

DNA Sequencing of the Seven Remaining Structural Genes of the Gene Cluster Encoding the Energy-Transducing NADH–Quinone Oxidoreductase of *Paracoccus denitrificans*^{†,‡}

Xuemin Xu, Akemi Matsuno-Yagi, and Takao Yagi*

Division of Biochemistry, Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, California 92037

Received August 26, 1992; Revised Manuscript Received October 16, 1992

ABSTRACT: In our previous papers, seven structural genes (*NQO1*–*7*) of the energy-transducing NADH–quinone (Q) oxidoreductase of *Paracoccus denitrificans* were characterized [Xu, X., Matsuno-Yagi, A., & Yagi, T. (1991a) *Biochemistry* 30, 8678–8684; (1991b) *Biochemistry* 30, 6422–6428; (1992a) *Biochemistry* 31, 6925–6932; (1992b) *Arch. Biochem. Biophys.* 296, 40–48]. This paper reports the identification, cloning, and sequencing of seven additional structural genes in the same gene cluster (*P. denitrificans* enzyme complex). These seven genes, designated *NQO8*–*14*, are composed of 1038, 492, 603, 306, 2112, 1542, and 1500 base pairs, respectively. The polypeptides encoded by the *NQO8*–*14* genes are homologous, respectively, to the ND1 product, the 23-kDa polypeptide, and the ND6, ND4L, ND5, ND4, and ND2 products of the bovine NADH–Q oxidoreductase. The order of the 14 structural genes of the *Paracoccus* energy-transducing NADH–Q oxidoreductase in the gene cluster is *NQO7*, *NQO6*, *NQO5*, *NQO4*, *NQO2*, *NQO1*, *NQO3*, *NQO8*, *NQO9*, *NQO10*, *NQO11*, *NQO12*, *NQO13*, and *NQO14*. Downstream from the *NQO14* gene an open reading frame (designated URF240) was detected which encodes a predicted polypeptide homologous to the biotin [acetyl-CoA–carboxylase] ligase of *Escherichia coli*. In addition, a putative terminal sequence motif was observed downstream of the *NQO14* gene, suggesting that the structural gene *NQO14* is the 3'-terminal gene of the *Paracoccus* NADH–Q oxidoreductase gene cluster. Nucleotide sequencing of the entire gene cluster revealed the presence of three unidentified reading frames: one between the *NQO3* and *NQO8* genes and other two between the *NQO9* and *NQO10* genes. These are designated URF4, URF5, and URF6 and are composed of 768, 393, and 405 base pairs, respectively. The possible functions of the putative proteins encoded by URF5 and URF6 are discussed.

The NADH–quinone (Q)¹ oxidoreductases present in the respiratory chain of various species of bacteria can be divided into two groups depending on whether these enzyme complexes bear an energy coupling site (Yagi, 1989, 1991, 1993; Yagi et al., 1988, 1992). Those enzymes that bear the coupling site are designated as NDH-1 and those that do not as NDH-2. The bacterial NDH-1 enzyme complexes appear to be related to the mammalian mitochondrial NADH–UQ oxidoreductase (complex I) as judged by similarities in prosthetic groups, polypeptide compositions, and specific inhibitors (Yagi, 1989, 1991, 1993; Yagi et al., 1992; Meinhardt et al., 1987).

Paracoccus denitrificans is a Gram-negative soil bacterium (John & Whatley, 1975, 1977). Aerobically grown *P. denitrificans* expresses a mammalian mitochondrial-type respiratory chain (Stouthamer, 1980) which appears to contain

NDH-1 but not NDH-2 (Yagi, 1986, 1991). This *Paracoccus* NDH-1 contains noncovalently bound FMN and 4–5 EPR-visible FeS clusters and appears to be composed of at least 10 unlike subunits (Yagi et al., 1992; Yagi, 1986; Meinhardt et al., 1987). Immuno-cross-reactivity has been reported between several of the *Paracoccus* NDH-1 and bovine complex I subunits (Yagi, 1986; Xu & Yagi, 1991; George et al., 1986; Yagi & Dinh, 1990; Xu et al., 1991a,b, 1992a,b). In addition, both the *Paracoccus* NDH-1 and the bovine complex I are inhibited by several of the same inhibitors (e.g., rotenone, piericidin A, capsaicin, and DCCD; Yagi, 1987, 1990, 1991; Yagi et al., 1992).

The NADH-binding subunit of the *Paracoccus* NDH-1 was previously identified in this laboratory using direct photoaffinity labeling with [³²P]NAD(H) (Yagi & Dinh, 1990). This subunit was purified and subjected to amino acid sequence analysis. On the basis of partial amino acid sequence information from this subunit, we succeeded in cloning the 18-kbp *Paracoccus* NDH-1 gene cluster. The DNA sequence of the entire gene cluster was derived from three DNA fragments of this gene cluster: 5.7-kbp *EcoRI*, 12-kbp *PstI*, and 3.4-kbp *HindIII* DNA fragments. The fragments are designated pXT-1, pXT-2, and pXT-3, respectively (Yagi, 1993; Yagi et al., 1992). The arrangement of these DNA fragments, from upstream, is pXT-3, pXT-1 and pXT-2. The 3'-terminal region of pXT-3 overlapped with the 5'-terminal region of pXT-1 by 220 bp. The 3'-terminal region of pXT-1 and the 5'-terminal region of pXT-2 overlapped each other by approximately 1.1 kbp. The DNA sequences of pXT-1

[†] This work was supported by U.S. Public Health Service Grant R01 GM33712. Facilities for computer were supported by U.S. Public Health Service Grant M01 RR00833 for the General Clinical Research Center. Synthesis of oligonucleotides was, in part, supported by the Sam & Rose Stein Charitable Trust. This is publication 7517-MEM from The Scripps Research Institute, La Jolla, CA.

[‡] The nucleotide sequence in this paper has been deposited with GenBank under Accession Number L02354.

* To whom correspondence should be addressed.

Abbreviations: Q, quinone; UQ, ubiquinone; NAD(H), nicotinamide adenine dinucleotide (reduced); NDH-1 or complex I, energy-transducing NADH–quinone oxidoreductase; NDH-2, NADH–quinone oxidoreductase lacking an energy coupling site; FMN, flavin mononucleotide; FP, IP, and HP, respectively, the flavoprotein, the iron–sulfur protein, and the hydrophobic protein fractions of complex I; DCCD, *N,N*-dicyclohexylcarbodiimide; (k)bp, (kilo)base pairs; FeS cluster, iron–sulfur cluster; EPR, electron paramagnetic resonance; mtDNA, mitochondrial DNA.

