).

Genetic studies of PQQ biosynthesis have been performed with a number of methylotrophs: Methylobacterium extorquens AM1 (14), Methylobacterium organophilum DSM760 and XX (1, 14), and Methylobacillus flagellatum (8). In these organisms, the genes for PQQ biosynthesis are required for growth on methanol. Seven PQQ complementation groups, in two clusters, have been identified in M. extorquens AM1: PqqDGCBA and PqqEF (14). The nucleotide sequence has been reported for pqqDG and part of pqqC in M. extorquens AM1 (14). Genes for PQQ biosynthesis have also been cloned and sequenced in two organisms that produce PQQ-linked glucose dehydrogenase. In Klebsiella pneumoniae six genes have been identified, pqqABCDEF (13), the first five of which correspond, respectively, to pqqDGCBA of M. extorquens AM1 (14). The same five genes, in the same gene order, also exist in Acinetobacter calcoaceticus (9). The product of the first gene is a peptide of 23 to 29 amino acids that contains conserved tyrosine and glutamic acid residues and is believed to be the peptide precursor from which PQQ is derived (10, 13, 25). The products of the other four genes, pqqGCBA in M. extorquens AM1 and pgqBCDE in K. pneumoniae, show no homology with other proteins, and their respective roles in PQQ biosynthesis are unknown (9, 13, 14). A recent study with K. pneumoniae demonstrated that the product of pqqC is required for the final step in the biosynthesis of PQQ, although the biochemical nature of that step was not determined (25). The only pqq gene product that is homologous to a protein of known function is PqqF of K. pneumoniae, which shares identity with a family of endopeptidases (13). It has been suggested that this protein is involved in processing of the peptide precursor during the biosynthesis of PQQ (13). In this study we present complementation analyses and nucleotide sequence of pqqEF of \hat{M} . extorquens AM1. M. extorquens AM1 pqqE is functionally equivalent to the PqqF endopeptidase of *K. pneumoniae*. The putative product of *M. extorquens* AM1 *pqqF* also shares some identity with endopeptidases.

All pagE mutants (EMS20 and EMS7-17) and pagF mutants (EMS7-27, EMS7-35, and EMS7-42) were derived from rifamycin-resistant M. extorquens AM1 (15). The K. pneumoniae pqqF mutant, KA216 (pAMH62/pts1103 pqqF17::Tn10 Kan^R), is a derivative of NCTC418 and was provided by P. W. Postma (University of Amsterdam, Amsterdam, The Netherlands). Escherichia coli DH5α was used for DNA manipulation (BRL, Inc.). Plasmids used for complementation of M. extorquens AM1 mutants (Fig. 1) are subclones of pELH2 (14), inserted into pRK310 (Tcr IncP1) (4). pELH2::TnphoA-E1 is described by Morris et al. (14), and a corrected map of the insert in this plasmid is shown in Fig. 1. Other plasmids used for complementation in K. pneumoniae, pEL41(HINDIII-CD), pDN102 (HINDIII-C), and pDN202(HINDIII-E), have been described (14, 15). The mobilizing plasmid used for conjugation was pRK2073 (Sm^r) (6). Plasmids used for sequencing were derivatives of pELH2 inserted into pUC19 (26).

M. extorquens AM1 strains were grown at 30°C on minimal medium described previously (7), containing 0.2% (vol/vol) methanol, 0.2% (wt/vol) methylamine, and 0.2% (wt/vol) succinate. K. pneumoniae KA216 was grown at 30°C in minimal medium A (20) with vitamins (22) or on EMB agar (Difco) supplemented with 0.2% (wt/vol) glucose or 0.2% (wt/vol) gluconate. E. coli was grown in L broth (19). Antibiotics were added at the following final concentrations (in milligrams per liter): ampicillin, 100; tetracycline, 12.5; kanamycin, 25; and streptomycin, 10. Bacto-agar, EMB agar, and Peptone were obtained from Difco Laboratories. Chemicals were obtained from Sigma Co.