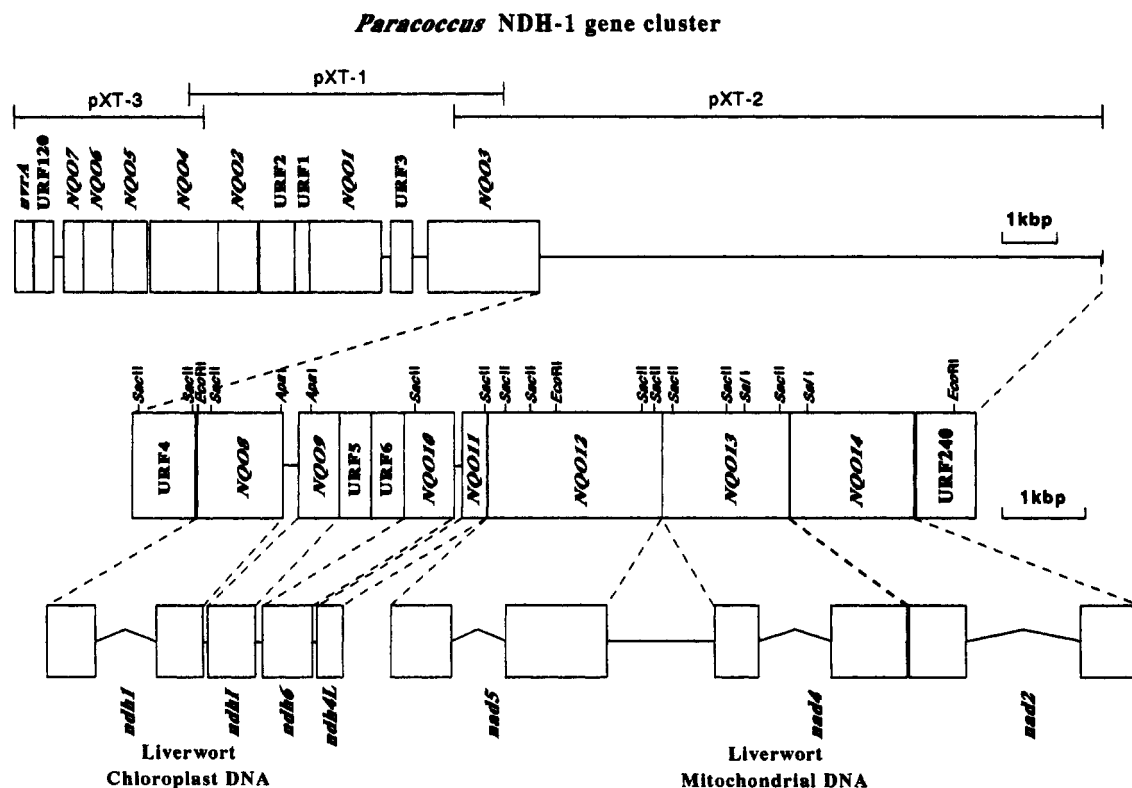


FIGURE 1: Gene map of the *Paracoccus* NDH-1 gene cluster. The gene map of the pXT-1, pXT-2, and pXT-3 DNA fragments is illustrated together with the URF4, *NQO8*–*10*, URF5, URF6, *NQO11*–*14*, and URF240 genes. The *NQO1*–*7* genes are the structural genes encoding the *Paracoccus* NDH-1 subunits homologous, respectively, to the 51-, 24-, 75-, 49-, 30-, and 20-kDa and the ND3 subunits of bovine complex I. For comparison, the gene arrangements of liverwort mitochondrial DNA (Oda et al., 1992) and liverwort chloroplast DNA (Ohyama et al., 1986) are shown. Gene arrangements similar to the liverwort chloroplast DNA have recently been reported in cyanobacterial DNA (K. Steinmüller, unpublished results, GenBank/EMBL Accession Number X62517) (Takahashi et al., 1991). — and ^ indicate, respectively, the noncoding region and the intron.

and pXT-3 have been published earlier (Xu et al., 1991a,b, 1992a,b).

This paper describes the nucleotide sequence of the remaining 10.3 kbp of pXT-2. This chromosomal DNA region contains seven structural genes encoding subunits of the *Paracoccus* NDH-1. These genes have been designated *NQO8* (1038 bp), *NQO9* (492 bp), *NQO10* (603 bp), *NQO11* (306 bp), *NQO12* (2112 bp), *NQO13* (1542 bp), and *NQO14* (1500 bp). In addition, three unidentified reading frames were observed in this gene cluster. Furthermore, an open reading frame (URF240) encoding a predicted polypeptide homologous to biotin [acetyl-CoA-carboxylase] ligase was detected downstream of the *NQO14* gene. The identification of the 3'-terminus of the gene cluster bearing the *Paracoccus* NDH-1 has been discussed.

MATERIALS AND METHODS

DNA Sequencing Strategy. General DNA techniques were carried out essentially according to Sambrook et al. (1989). As described previously (Xu et al., 1992b), we have isolated a plasmid, designated pXT-2, containing an approximately 12-kbp *Pst*I *Paracoccus* DNA fragment. pXT-2 was utilized as template for DNA sequencing in this paper. The isolated pXT-2 insert was digested with *Eco*RI, *Sal*I, *Sac*II, and *Apa*I. All the fragments thus produced were subcloned into the pBluescript KS phagemid vector and amplified in *Escherichia coli* JM109 as described previously (Xu et al., 1991b).

The DNA sequencing was performed by the dideoxynucleotide method of Sanger et al. (1977) using the 7-deaza-dGTP sequencing kit from Pharmacia. Deletion mutants for nucleotide sequencing were created with the exo III/mung bean nuclease deletion kit from Stratagene and Takara. The

universal primers (T3, T7, and M13 reverse primers) and unique internal primers (18 bases in length) were used in these experiments. When compression regions were encountered, the dITP sequencing kit (Pharmacia) was used instead of the 7-deaza-dGTP. Both complementary DNA strands were sequenced at least three times. Sequenase version 2 was used as polymerase for the DNA sequencing.

Analysis of Nucleotide and Protein Sequences. As described previously (Xu et al., 1991a,b, 1992a,b), the University of Wisconsin Genetic Computer Group's software programs were used to analyze the sequence data (Devereux et al., 1984). The sequence was searched for open reading frames and terminators by the CODONPREFERENCE and TERMINATOR programs, respectively. A comparison of the polypeptides was conducted with the BESTFIT and PILEUP programs. The FASTA and PROFILESEARCH programs were used to search the GenBank/EMBL sequence databases for proteins having some homology to the polypeptides predicted by the structural genes and the open reading frames.

Materials. Acrylamide, bisacrylamide, and SDS were from Bio-Rad; 7-deaza-dGTP and dITP sequencing kits and universal primers were from Pharmacia LKB; [α - 35 S]-thio-dATP was from Amersham; Sequenase version 2 was from U.S. Biochemical Corp.; deletion mutant kits, pBluescript, and conventional restriction enzymes were from Stratagene; λ /HindIII DNA standard was from New England Biolabs; and additional deletion mutant kits were from Takara.

RESULTS AND DISCUSSION

Sequence Analysis. As shown in Figure 1, we have previously sequenced the *P. denitrificans* DNA fragments pXT-1 and pXT-3. These fragments were obtained by

Table I: Characteristics of the *NQO8*, *NQO9*, *NQO10*, *NQO11*, *NQO12*, *NQO13*, *NQO14*, *URF4*, *URF5*, *URF6*, and *URF240* Genes and Gene Products

gene	DNA length (bp)	characteristics of polypeptide ^a product				homologue
		no. of amino acids	MW	pI		
NQO8	1038	345	38751	5.39	ND1 product (HP) ^b	
NQO9	492	163	18959	6.49	23 kDa (HP) ^b	
NQO10	603	200	21819	6.53	ND6 product (HP?) ^b	
NQO11	306	101	10856	7.69	ND4L product (HP) ^b	
NQO12	2112	703	77705	6.71	ND5 product (HP?) ^b	
NQO13	1542	513	56417	7.64	ND4 product (HP?) ^b	
NQO14	1500	499	52536	5.42	ND2 product (HP) ^b	
URF4	768	255	26869	9.67		
URF5	393	130	13765	6.26	γ -carboxymuconolactone decarboxylase ^c	
URF6	405	134	15002	4.72	γ -carboxymuconolactone decarboxylase ^c	
URF240	723	240	24375	5.38	biotin [acetyl-CoA-carboxylase] ligase ^d	

^a Predicted from nucleotide sequence. ^b Bovine mitochondria. ^c *A. calcoaceticus*. ^d *E. coli*.

digesting *P. denitrificans* genomic DNA with *EcoRI* and *HindIII*. The DNA fragments of interest (pXT-1, pXT-2, and pXT-3) were isolated, amplified, and sequenced. The pXT-1 and pXT-3 fragments were found to contain the *NQO1*–*7* genes, corresponding to several of the structural genes of the *Paracoccus* NDH-1 enzyme complex (Xu et al., 1991a,b, 1992a,b). These genes are arranged in the *Paracoccus* DNA in the following order: *NQO7*, *NQO6*, *NQO5*, *NQO4*, *NQO2*, *NQO1*, and *NQO3*. Unfortunately, because the *NQO3* gene was truncated at the 3'-terminus of pXT-1 it was not possible to obtain the complete sequence of this gene from pXT-1. The remaining sequence of the *NQO3* gene was detected in the *PstI* 12-kbp DNA fragment designated pXT-2. In addition to completing the sequence of *NQO3*, DNA sequencing downstream from the *NQO3* gene indicated the presence of polypeptides homologous to a number of mtDNA gene products of bovine complex I (Xu et al., 1992b). Complete DNA sequencing of pXT-2 revealed (Figure 1) the presence of seven structural genes encoding polypeptides homologous to polypeptides of bovine complex I. These structural genes are located downstream from the *NQO3* gene and have been designated *NQO8*–*14* (see Figure 1). These genes encode homologues of the bovine complex I ND1 product, 23-kDa polypeptide, and ND6, ND4L, ND5, ND4, and ND2 products, respectively (see Table I).

We have shown in a previous paper that the *NQO7* gene is the 5'-terminal structural gene of the *Paracoccus* NDH-1 gene cluster (Xu et al., 1992a). It was of interest to locate the 3'-terminus of this gene cluster. This was done by searching the sequences downstream from *NQO14* with the TERMINATOR program from the University of Wisconsin Genetic Computer Group. This program identified putative terminator sequences downstream of the *NQO14* gene (see Figure 2), suggesting that the *NQO14* gene may be the 3'-terminal structural gene of the gene cluster encoding the *Paracoccus* NDH-1. In addition to identifying the 3'-terminus of the gene cluster, Figure 1 indicates that one open reading frame (URF240), located downstream of the *NQO14* gene, was also detected. When the putative amino acid sequence encoded by the URF240 was compared with amino acid sequences present in the GenBank/EMBL database, significant homology was found (31% identity) between the gene product of URF240 and the region from residue 83 to the C-terminus of a bifunctional protein encoded by the *birA* gene of *E. coli*

(Howard et al., 1985) (data not shown). According to Howard et al. (1985), the regions corresponding to the residues 1–80 and 83–321 operate, respectively, as a biotin repressor and a biotin [acetyl-CoA-carboxylase] ligase. With the exception of the *birA* protein of *E. coli*, no polypeptides having any significant similarity to the predicted polypeptide encoded by URF240 were found in the GenBank/EMBL database. This result appears to support that the *NQO14* gene is the 3'-terminal structural gene of the *Paracoccus* NDH-1 gene cluster.

As shown in Figure 2, the initiation codons of the *NQO8*–*14* genes were all found to be preceded by sequences resembling the Shine–Dalgarno ribosome-binding site (Shine & Dalgarno, 1975). The G-C content of the *NQO8* (61.8%), *NQO9* (62.4%), *NQO10* (65.0%), *NQO11* (57.2%), *NQO12* (63.8%), *NQO13* (63.7%), and *NQO14* genes (66.1%) and of URF240 (72.5%) was high. This has also been found to be the case for other genes of *P. denitrificans* including not only these that code for NDH-1 but also those that code for complex III, cytochrome oxidase, methylamine dehydrogenase, etc. (Harms et al., 1985, 1987; Van Spanning et al., 1990a,b; Xu et al., 1991a,b, 1992a,b; Kurowski & Ludwig, 1987; Raitio et al., 1987).

Three open reading frames were detected by the CODON-PREFERENCE program. One is located between the *NQO3* and *NQO8* genes and is designated URF4 [URFs 1–3 were identified in pXT-1 and pXT-3 and are described in Xu et al. (1991a,b)]. The other two are present between the *NQO9* and *NQO10* genes and are designated URF5 and URF6, respectively (see Figure 1). The G-C content of URF4 (67.7%), URF5 (68.2%), and URF6 (63.7%) was also high as expected from other *Paracoccus* structural genes. No open reading frames were evident on the complementary strand (data not shown).

Characterization of the *NQO8*–*14* Gene Products and Comparison to Their Mammalian Complex I and Chloroplast NDH-1 Counterparts. Bovine complex I can be resolved into a water-soluble fraction and a water-insoluble fraction (hydrophobic protein fraction, HP) by treatment with NaClO₄ (Hatefi et al., 1985; Hatefi, 1985). The water-soluble fraction can be further resolved into two fractions by ammonium sulfate fractionation (Ragan, 1987; Hatefi et al., 1985; Hatefi, 1985). The two water-soluble fractions have been designated the flavoprotein fraction (FP) and the iron-sulfur protein fraction (IP). On the basis of results illustrating that when human mitochondria were resolved by deoxycholate and ammonium acetate fractionation, the polypeptides encoded by these genes were present in a fraction accompanying the NADH dehydrogenase activity (Chomyn et al., 1985b), the seven mtDNA ND gene products are believed to encode several of the structural polypeptides of mammalian complex I. The ND4L, ND3, and ND1 gene products have been shown immunochemically to be located in the bovine HP (Chomyn et al., 1985a; Yagi & Hatefi, 1988). The ND2 product, which can be extracted in highly purified form by chloroform-methanol extraction from complex I, is also located in the bovine HP (Yagi and Hatefi, unpublished results). However, with the exception of the ND1 product, which has been identified as the DCCD- and rotenone-binding subunit, the structural or functional roles of these polypeptides remain to be clarified (Yagi, 1987; Yagi & Hatefi, 1988; Earley et al., 1987). Likewise, it also remains unclear as to whether all the ND polypeptides are components of complex I or whether some of them are only coprecipitating contaminants. In an attempt to clarify these issues, the deduced amino acid sequences of

1 ^{NQ03} TCGGCGATGGCCGCGGCCCAAGGCCCGCGCGTTGGCGGCGAGT ^{URF4} GACGGGGATGCGGCCCGCCTTGCCCTTCTGAGCGGTATCGGCGCCCTGGC
 S A M A A A R K A P A P L A A E * M R P R L A L L S G I G A L A

101 CCTGCTGGCGGCTGCGCCGAGACCAAGGATGCGCGCCGCGGCCCAAGCGGCCGAACGGGCGTGGCGGTGCTGCAAAGTGACGCAGGCGCGCCGGAC
 L L A G C A E T K D A R P R P K P A E L G V A V L Q S D A G A P D

201 AAGAAGACGCTGACCAGGGGCGGATCTGACCAAATCCGGCGAGATCCTGATTCTCGAACCACGGGTTCGGTCAGCGAGATGGCGCTGGACAGCCCCG
 K K T L T R G R I S T K S G E I L I L E P D G S V S E M A L D S P E

301 AGGGCCGCGACGCCCTTCGCGGTGACCGAGGCGGAGTTGATGGCTCTGAACCGGAATCTGGAGCTGGACCTGTCCGGCCTGCCAAACATGGCCCCGGTGGC
 G R D A F A V T E A E L M A L N A N L E L D L S G L P N M A P V R

401 CAAGCGCGAGCCACGGCGCAGGAAAAGGCGTGGCTGCATTCTCGGCCCGTTCCCGCCGTTGCGGCTGAACCTGCCGTGAGTTTCGAGGCCGATCCC
 K R E P T A Q E K A L A A F S A R S R P L R L N L P V S F E A D P

501 CAGGATTTCCAGGGCGCGCAGGTGCGCGCGTGGCCAGTTCGCGCAAGGGGGCGGACAGCCTGGTGAAGTTACGGCAAACCTGCGCGGCGGGGTGGACG
 Q D F Q G A Q V A A L A S S R K G A D S L V E V T A N L R G G V D A

601 CCGACACCGCCTTCGCCTATGCCACATGCGCGCTGGCCAGCTGGGCCGATGCGAAAGGGCGTCTACGCCCGGCACATCCGCACGCTCCGGGACAAGCG
 D T A F A Y A T C A L A S W A D A K G A S Y A R H I R T L R D K R

701 CAACGGCAAGATGGTCGTCGGTTCGGTATTACGCTTTCCGAGAAAAACCGATGGGCCTGACGGTGATGACAACGAAACAGACCTTGCAAGATGCAAG
 N G K M V V G S V F T L S E K K P M G L T V M T T K Q T L Q E C K

801 ^{NQ08} GCACGCGCAATCCCGCGCGTGATGAGAGGACGAAAGATCATGGCTGAATTCTGGGCCTCGCCCTACGGTTTCGCGCTCAGCATGCTCTGACGGGTG
 A R G I P A A * M A E F W A S P Y G F A L S M L L Q G L

901 TGGCGGTCAATCGCCTTCGTATGGGTTGCTGATCTTCAATGGTCTATGGCGACCGCAAGATCTGGGCGGCGGTGCAGATGCGCGCGCGCCGAACGTGGT
 A V I A F V M G S L I F M V Y G D R K I W A A V Q M R R G P N V V

1001 CGGCCCTGGGGCTTGTGACACCTTCGCGACGCGCTGAAATACATCGTCAAGGAAATCGTCATTCCGGCCGCGCGGACAAGTTCTGCTATTTCTGT
 G P W G L L Q T F A D A L K Y I V K E I V I P A G A D K F V Y F L

1101 GCGCCCTTCCTGTCGATGATGCTGGCGCTGTTGCGCTTTGTGGTCATCCCTTCGACGAAGGCTGGGTGATGGCGAACATCAACGTGGGCATCCTGTTCA
 A P F L S M M L A L F A F V V I P F D E G W V M A N I N V G I L F I

1201 TCTTTGCCGCTCCTCTCTCGAGGTCTACGGCGTCATCATGGCGGCTGGGCCCTCGAACTCGAAATATCCGTTCTCGCCTCGTTGCGGTGCGGCGCGCA
 F A A S S L E V Y G V I M G G W A S N S K Y P F L A S L R S A A Q

1301 GATGATCTCTACGAGGTGCTGCTGGGCTGATCATCATCGGGATCATCATCTCGACCGGCTCGATGAACCTGACCGCCATCGTCGAGGCGCACGGCGGC
 M I S Y E V S L G L I I I G I I I S T G S M N L T A I V E A H G G