Plasmids were transferred into M. extorquens AM1 strains by conjugation by using triparental matings as described previously (21). Plasmids were transferred into K. pneumoniae KA216 in a similar manner, except that strains were spotted onto Peptone agar plates (Peptone, 10 g/liter; K₂HPO₄, 2 g/liter) and transconjugants were plated onto minimal medium A containing either glucose or gluconate and onto glucose EMB agar, each supplemented with tetracycline and kanamycin to select for, respectively, the conjugated plasmid and the Kan^r insertion in KA216. For selection of pELH2::TnphoA-E1 transconjugants in KA216, kanamycin and ampicillin were used to select for, respectively, the TnphoA and an Amp^r plasmid in KA216. Recombinational rescue was scored as described previously (21). Construction of subclones was performed following the methods of Sambrook et al. (19), or as described previously (14, 21). Enzymes were obtained from New England Biolabs and Boehringer-GmbH (Mannheim, Germany).

^{*} Corresponding author. Phone: (818) 395-2132. Fax: (818) 395-2940. Electronic mail address: lidstrom@cco.caltech.edu.

[†] Present address: Center for Bioengineering, University of Washington, Seattle, WA 98195.

Vol. 178, 1996 NOTES 2155

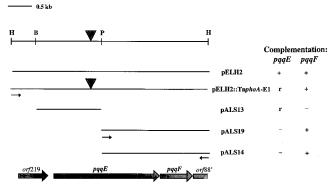


FIG. 1. Restriction map and complementation analysis of the pqqEF region using derivatives of pELH2 as indicated. Small arrows indicate direction of the lacZ promoter on the pRK310 vector, where determined. The filled triangle represents the TnphoA-E1 insertion. Complementation was scored as the ability of the clone to restore growth of the respective mutant on methanol; an r in the complementation table indicates recombinational rescue. The open reading frames are indicated at the bottom of the figure as large arrows with gene designations. Abbreviations: B = BamHI, H = HindIII, P = PstI.

Sequence was obtained from *pqqEF* subclones by using the M13 forward and reverse primers designed for pUC19 (U.S. Biochemicals) and by using single-stranded oligonucleotide primers synthesized by the Caltech Microsynthesis facility. Sequencing reactions were performed at the University of California at Los Angeles Sequencing Facility or using Dye Terminator cycle-sequencing kit (Applied Biosystems, Inc.) and an Ericomp TwinBlock System. Cycle-sequencing reactions were analyzed at the Caltech DNA Sequencing Facility. Sequence analyses were carried out using the Genetics Computer Group (GCG) program (Madison, Wis.). The GenBank accession number for the *pqqEF* sequence is L43136.

Complementation analyses of PqqE and PqqF. The complementation groups PqqE and PqqF were previously mapped to the 4.1-kb fragment in pELH2 (14). To further localize these complementation groups, subclones of pELH2 were constructed. Complementation was scored as the ability to restore growth of Pqq mutants on methanol. pALS13 (Fig. 1) contains the 1.4-kb BamHI-PstI fragment from pELH2 inserted into pRK310 and did not permit growth of PqqE or PqqF mutants on methanol, indicating that neither complementation group was completely contained within this fragment. pALS13 did recombinationally rescue PqqE mutants, suggesting that at least part of pqqE is contained within this insert. Subclones pALS14 and pALS19 (Fig. 1) contain the 2.3-kb HindIII-PstI insert from pELH2, inserted in opposite orientations into pRK310, both plasmids complemented PqqF mutants but did not complement PqqE mutants. These results indicate that pqqE lies across the PstI site and that pqqF is contained within the 2.3-kb HindIII-PstI fragment. The transposon in pELH2:: TnphoA-E1 was reported to map near the left-hand end of the pELH2 insert (14). A more complete analysis of restriction sites performed in this study revealed that the transposon was in fact 1.4 kb from the left-hand end of this insert (Fig. 1), and complementation analyses confirm that the insertion site is within pqqE.