1401 GATTACGGGTGCTGAACTGGTATTGGCTGCCGACCTGCCGATGGTGGTGTCTTCTGCTCGGCGCTCGCCGAGTGCAACCGCCACCCCTTCGACC
 D Y G L L N W Y W L P H L P M V V L F F V S A L A E C N R P P F D L

1501 TGGTCGAGGCGAGTCCGAACCTGGTCGCGGCTTCATGACCGAATATTCGTCACGCCCTACCTGCTGTTCATGGCGGGCGAATATATCGCCATGTATCT
 V E A E S E L V A G F M T E Y S S T P Y L L F M A G E Y I A M Y L

1601 GATGTGCGGCTCCTGTGCTGCTGTTCTTCGGCGGCTGGCTGTGCGCGGTGCGGTTTCATCGCCGACGGCTGGTGGTGGATGGTCATCAAGATGTGGTTC
 M C A L L S L L F F G G W L S P V P F I A D G W W W M V I K M W F

1701 TGGTCTACATGTTCCCATGGTCAAGGCGATCGTGCCGCTACCGCTACGACAGCTGATGCGCATCGGCTGGAAGGTGTTCTGCCCTGTGCTGG
 W F Y M F A M V K A I V P R Y R Y D Q L M R I G W K V F L P L S L G

1801 GCTGGGTGGTGTGGTGGCGATCCTGGCCGCTACGAAATCCTGGGCGGCTTCTGGGCCGCTTTCGGGTGCGAGGATAAGCCATGTATTCCAACAAGCT
 W V V L V A I L A R Y E I L G G F W A R F A V G G *

1901 AGCCGAGGCATTTGCCAAAGCCAGTGAGACCGACAAGGTGGAGCTGGCGCATTTCTCGCCAACAATCTGGACAGTGGGAACAAGCGCAGGACGGCGCG
 2001 CGCATTCCGGGATTATGTCGCGCGTTCGACCGTTTCGCGGCCCGCAGGGAAGGGGAGCGTTTCTGATGGCCTTCGATTTCGCCCGTGGACCAATATT
 M A F D F A R A T K Y F

2101 CCTGATGTGGGATTTTCATCAAGGCTTCGGCCTGGGGATGCGCTATTTCTGTGCGCCCAAGCGGACGCTGAACCTACCGCATGAAAGGGCGCGCTGTG
 L M W D F I K G F G L G M R Y F V S P K P T L N Y P H E K G P L S

- 2201 CCGCGCTTCCGGGGCGAACACGCACTGCGCCGCTATCCCAACGGCGAGGAAACGCTGCATCGCCTGCAAGCTGTGCGAGGCGGTCTGCCCGGCACAGGCCA
P R F R G E H A L R R Y P N G E E R C I A C K L C E A V C P A Q A I
- 2301 TCACCATCGACGCGGAACGCGCGAGGACGGCTCGCGCCGGACCACGCGCTACGACATCGACATGACGAAATGCATCTATTGCGGCTTCTGCCAGGAAGC
T I D A E R R E D G S R R T T R Y D I D M T K C I Y C G F C Q E A
- 2401 CTGCCCCGTGATGCCATCGTCGAGGGTCCGAACCTCGAATATGCGACCGAAACCCGCGAGGAGCTGTTCTACGACAAGCAGAAGCTGCTTCCGAACGGC
C P V D A I V E G P N F E Y A T E T R E E L F Y D K Q K L L A N G
- 2501 GAACGCTGGGAGGCGGAGATCGCCCGCAACCTGCAACTGGATGCGCCCTACAGATGACCGCTGACGCGCATATCAGCGCGGCGCGCCCTGGCCGAGCA
E R W E A E I A R N L Q L D A P Y R *
M T A D A A Y Q R G R A L A E H
- 2601 TCTCAATCCGGGGATGGAGGTGGCCCTGCGCGACCGCTATGGCCGCTGGCTGCCCGATGCGGTGGCCGAAACGGTCGTGGGTACGCGGATGGCGAGGT
L N P G M E V A L R D R Y G R W L P D A V A E T V V G H G M G E V
- 2701 TATGCGCGCGAGGGGCTGGATCTGAAGACCCGGCTGCTGGTCACCGTCGGCGCATTGGCGCGATGGCGGGCAGACCCGGCCGAGCTCAAGGTGAACG
Y A R E G L D L K T R L L V T V G A L A A M G G Q T R P Q L K V N V
- 2801 TGGCCTCGGCCCTGCGCGCGGGCGCAAGCGCGGGGAAATCTGCGAGGCGATCTCCAGATGCACCTCTATGGCGGGATGCCGGCGCGATCAACGCGTT
A S A L R A G A S A R E I C E A I F Q M H L Y G G M P A A I N A L
- 2901 GAACGCCGCTATCGAGGTTTTCGAAGCGGAAGGAACAAGTCCATGACCGATGCCCTCCAGAAGCTGTTCCAGCAGATGCTGCAATCCGGCCAGGAGATGG
N A A I E V F E A E G T S P *
M T D A F Q K L F Q Q M L Q S G Q E M A
- 3001 CGCGCGCCTTCAACCCGGCGCTCGAGCATTTGACATGCGCGCCATGGAAAGCTGGTGCCGACCATTCCTGCCGACATGCTGGAAATGTGGTTCCGGCA
R A F N P A L E H F D M R A M E K L V P T I P A D M L E M W F G K
- 3101 GACCTTCAACCGCGAGGGGCTGGACGCCAAGACGCGCCTGCTTCTCACCATCGGCGCGATCACCGTTACGGGCGCGCTGGCCGAGCCGCAACTGCGCATG
T F N R E G L D A K T R L L L T I G A I T V Q G A L A E P Q L R M
- 3201 ACCGTCCGTGAGGCGCTGGCGGCGGCGCCACCAAACGCGAGATCGCGGAGACGATCTTTCAGATGAGCATGTTCCGGCGGCTGCCCGCATGCAGAAGG
T V R Q A L A A G A T K R E I A E T I F Q M S M F G G L P A M Q K A
- 3301 CGCTGGAATCGCCAGAGCGTCTATGCCGAGGAGGACGAGGAATGATGACCTTCGCTTCTACCTGTTGCGATCAGCGCTGCGTGGCGGGTTCATG
L E I A Q S V Y A E E D E E *
M M T F A F Y L F A I S A C V A G F M
- 3401 GTCGTGATCGGCCCAACCCGGTGCAATTCGGTGTGTGGCTGATCCTGGCCTTCTGTCCGCGGCGGGGCTGTTCTGCTGCTGCAAGGCGCGGAGTTCTGCG
V V I G R N P V H S V L W L I L A F L S A A G L F V L Q G A E F V A
- 3501 CCATGCTGCTGGTCTGCTGCTATGTCGGCGCGTGGCGGTGCTGTTCTGTTCTGCTGCTGATGCTGGACGTGGATTTCCGCCGAGCTGAAGGGCGAACT
M L L V V V Y V G A V A V L F L F V V M M L D V D F A E L K G E L
- 3601 GGCGCGCTACTGCGCGTGGCGCTGGTGAITGGCGTCTGCTGCGCGCAACTGGGCAITCGCCTTCTCGGGCTGGACGCCCTCGGACCAGGCCGAAAGC
A R Y L P L A L V I G V V L L A Q L G I A F S G W T P S D Q A E S
- 3701 CTGCGCGCCGCTCCGGTGGATGCGGCGGTGAGAACACCTGGGCGCTCGGCTCGTGTCTTACGACCGCTATGTGCTGATGTTCCAGCTTGCCGGCCTGG
L R A A P V D A A V E N T L G L G L V L Y D R Y V L M F Q L A G L V
- 3801 TGCTGCTGGTGGCGATGATCGGGCGGATCGTCTGACCATGCGCCACCGCAAGGACGTCAGCGCCAGAACGTCGCTGGAACAGATGTGGCGCGACCCGGC
L L V A M I G A I V L T M R H R K D V K R Q N V L E Q M W R D P A
- 3901 CAAGACGATGGAATTGAAGGACGTGAAGCCGGGCGAGGGCTTTGAACGCGCGAGGCTGGACCGGATAGCCATCGGGATCGTCACGGCGGTCTAGTAG
K T M E L K D V K P G Q G L *
- 4001 CAAACGGATATTTCAACGGGTGAACGACCCGGAGGACAGGACGATGATGCGGATTGACTCATTATCTTGTGTTGGGGGCGATATTGTTCTGTCACCGGCA
M I G L T H Y L V V G A I L F V T G I
- 4101 TTTTCGGGATCTTCTGTAACCGAAAGACGTATCGTCACTCTGATGTGATCGAACTGATGCTGCTGGCGGTGAACATCAACTCTGTCGCTTCTCGAC
F G I F V N R K N V I V I L M S I E L M L L A V N I N F V A F S T
- 4201 CCATCTGGGCGATCTGGCCGGGCGAGGTCTTACCATGTTCTGCTGACCGTCCGCCCGCGGAGGCCGCGCATCGGCTGGCGATCCTGGTGGTCTTCTTC
H L G D L A G Q V F T M F V L T V A A A E A A I G L A I L V V F F

NQ012

4301 CGCAACCGCGGCACCATCGCGGTGGAAGACGTCAACGTGATGAAGGGATAAGCGGCCATGGAAAAATTCGTCTCTTTGCGCGCTCATCGCGTCGCTGA
R N R G T I A V E D V N V M K G * M E K F V L F A P L I A S L I

4401 TCGCCGGGCTGGGCTGGCGGGCGATCGGCGAGAAGGCAGCGCAATACCTGACCACGGGCGTGCTGTTCTCAGCTGCCTGATCAGCTGGTACCTGTTTCT
A G L G W R A I G E K A A Q Y L T T G V L F L S C L I S W Y L F L

4501 GTCCTTCGACGCGTGCCGCGGCATATCCCGGTGCTCGACTGGGTGCTGACCGGCGATTTCATGCCGAATGGGCGATCCGGCTGGACCGGCTGACCGCG
S F D G V P R H I P V L D W V V T G D F H A E W A I R L D R L T A

4601 ATCATGCTGATCGTCGTGACACCGTTTCGGCGCTCGTGACATGTATTGCTGGGCTACATGGCCCATGACGACAACTGGACGCATGACGAGCACTACA
I M L I V V T T V S A L V H M Y S L G Y M A H D D N W T H D E H Y K

4701 AGGCCCGCTTCTTCGCTATCTCTCGTTCTTACACTTCGCGCATGTGATGCTGGTGACGGCGGACAACCTGTTGCAGATGTTCTTCGGCTGGGAGGCGGT
A R F F A Y L S F F T F A M L M L V T A D N L L Q M F F G W E G V

4801 GGGCGTCGCTCTACCTACTGATCGGATCTACTACAAGAAGGCCAGCGCCAATGCCGCGCGATGAAGCGGTTTCATCGTCAACCGGGTGGGCGACTTC
G V A S Y L L I G F Y Y K K A S A N A A A M K A F I V N R V G D F

4901 GGCTTCCTGCTGGGTATCTTCGGCATCTACTGGCTGACCGGCTCGGTCCAGTTCGACGAGATCTTCGCCAGGTGCCGAGCTCGCGCAGACCGAGATCG
G F L L G I F G I Y W L T G S V Q F D E I F R Q V P Q L A Q T E I D

5001 ATTTCTGTGGCGGACTGGAACGCCGCGAACCTGCTGGGTTTCTGCTGTTCTGTCGGCGCCATGGGCAAGTCGGCGCAGCTCTTGCTGCACACCTGGCT
F L W R D W N A A N L L G F L L F V G A M G K S A Q L L L H T W L

5101 GCCGGACGCGATGGAAGGCCGACCCCGGTTTCGGCGCTGATCCATGCCGCGACCATGGTGACGGCGGGCTGTTCTGGTCTGCCCATGTGCCCGCTT
P D A M E G P T P V S A L I H A A T M V T A G V F L V C R M S P L

5201 TACGAATTGCGCGCGATGCCAAGAAGTTCATCGTCATCATCGGCGCCACACCGCTTCTTCGCCGCGACCGTGGGCTGGTGCGAAGACATCAAGC
Y E F A P D A K N F I V I I G A T T A F F A A T V G L V Q N D I K R

5301 GCGTCATCGCTATTCGACCTGTTGCGAGTTGGGCTACATGTTCTGGCGGGCGGTGTGCGGCTCTATTGCGCGGCCATGTTCCACCTGCTGACGCACGC
V I A Y S T C S Q L G Y M F V A A G V G V Y S A A M F H L L T H A

5401 TTTCTTCAAGGCCATGCTGTTCTGGGCGCGGGCTCGGTGATCCATGCCATGCACCATGAACAGGACATGCGGAAGTATGGCGGGCTGCCGAAGAAGATC
F F K A M L F L G A G S V I H A M H H E Q D M R N Y G G L R K K I

5501 CCGCTGACCTTCTGGGCGATGATGATCGGCACCTTCGCCATCACCGGCGTGGGCATTCCGCTGACCCACCTGGGCTTTGCGGTTTCTGTCCAAGGACG
P L T F W A M M I G T F A I T G V G I P L T H L G F A G F L S K D A

5601 CGATCATCGAAAGCGCTATGCGGGCAGCGGCTATGCCTTCTGGCTGCTGGTCATCGCGGCTGTTTCAACAGCTTCTACAGCTGGCGGTGATCTTCTT
I I E S A Y A G S G Y A F W L L V I A A C F T S F Y S W R L I F L

5701 GACCTTCTACGGCAAGCGCGCGCGGACCATCACGCCATGATCATGCGCATGAAAGCCCGCGGTGATGACGATCCCGCTGGGCGTGCTGGCCATCGGC
T F Y G K P R G D H H A H D H A H E S P P V M T P L G V L A I G

5801 GCGGTCTTTGCCGCGATGGTCTGGTATGGTCCCTTCTTGGCGACCATCAAGGTCACGGAATATTTCCACATCGCCGCGCGCATCAGGAAGCGGCGG
A V F A G M V W Y G P F F G D H H K V T E Y F H I A G A H H E A A E

5901 AGGGCGAAGAAGCAGAGCACGCGACCGCGAGGCTCCGGTGAGCACGCGTCCGCGATACCGCGACCGCGAGGGCGAAGCGGCCCGGAGGCCGAACA
G E E A E H A T A E A P V E H A V A D T A T A E G E A A A E A E H

6001 TGGCGAGATCGCGCGCGGCTGGGCGGTGCGATCTACATGCATCCCGACAACCACATCATGGACGAGGCGCACCATGCGCGGCGCTGGGTGAAGGTCTCG
A E I A A P V G G A I Y M H P D N H I M D E A H H A P A W V K V S

6101 CCCTTCGTGCGCATGGTGTGGGCTGATCACCGCTGGACATTCTACATCGCCAACCCGTGCTGCCGCGGCGGCTGGCGGCGCACGAACCGGCGCTGT
P F V A M V L G L I T A W T F Y I A N P S L P R R L A A H E P A L Y

6201 ATCGGTTCTGCTCAACAAATGGTATTTGACGAGATCTACGAGTTCATCTTCGTCGCCCGCGGGAATGGCTTGGCGCGTGCTGTGGAAGGGCGGCGA
R F L L N K W Y F D E I Y E F I F V R P A K W L G R V L W K G G D

6301 CGGCGCGCTCATCGACGGCACCATCAACGGGGTGGCCATGGGCTTGATCCCGCGGCTGACCCGCGCGGCGTCCGGGTGCAATCGGGCTATCTGTTCCAC
G A V I D G T I N G V A M G L I P R L T R A A V R V Q S G Y L F H

NQ013

6401 TACGCCCTCGCGATGGTCTGGGCGATCGTGGGCTGCTGATCGGGTGATGATCGGGGCGCGCACTGACATGACGAACCTTCTTTCCATCATCACCTTC
Y A F A M V L G I V G L L I W V M M R G A H * M T N L L S I I T F