Complementation analyses were also performed with the *K. pneumoniae* mutant, KA216, which is unable to ferment glucose because of mutations in both *pqqF* and in the PTS system (13). Transconjugants were tested for color change on EMB-glucose medium, an indicator of PQQ-linked glucose dehydrogenase activity. Transconjugants containing pELH2, the clone of *M. extorquens* AM1 *pqqEF*, turned purple on EMB-glucose

TABLE 1. Complementation of *K. pneumoniae* KA216(pAMH62/ pts1103, pqqF17::Tn10 Kan^r) with *M. extorquens* AM1 DNA

M. extorquens	M. extorquens	Color of KA216 on
AM1 clone	AM1 gene(s)	EMB glucose
pELH2	pqqEF	Purple
pELH2::TnphoA-E1	pqqF	White
pEL41	pqqDGCBA	White
pDN102	pqq'CBA	White
pDN172	mxcQE	White

plates (Table 1), indicating that pELH2 is capable of restoring PQQ-linked glucose dehydrogenase activity in KA216. Negative controls, KA216 containing M. extorquens AM1 plasmids with mox (genes required for methanol oxidation) and other pqq genes, pEL41, pDN102, and pDN202, did not show a color change on EMB glucose plates (Table 1). The KA216 transconjugants containing pELH2 were unable to grow on minimal glucose plates, suggesting that PQQ activity is not fully restored in these strains. It is possible that this is due to differences in expression from the M. extorquens DNA or because of a low degree of identity between the proteins in the respective organisms. To determine which of the two pqq genes in pELH2 was responsible for complementation, pELH2::TnphoA-E1 was introduced into KA216. This plasmid contains an insertion in pqqE but expresses functional pqqF (14). Transconjugants containing pELH2::TnphoA-E1 were white on EMB-glucose plates, indicating that M. extorquens pqqE is the gene that is functionally equivalent to K. pneumoniae pggF.

Nucleotide sequence of *pqqEF* and surrounding region. The sequence of the pELH2 insert revealed three intact open reading frames and one partial open reading frame in the direction shown from left to right in Fig. 1. The first open reading frame begins 171 bp from the leftmost *HindIII* site and encodes a putative product of 219 amino acids. This putative product of *orf219* does not have significant identity to any known proteins.

The second open reading frame begins 972 bp into the sequence, 142 bp downstream of the end of orf219. This open reading frame encodes a putative product of 709 amino acids and extends across the PstI site, indicating that this is pqqE. The predicted *M. extorquens* AM1 PqqE shares identity with a superfamily of divalent cation-containing endopeptidases, showing 22% identity over 414 amino acids with the mitochondrial processing peptidase β-subunit (MPP) from Neurospora crassa (GenBank accession number M20928). The identity is highest in the N-terminal portion of the protein, including the conserved His-X-X-Glu-His sequence believed to be involved in metal binding, and a conserved Glu required for catalysis (16, 17). There is also extensive identity to a putative MPP-like peptidase of Mycobacterium fortuitum (29% over 399 amino acids; GenBank accession number L25634), to E. coli pitrilysin (25% over 227 amino acids; GenBank accession number M17095), and to several insulin-degrading endoproteinases (IDEs; GenBank accession numbers X67269, M58465, and M21188). M. extorquens AM1 PqqE shows 24% identity over 237 amino acids to K. pneumoniae PqqF, providing further evidence that K. pneumoniae PqqF and M. extorquens AM1 PqqE are homologs. Many members of this endopeptidase family are involved in processing small peptides (17). Phylogenetic analysis suggested that M. extorquens AM1 PqqE is related to the MPP subfamily, whereas K. pneumoniae PqqF is fairly distant from the known subfamilies but is most closely related to the pitrilysin/IDE subfamily (17) (Fig. 2). Members of the MPP subfamily form heterodimers with a second subunit, α , that shares a high degree of identity with the first 2156 NOTES J. BACTERIOL.

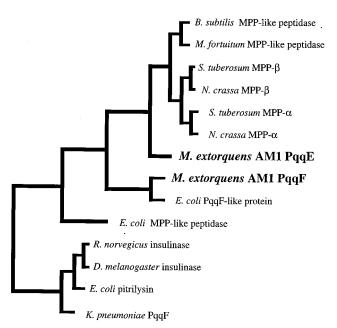


FIG. 2. Phylogenetic tree showing relationships of representative members of the MPP and Pitrilysin/IDE subfamilies, M. extorquens AM1 PqqE, M. extorquens AM1 PqqF, and K. pneumoniae PqqF. Distance between sequences was calculated by using the GCG sequence analysis package Kimura two-parameter distance algorithm, and the dendrogram was produced using the UPGMA algorithm from the Growtree program. GenBank accession numbers for these sequences are listed in the text, except for N. crassa MPP- α (J05484), S. tuberosum MPP- α (X66284), and S. tuberosum MPP- β (X80237).

subunit but lacks the catalytic residue (17). By homology, *M. extorquens* AM1 PqqE might also be expected to associate with a second subunit.