6501 CTGCCGATCGTCGCCGCATCATATGGCGCTGTTCTCGCGGCCAGGACGAGGCCGCGCGCAACGCCAAATGGCTGGCGCTGCTGACCACCACGG
 L P I V A A I I M A L F L R G Q D E A A A R N A K W L A L L T T T A
 6601 CGACCTTCGTGATCTCGCTTTTCGTGCTGTTCCGCTTCGACCCGGCGAAACACCGGCTTCAGTTCGTGAGGATCACGCCTGGATCATGGCGCTGCTGCTA
 T F V I S L F V L F R F D P A N T G F Q F V E D H A W I M G V C Y
 6701 CAAGATGGGCGTGGACGGCATCTCGGTGCTGTTCTGCTGCTGACCACTTCATGATGCCGCTGACCATCTTCGACCTGGCAGGTCCAGGACAAGGTC
 K M G V D G I S V L F V L L T T F M M P L T I L S T W Q V Q D K V
 6801 AAGGAATACATGATCGCCTTCCTGGTGTGGAAGGGCTGATGATCGGCGTCTTCACCGCGCTCGATCTGGTGTCTTCTACCTGTTCTTCGAGGCGGGGC
 K E Y M I A F L V L E G L M I G V F T A L D L V L F Y L F F E A G L
 6901 TGATCCGATGTTCTGATCATCGGCATCTGGGCGGCAAGACCGCATCTATGCGTCTTCAAGTTCTTCTCTACACCTTCCTCGGCTCGGTGCTGAT
 I P M F L I I G I W G G K D R I Y A S F K F F L Y T F L G S V L M
 7001 GCTGGTGGCGATGATCGCCATGTATCGCATGGCGGGCACCACCGACATCCGACGCTGCTGACCTTCGACTCCCGTCCGAGAACTTCGCTCTTGGG
 L V A M I A M Y R M A G T T D I P T L L T F D F P S E N F R L L G
 7101 ATGACCGTGGTGGCGGCATGAGATGCTGCTGTTCTGGCCTTCTTGGCAGCTTCGCGGTCAAGATGCCGATGTGGCCGGTCCATACCTGGCTGCCCC
 M T V V G G M Q M L L F L A F F A S F A V K M P M W P V H T W L P D
 7201 ACGCGCACGTCCAGGCGCGGACCGCGGCTCGGTCTGCTGGCGCGGTGCTGCTGAAGATGGCGGGTACGGCTTCCTGCGCTTCTCGTGCCTGATGTT
 A H V Q A P T A G S V L L A A V L L K M G G Y G F L R F S L P M F
 7301 CCCGGTCCGCGAGCGCGTGGCCAGCCCTATGTCTTCTGGTGTGGCCATCGCCATCGTCTACACTTCGCTGGTGGCGCTGGCGCAGTCCGACATGAAG
 P V A S G V A Q P Y V F W L S A I A I V Y T S L V A L A Q S D M K
 7401 AAGGTCATCGCCTATTCTCGGTGCGCCATATGGGCTACGTGACCATGGCGCTTTTCCCGCAACCAGATCGGCGTCCGCGCGCATCTCCAGATGC
 K V I A Y S S V A H M G Y V T M G V F A A N Q I G V D G A I F Q M L
 7501 TGTGCGACGGCTTTATCTCGGGCGCGCTGTTCTGTGCGTGGCGGTGATCTACGACCGCATGCACACGCGGAGATCGACCGCTATGGCGGGTTGGTGA
 S H G F I S G A L F L C V G V I Y D R M H T R E I D A Y G G L V N
 7601 CCGGATGCCGGCTATGCGCGGTGTTCACTGTTCTTACCATGGCGAAGTGGGCGTCCGCGGACCTCGGGTTTCGTGGCGAGTTCCTGACGCTGATG
 R M P A Y A A V F M F F T M A N V G L P G T S G F V G E F L T L M
 7701 GGGGTCTTCCGCGTCGATACCTGGGTGGCGCTGGTGGCGACCTCGGGGTGATCCTGTGCGCGGCTATGCGCTCTGGCTGTATCGCGGGTGACGCTGG
 G V F R V D T W V A L V A T S G V I L S A A Y A L W L Y R R V T L G
 7801 GCCAGCTCATCAAGGAAAGCCTCAAGTCGATCACCGACATGACCCCGCGCAACGCTGGGTGTTTCATCCCGCTGATCGCCATGACGCTGATCCTGGGCGT
 Q L I K E S L K S I T D M T P R E R W V F I P L I A M T L I L G V
 7901 CTATCCGCGGCTCGTCACCGACGTGACCGGCCCGCGGTGCGAGCACTCGTGAAGACTACAACCAGAGCCAGCCGGCGGCGACCCGTGCGCACGGCCCAA
 Y P R L V T D V T G P A V A A L V Q D Y N Q S Q P A A P V A T A Q
 8001 GCCAGCCATAGGGGACGGACATGACCTCGCTCGATTTCTCGACATTCTGCCGAGGTGGTGTGGCGGTTACGCCCTTGCCGCGCTGATGGCGGGG
 A S H * M T S L D F S T I L P E V V L A G Y A L A A L M A G A
 8101 CCTATCTCGGCAAGGACAGGCTGGCCCGACGCTGCTTTGGGTACGGTCCGGGCTTCTGGTGTGCGCGCATGGTGGGCTCGGCAACCATGTGGA
 Y L G K D R L A R T L L W V T V A A F L V V A A M V G L G N H V D
 8201 CGGCGCCGCTTCCACGGCATGTTTCATCGACGACGGCTTCTCGCGCTTTGCCAAGGTGGTGACGCTGGTCCGCGCGCGCGCTGCTGGCGATGAGCGCC
 G A A F H G M F I D D G F S R F A K V V T L V A A A G V L A M S A
 8301 GATTACATGACGCGCGCAACATGCTGCGCTTCAGTTCCCGATCATGTTGCACTGGCGGTGCTGGCATGATGTTCACTGGTCTCGGCGCGCGACCTGC
 D Y M Q R R N M L R F E F P I I V A L A V L G M M F M V S A G D L L
 8401 TGACGCTCTACATGGGCTGGAGTGAATCGCTGGCGCTTACGTCGTGCGCGCATGGCGCGGACTCGGTGCGCTCGTCCGAGGCGGGGTGAAATA
 T L Y M G L E L Q S L A L Y V V A A M R R D S V R S S E A G L K Y
 8501 CTTCTGCTGGGATCGCTTTCTCGGGCTGCTGCTCTATGGCGCTCGTGGTCTACGGTTTTGCCGGCACCGGGTTCGAGGGCATCATCTCGACC
 F V L G S L S S G L L L Y G A S L V Y G F A G T T G F E G I I S T
 8601 ATCGAGGCCGGCCACCTGTCCCTGGGCGTGTGTTGGCGCTGGTGTCACTGGTGGGCTGCTGTTCAAGGTTTCGGCCGTGCCCTTCACATGTGGA
 I E A G H L S L G V L F G L V F M L V G L S F K V S A V P F H M W T

```

8701  CGCCCCAGCTCTACGAGGGCTCGCCGACGCGGTACCGCTTCTTCGCCACCGCGCCCAAGGTCGGGCGATGGCGCTGATCGCGCGGTGGTCTTCGA
      P D V Y E G S P T P V T A F F A T A P K V A A M A L I A R L V F D

8801  CGCCTTCGGCCATGTCATCGGGACTGGAGCCAGATCGTCGGGCGCTGGCGGTGATGTCGATGTTCTGGGCTCGATCGCCGGGATCGGCCAGACCAAC
      A F G H V I G D W S Q I V A A L A V M S M F L G S I A G I G Q T N

8901  ATCAAGCGGCTGATGGCCTATTCTCGATCGCGCATATGGGCTTTGCCCTTGTGGGCTTGGCCCGGCACCGCCATCGGCGTGCAGAACATGCTGCTTT
      I K R L M A Y S S I A H M G F A L V G L A A G T A I G V Q N M L L Y

9001  ACATGACCATCTATGCGGTGATGAACATCGGCACCTTCGCTTATCCTGTGATGGAGCGCGACGGCGTGCCGGTGACCGACCTTGGCGCGTGAACCG
      M T I Y A V M N I G T F A F I L S M E R D G V P V T D L A A L N R

9101  CTTTGCCTGGACCGATCCGGTCAAGGCGCTGGCGATGCTGGTCTGATGTTACGCTTGCAGGCGTGCCGCGACGCTGGGCTTCTTTGCCAAGTTCGGC
      F A W T D P V K A L A M L V L M F S L A G V P P T L G F F A K F G

9201  GTGCTGACGGCGCGGTGATGCGGATGGGCTGGCTGGCGGTGCTGGGCGTGATCGCTCGGTGATCGCGCGCTTCTATTACCTGCGCATCGTCTATT
      V L T A A V D A G M G W L A V L G V I A S V I G A F Y Y L R I V Y Y

9301  ACATGTATTTCGGCGGCGAGTCCGAGGGCATGACCTCGCGCATGGGCGCGGTGCAATACCTGGCGGTGATGGTGCCGGCGCTGGCGATGCTGGTGGGTGC
      M Y F G G E S E G M T S R M G A V Q Y L A L M V P A L A M L V G A

9401  GATCAGCATGTTTCGGCGTCAATTCCGCGCGCGGCGCGCGCGGAGACCTGGTGGTTCGGTGGCGGCCATCGAGCAGCCCGCGAAGCCGCGCAAGCC
      I S M F G V D S A A G R A A E T L V G P V A A I E Q P A E A A Q A

9501  GAGCCCGTCAAGGTGAGTGACGCTGGCCCGAGGGCGTGGCCCGCATGTGCTGGCCCGCACCGACAGCACCAATGCCGAGGCGCTGAAGCTGGCGGCC
      E P V Q G E *                               URF240
                                              V L A R T D S T N A E A L K L A P

9601  GGCCTTTCCGGTTCGCTGGGCTGGTGGCGCGAGCAATTCGGGGCGGTGGCGCTCGCGGGCGCAATGGGTGATGCCGGCGGAACTTTGCCGGCA
      G L S G S A W V L A R E Q F A G R G R R G R E W V M P A G N F A G T

9701  CGCTGGTTCGCGGCCCCAGGGAGGGGCGCTGGCGGGCGGCAACTGTCTTCGTCGCCGCGCTGGCGCTCTACGATGCGCTTGGTCTGGCCTCGCGGCC
      L V L R P Q G G A L A A A Q L S F V A A L A L Y D A L G L A C G P

9801  GGCCGCGCGGTGGCGATCAATGGCCGAACGACGTGCTGCTGAACGGCGGCAAGGTGGCGGGCATCTGCTGGAAGCTCGGGCTCCGGCCCCGGCGTG
      A A R L A I K W P N D V L L N G G K V A G I L L E S S G S G P G V

9901  CAGGCGCTCGGGTCGGCATCGGCGTCAACTGGCCGGCGCGCCGATGCGCGCGCGGTGAGCCCGGTGCCACCCCGCGGTTTCCGTGCAAGGCGAGA
      Q A V A V G I G V N L A G A P D A G A V E P G A T P P V S V Q G E T

10001 CTGGCCACGCGCTCGATCCCGAGGAATCTTGGACCTGCTGGCGCGGCTTTGCACGCTGGCAGGCGCAACTCGACACCTATGGCTTCGCGCGATCCG
      G H A V D P E E F L D L L A P A F A R W Q A Q L D T Y G F A P I R

10101 CAATGCCTGGTGGCCCGCGCGGCGGCTGGGCGAGCGGATCATCGCCGCGACCGGCACTGCCGAAAGCCACGGCATCTTCGAGGGCATCGACGACAGC
      N A W L A R A A R L G E P I I A R T G T A E S H G I F E G I D D S

10201 GGCGCGCTGATCTGCGCGGCGCGGCGGGCGGCGAGGTCAITCCCGCGCGCAAGTCTTTTCGGAGGCTAGGCCATGCTGCTTTGCATCGACACCGGCA
      G A L I L R G P A G R Q V I P A A E V F F G G *

10301 ACACCAACACGGTCTTTTCGTCTGGGACGGCGAGAAATTCGTGGGCTCTGGCGCATCTCGACCGACCACCGGCGCACGGCGACGAGTATTCGTCTG
      G

10401 GCTATCCACGCTGATTCCCTGCAG 10425