PqqE also shares 25% identity over 266 amino acids with an open reading frame in E. coli (GenBank accession number X71917), in a region near gadB that complements pqqE and pqqF mutants of M. organophilum DSM760 (23). This open reading frame encodes a protein with identity to the MPP subfamily and shares no significant identity with K. pneumoniae PqqF. The identity to M. extorquens AM1 PqqE suggests that this open reading frame is responsible for complementing M. organophilum DSM760 pqqE mutants. The role of this putative MPP-like peptidase is unknown. E. coli does not produce PQQ under standard laboratory conditions, although it does produce an inactive glucose dehydrogenase apoprotein that can be activated by exogenous PQQ (11), and an E. coli mutant has been reported that synthesizes PQQ (2). Preliminary transcriptional studies suggest the presence of a weak C₁-inducible promoter upstream of pqqE, expression from this promoter was not affected by mxb, mxc, or mxaB regulatory mutants

The third open reading frame begins 3,164 bp into the sequence, 63 bp from the end of pqqE. This open reading frame encodes a putative product of 241 amino acids, and must be pqqF, since it is the only open reading frame completely encoded within the 2.3-kb fragment that complements pqqF mutants. M. extorquens AM1 PqqF shows identity to the C-terminal half of the putative MPP-like peptidase from M. fortuitum described above (29% over 171 amino acids) and to the C-terminal half of a putative MPP-like peptidase from Bacillus subtilis (23% over 172 amino acids; GenBank accession number L08471). M. extorquens AM1 PqqF shows very low identity (16% over 170 amino acids) with the C-terminal portion of M.

BsuPep MfoPep MexPqqE MexPqqF	KETEQA LWYDRDGEQS IAWADPKVEQ IVVDLDVPQS	HVST.GVR PTLQR.TYLT	FKGLEVGHER .TPGRHWEHR PSCMTA.RDG PGVAWRDP	I.YDLININN W. ALSVENT EGYALELEAE DFIPAYVENH	VLGGS MSSR ALGGG LSSR VVGGG STSF ILGGGAFTSR	278 266 347 101
			GPLRDSGALS GSAMDDTRFA		QQLSET.IQE DEVVRVTTEV LEATEFHIDR VEAT.DVIGD	326 315 397 149
BsuPep MfoPep MexPqqE MexPqqF	.LEGVARDGI VLRRVPEALG		GSLROGLVIC IRLMAETVYS	LEDSGSRMHR .SDSQSSLAR	IGRSELNYGE IYGSALAIGE	375 363 444 198
	HKTLDEI.IN HRTIDHT.LA TVEEVRRWPV MDYIARR.ND	elnavnherv Oteavtheev Eteavthdrl LVASVTQADI	NAVAHOLLSR VAVAARYLVP		ARDPDVAIAE	408 396 494 241

FIG. 3. Alignment of portions of *M. extorquens* AM1 PqqF (MexPqqF) and PqqE (MexPqqE) with the putative peptidase from *M. fortuitum* (MfoPep) and the putative peptidase from *B. subtilis* (BsuPep). Conserved residues are indicated by gray highlighting; numbers on right refer to the residue number.

extorquens AM1 PqqE, and many of the residues conserved between these are conserved in all four proteins (Fig. 3). Since M. extorquens AM1 PqqF shows identity to two members of the MPP subfamily but lacks the catalytic portion of the peptidase, it may have a role similar to the α subunit of MPP, although no identity to the α subunits of MPP was observed. The C termini are not conserved among these peptidases, which may explain why M. extorquens AM1 PqqF does not share identity with many other peptidases. M. extorquens AM1 PqqF shows low identity (19% over 134 amino acids) to an open reading frame in E. coli immediately downstream of the gene encoding an MPP-like peptidase described above. Many of the residues conserved between M. extorquens AM1 PggF and this PggFlike protein in E. coli are the same as those conserved in Fig. 3. This open reading frame in E. coli is probably the one responsible for complementing pqqF mutants of M. organophilum DSM760 (23).