```

FIGURE 2: Nucleotide and predicted amino acid sequences of URF4, *NQO8*, *NQO9*, URF5, URF6, *NQO10–14*, and URF240. The putative Shine–Dalgarno sequences are doubly overlined. The putative terminator sequences are marked by dashed lines.

the *Paracoccus NQO8–14* gene products were used to search the GenBank/EMBL database. Significant similarities were found between these gene products and the ND1, 23-kDa (homologue of *ndhI* of chloroplast DNA and cyanobacteria), ND6, ND4L, ND5, ND4, and ND2 polypeptides of bovine complex I (Anderson et al., 1982). The fact that these structural genes encoding the homologues of the seven ND products and the HP 23-kDa polypeptide are located in the gene cluster encoding the *Paracoccus* NDH-1 strongly support both the hypothesis of Chomyn et al. (1985b, 1986) that the

seven ND products are subunits of complex I and also the speculation that the 23-kDa polypeptide is a subunit of the mammalian NADH dehydrogenase complex (Dupuis et al., 1991).

Hydropathy plots of the polypeptides encoded by the *NQO7*, *NQO8*, and *NQO10–14* genes show that these polypeptides contain membrane-spanning regions similar to their bovine counterparts (data not shown). In addition, although the hydropathy plot of the *NQO9* polypeptide does not appear to contain any membrane-spanning regions, since the bovine 23-

Table II: Comparison Matrices for the *NQO8*, *NQO9*, *NQO10*, *NQO11*, *NQO12*, *NQO13*, and *NQO14* Subunits of the Energy-Transducing NADH-Quinone Oxidoreductase from Various Species

subunit	source	% identity						
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>NQO8</i>	(1) <i>P. denitrificans</i>	100	77.1	38.7	47.9	35.6	33.8	31.5
	(2) <i>R. capsulatus</i>		100	40.6	52.4	37.1	34.8	32.5
	(3) bovine mit ^a			100	48.4	34.3	35.2	32.4
	(4) liverwort mit				100	36.9	39.0	39.3
	(5) <i>Synechocystis</i>					100	59.8	56.6
	(6) liverwort chl ^b						100	69.3
	(7) rice chl							100
<i>NQO9</i>	(1) <i>P. denitrificans</i>	100	88.7	72.0	27.4	28.1	28.1	
	(2) <i>R. capsulatus</i>		100	72.4	27.6	27.6	28.2	
	(3) bovine mit			100	25.6	26.7	30.0	
	(4) <i>Synechocystis</i>				100	66.7	57.3	
	(5) liverwort chl					100	73.0	
	(6) rice chl						100	
	(7) rice chl							100
<i>NQO10</i>	(1) <i>P. denitrificans</i>	100	13.7	42.2	24.4	23.6	19.9	
	(2) bovine mit		100	13.1	12.0	13.1	11.4	
	(3) liverwort mit			100	21.6	27.8	24.4	
	(4) <i>Synechocystis</i>				100	41.4	41.5	
	(5) liverwort chl					100	55.7	
	(6) rice chl						100	
	(7) rice chl							100
<i>NQO11</i>	(1) <i>P. denitrificans</i>	100	23.5	66.0	47.1	43.0	37.6	
	(2) bovine mit		100	22.5	21.4	21.4	23.5	
	(3) liverwort mit			100	48.0	38.0	38.0	
	(4) <i>Synechocystis</i>				100	66.0	52.5	
	(5) liverwort chl					100	66.0	
	(6) rice chl						100	
	(7) rice chl							100
<i>NQO12</i>	(1) <i>P. denitrificans</i>	100	32.5	41.4	32.1	33.2		
	(2) bovine mit		100	32.7	28.4	27.4		
	(3) liverwort mit			100	31.1	31.8		
	(4) liverwort chl				100	61.6		
	(5) rice chl					100		
	(6) rice chl						100	
	(7) rice chl							100
<i>NQO13</i>	(1) <i>P. denitrificans</i>	100	29.2	47.7	34.1	32.5	32.0	
	(2) bovine mit		100	29.4	26.4	25.9	25.7	
	(3) liverwort mit			100	33.1	31.5	30.3	
	(4) <i>Synechocystis</i>				100	58.7	52.6	
	(5) liverwort chl					100	71.3	
	(6) rice chl						100	
	(7) rice chl							100
<i>NQO14</i>	(1) <i>P. denitrificans</i>	100	23.6	36.0	31.8	29.4	28.6	
	(2) bovine mit		100	20.5	22.5	24.5	22.8	
	(3) liverwort mit			100	28.6	30.5	29.5	
	(4) <i>Synechococcus</i>				100	50.3	47.3	
	(5) liverwort chl					100	66.5	
	(6) rice chl						100	
	(7) rice chl							100

^a mit, mitochondrial subunit. ^b chl, chloroplast subunit.

kDa polypeptide is present in the HP fraction (Dupuis et al., 1991) it seems possible that its homologue, the *NQO9* polypeptide, may also be located in the membrane sector. Additional support for the hypothesis that the *NQO7-14*-encoded polypeptides constitute the membrane sector of the *Paracoccus* NDH-1 is found in the discovery that when *Neurospora crassa* was grown in the presence of chloramphenicol, an inhibitor of mitochondrial protein synthesis, the large membrane-embedded domain, the rotenone sensitivity, and the high-potential FeS cluster N2 signal of complex I were all lost (Tuschen et al., 1990; Weiss et al., 1991). Taken together, these data suggest that the polypeptides encoded by *NQO7-14* constitute the membrane sector of the *Paracoccus* NDH-1.