A fourth partial open reading frame starts 3,991 bp into the sequence and extends 88 amino acids to the end of the published sequence. The putative product of *orf88'* shares identity with a number of dioxygenases, including 26% identity over 86 amino acids with catechol 2,3-dioxygenase from *Rhodococcus rhodochrous* (GenBank accession number X69504) and 21% identity over 70 amino acids with catechol 2,3-dioxygenase from *Bacillus stearothermophilus* (GenBank accession number X67860). The regions sharing identity correspond to the N-terminal ends of these two catechol dioxygenases. It is not known if this putative dioxygenase is involved in PQQ biosynthesis because no mutations exist in *orf88'*.

The results discussed here show that *M. extorquens* AM1 pqqE is the functional equivalent of *K. pneumoniae pqqF* with respect to PQQ biosynthesis. Sequence analysis indicates that *M. extorquens* AM1 pqqE encodes a member of a family of endopeptidases. *K. pneumoniae* PqqF is also a member of this family, although the *M. extorquens* AM1 and *K. pneumoniae* proteins are members of different subfamilies. These peptidases are likely to be involved in processing the template peptide during PQQ biosynthesis. The putative product of *M. extorquens* AM1 pqqF shares some identity with two members of the endopeptidase family and may serve as a noncatalytic subunit associated with PqqE. These results suggest that *M. extorquens* AM1 and *K. pneumoniae* use divergent systems for processing the template peptide.

This work was supported by a grant from the Department of Energy (DEF03-87ER13757).

Vol. 178, 1996 NOTES 2157

We thank P. W. Postma for providing *K. pneumoniae* KA216 and for advice on the indicator medium.

REFERENCES

- Biville, F., E. Turlin, and F. Gasser. 1989. Cloning and genetic analysis of six pyrroloquinoline quinone biosynthesis genes in *Methylobacterium organophilum* DSM760. J. Gen. Microbiol. 135:2917–2929.
- Biville, F., E. Turlin, and F. Gasser. 1991. Mutants of Escherichia coli producing pyrroloquinoline quinone. J. Gen. Microbiol. 137:1775–1783.
- Davidson, V. 1992. A brief history of quinoproteins, p. 1–22. In V. Davidson (ed.), Principles and application of quinoproteins. Marcel Dekker, New York.
- Ditta, G., T. Schmidhauser, E. Yakobson, P. Lu, X.-W. Liang, D. R. Finlay, D. Guiney, and D. R. Helinski. 1985. Plasmids related to the broad host range vector pRK290, useful for gene cloning and for monitoring gene expression. Plasmid 13:149-153.
- Duine, J. A., J. Frank, and P. E. J. Verwiel. 1980. Structure and activity of the prosthetic group of methanol dehydrogenase. Eur. J. Biochem. 108:187–192.
- Figurski, D. H., and D. R. Helinski. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc. Natl. Acad. Sci. USA 76:1648–1652.
- Fulton, G. L., D. N. Nunn, and M. E. Lidstrom. 1984. Molecular cloning of a malyl CoA lyase gene from *Pseudomonas* AM1, a facultative methylotroph. J. Bacteriol. 160:718–723.
- 8. Gomelsky, M., F. Biville, F. Gasser, and Y. Tsygankov. 1992. Cloning and study of the genes involved in PQQ biosynthesis in an obligate methylotroph, *Methylobacillus flagellatum*, abstr. B81. *In* Program and Abstracts of the 7th International Symposium on Microbial Growth on C₁ Compounds.
- Goosen, N., H. P. A. Horsman, R. G. M. Huinen, and P. van de Putte. 1989. Acinetobacter calcoaceticus genes involved in biosynthesis of the coenzyme pyrroloquinoline quinone: nucleotide sequence and expression in Esche-richia coli K-12. J. Bacteriol. 171:447–455.
- Goosen, N., R. G. M. Huinen, and P. van de Putte. 1992. A 24-amino-acid polypeptide is essential for the biosynthesis of the coenzyme pyrroloquinoline quinone. J. Bacteriol. 169:303–307.
- Hommes, R. J. W., P. W. Postma, O. M. Neijssel, D. W. Tempest, P. Dokter, and J. A. Duine. 1984. Evidence of a quinoprotein glucose dehydrogenase apoenzyme in several strains of *Escherichia coli*. FEMS Microbiol. Lett. 24:329–333
- Houck, D. R., J. L. Hanners, and C. J. Unkefer. 1991. Biosynthesis of pyrroloquinoline quinone. 2. Biosynthetic assembly from glutamate and tyrosine. J. Am. Chem. Soc. 113:3162–3166.
- Meulenberg, J. J. M., E. Sellink, N. H. Riegman, and P. W. Postma. 1992. Nucleotide sequence and structure of the *Klebsiella pneumoniae pqq* operon. Mol. Gen. Genet. 232:284–294.

14. Morris, C. J., F. Biville, E. Turlin, E. Lee, K. Ellermann, W.-H. Fan, R. Ramamoorthi, A. L. Springer, and M. E. Lidstrom. 1994. Isolation, phenotypic characterization, and complementation analysis of mutants of *Methylobacterium extorquens* AM1 unable to synthesize pyrroloquinoline quinone and sequence of *pqqD*, *pqqG*, and *pqqC*. J. Bacteriol. 176:1746–1755.

- Nunn, D. N., and M. E. Lidstrom. 1986. Isolation and complementation analysis of 10 methanol oxidation mutant classes and identification of the methanol dehydrogenase structural gene of *Methylobacterium* sp. strain AM1. J. Bacteriol. 166:581–590.
- Rawlings, N. D., and A. J. Barrett. 1991. Homologues of insulinase, a new superfamily of metalloendopeptidases. Biochem. J. 275:389–391.
- Rawlings, N. D., and A. J. Barrett. 1995. Evolutionary families of metallopeptidases. Methods Enzymol. 248:183–228.
- Salisbury, S. A., H. S. Forrest, W. B. T. Cruse, and O. Kennard. 1979. A novel coenzyme from bacterial primary alcohol dehydrogenases. Nature (London) 280:843–844.
- Sambrook, T., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Scholte, B. J., A. R. J. Schuitema, and P. W. Postma. 1982. Characterization
 of factor III^{Gic} in catabolite repression-resistant (crr) mutants of Salmonella
 typhimurium. J. Bacteriol. 149:576–586.
- Springer, A. L., H.-H. Chou, W.-H. Fan, E. Lee, and M. E. Lidstrom. 1995. Methanol oxidation mutants in *Methylobacterium extorquens* AMI: identification of new genetic complementation groups. Microbiology 141:2985–2993
- Staley, J. T. 1981. The genera *Prosthecomicrobium* and *Ancalamicrobium*, p. 456–460. *In* M. P. Starr, H. Stolp, H. G. Trüper, A. Balows, and H. G. Schlegel (ed.), The procaryotes, vol. 1. Springer-Verlag, New York.
- Turlin, E., F. Biville, and F. Gasser. 1991. Complementation of Methylobacterium organophilum mutants affected in pyrroloquinoline quinone biosynthesis genes paqE and paqF by cloned Escherichia coli chromosomal DNA. FEMS Microbiol. Lett. 83:59–64.
- van Kleef, M. A. G., and J. A. Duine. 1988. L-tyrosine is the precursor of PQQ biosynthesis in *Hyphomicrobium X*. FEBS Lett. 237:91–97.
- Velterop, J. S., E. Sellink, J. J. M. Meulenberg, S. David, I. Bulder, and P. W. Postma. 1995. Synthesis of pyrroloquinoline quinone in vivo and in vitro and detection of an intermediate in the biosynthetic pathway. J. Bacteriol. 177: 5088–5098
- Yanisch-Perron, C. L., J. Vieira, and J. Messing. 1985. Improved cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119.