Kikuno and Miyata (1985), on the basis of amino acid sequence homologies among the ND gene products, have hypothesized that the ND2, ND4, and ND5 genes evolved from a single ancestral gene by a series of gene duplications. The amino acid sequence similarities among these *Paracoccus* NDH-1 subunits encoded by the *NQO12*, *NQO13*, and *NQO14* genes have been analyzed by means of dot matrices (data not shown). These analyses showed similarity among these

Paracoccus NDH-1 subunits and appear to lend support to the hypothesis of Kikuno and Miyata. If their hypothesis concerning the evolutionary relationship among the ND2, ND4, and ND5 genes is correct, these results suggest that the gene duplications might have occurred before the endosymbiosis. Nevertheless, we must also consider the possibility that the observed homology among the ND2, ND4, and ND5 products reflect only the inclusion of those structural motifs essential for the functioning of NDH-1 (e.g., FeS cluster binding site) and not their evolutionary relationship. Clarification of this issue may require the identification of the ancestral genes of ND2, ND4, and ND5.

A comparison of the amino acid sequences of the *Paracoccus* subunits with those of their mitochondrial, chloroplast, and bacterial counterparts is illustrated in Table II. Although only two structural genes encoding *Rhodobacter capsulatus* NDH-1 have been determined at the present time (Dupuis, 1992), the *Paracoccus* *NQO8* and *NQO9* gene products have significant similarities to their *R. capsulatus* counterparts. These results support the hypothesis that *R. capsulatus* may be closely related, evolutionally, to *P. denitrificans* (Xu et al., 1991b; Stouthamer, 1992). With respect to the *NQO8* and

NQO10–14 products, the *Paracoccus* NDH-1 subunits have greater sequence similarity to their liverwort mitochondrial counterparts than to their mammalian mitochondrial and chloroplast homologues. Furthermore, a similar trend has been observed for the *Paracoccus NQO7* product (Xu et al., 1992a). In light of these results, the *Paracoccus* NDH-1 may have a closer evolutionary relationship to plant mitochondrial complex I than to mammalian mitochondrial complex I. This appears to be particularly true for the hydrophobic subunits. Unfortunately, at the present time, liverwort is the only plant species for which the complete ND gene sequence is available. In order to justify this speculation, it will be necessary to accumulate complete ND gene sequences from several other plant species.

In a previous report we compared the primary structures of the *NQO4*, *NQO5*, and *NQO6* products from the respiratory NDH-1 complexes to those of the putative chloroplast NDH-1-type enzymes (Xu et al., 1992a). In the present study we have conducted a similar comparison for the *NQO8–14* gene products and their homologues in both the respiratory and chloroplast systems. As seen in Figure 3, unique differences have been detected in individual amino acid residues of the *NQO8–14* homologues of the respiratory NDH-1 and the putative NDH-1 of chloroplasts. In addition, among the *NQO9* homologues, the mitochondrial subunits contain three stretches of amino acids ($M_{14}WDFI_{18}$, $E_{51}HALR_{55}$, and $L_{141-x}NG_{144}$) which are absent in the chloroplast subunits. On the other hand, the chloroplast *NQO9* homologues bear a stretch of approximately 20 amino acid residues at the C-terminal region which is not found in the respiratory NDH-1 (or complex I) subunits. Furthermore, the *NQO12* homologues of the respiratory NDH-1 lack the sequences (xxQOI and EPxVGyXP) which are found in the chloroplast subunits. Interestingly, the subunits of the cyanobacterium *Synechocystis* NDH-1 type enzyme appear to be considerably more similar to their chloroplast NDH-1-type enzyme homologues than to their respiratory NDH-1 counterparts, at least in terms of the *NQO8–11* and *NQO13* homologues. Similar trends were also observed in the *NQO4*, *NQO5*, and *NQO6* homologues (Xu et al., 1992a). These data suggest that the *Synechocystis* NDH-1-type enzyme is more closely related to the putative NDH-1-type enzymes of chloroplasts than to the respiratory NDH-1.

Iron–Sulfur Clusters. Analysis of the *NQO8–14* sequences revealed that only the *NQO9* gene product contains the two structural motifs typically associated with the tetranuclear iron–sulfur clusters found in ferredoxins of either bacteria or chloroplasts (Yasunobu & Tanaka, 1980). These are the sequences $C_{64}xxC_{67}xxC_{70}xxxC_{74}P_{75}$ and $C_{103}xxC_{106}xxC_{109}xxxP_{114}$. According to the available literature (Hatefi et al., 1985; Ohnishi et al., 1985), the bovine complex I HP is believed to contain one or two binuclear and one tetranuclear FeS cluster. The tetranuclear cluster has been provisionally designated as cluster N2 as this component exhibits a phospholipid-dependent midpoint potential and is the likely electron donor to quinone in the lipid-phase membrane. As described above, the bovine 23-kDa polypeptide, which is homologous to the *Paracoccus NQO9* polypeptide, is present in HP. Taken together, these data suggest not only that one of the putative iron–sulfur binding sites in the bovine complex I may involve the 23-kDa polypeptide but also that this may be the binding site of the high-potential cluster N2. A high-potential cluster resembling N2 has also been detected in the *Paracoccus* NDH-1 (Meinhardt et al., 1987). These data suggest that its binding site may involve the *NQO9* gene product. This hypothesis is supported by the finding that

other polypeptides composing the membrane sector of the *Paracoccus* NDH-1 do not bear any putative FeS cluster binding sites. However, this assignment must be verified by direct experiment.

Characterization of the URF4, URF5, and URF6 Polypeptides. No significant homology was found between the URF4 and any of the sequences in the GenBank/EMBL database. Nor was any significant relationship found between URF4 and the known amino acid sequences of any of the other NADH–UQ oxidoreductase proteins. Assuming that the putative polypeptide encoded by URF4 is synthesized, it is a basic polypeptide (see Table I) and contains a hydrophobic stretch long enough to span the membrane.

On the basis of the hydropathy plots of their respective polypeptides, URFs 5 and 6 appear to code for peripheral proteins. As shown in Figure 4, the putative polypeptides encoded by URFs 5 and 6 are similar to each other, especially in the region of residues 50–120. This region is also similar to the primary structure of the γ -carboxymuconolactone decarboxylase of *Acinetobacter calcoaceticus* (Hartnett et al., 1990; Yeh et al., 1980; Parke, 1979). This enzyme is involved in the metabolism of β -ketoadipate and catalyzes the decarboxylation of γ -carboxymuconolactone to produce β -ketoadipate enol lactone (Hartnett et al., 1990; Yeh et al., 1980; Parke, 1979). Both quinone and lactone have been reported to induce expression of the FMN-dependent NAD(P)H-diaphorase in *E. coli* (Hayashi et al., 1990; Unemoto et al., 1992). It could be speculated, therefore, that the URF5 and URF6 products might be related to the regulation of NDH-1 biosynthesis.

Gene Arrangement. As shown in Figure 1, the *Paracoccus* NDH-1 structural genes encoding homologues of the bovine IP polypeptides are arranged in such a way that they surround the structural genes encoding the homologues of the bovine FP polypeptides. On the other hand, the structural genes encoding the *Paracoccus* NDH-1 homologues of the bovine HP polypeptides constitute a single cluster. The one exception is the *NQO7* gene. Although the *NQO7* gene encodes the *Paracoccus* counterpart of a bovine HP subunit, this gene is located at the 5'-terminus of the *Paracoccus* NDH-1 gene cluster rather than with the other HP-like subunit genes. As described previously (Xu et al., 1992a), the arrangement of the *NQO7*, *NQO6*, and *NQO5* genes in *Paracoccus* is the same as that detected in *Synechocystis* and chloroplast DNA (Ohya et al., 1986; Steinmüller et al., 1989). In this study, the ordering of the *NQO8–11* genes was also identical to that found in chloroplast and cyanobacterial DNA (Ohya et al., 1986; Takahashi et al., 1991). Recently, the complete DNA sequence of liverwort mitochondria was published (Oda et al., 1992). We compared the gene arrangements of the liverwort mitochondrial DNA with that of the *Paracoccus* NDH-1 gene cluster and found that the arrangement of the *NQO12*, *NQO13*, and *NQO14* genes is identical in the liverwort mitochondrial DNA and the *Paracoccus* gene cluster (see Figure 1). Preliminary results of DNA sequencing of the gene cluster encoding the *E. coli* and *R. capsulatus* NDH-1 also suggest that their gene arrangements are similar to that of *Paracoccus* NDH-1 (Weidner et al., 1992; Dupuis, 1992). Such similarities will provide useful information for investigating the evolutionary relationship of NDH-1, the mitochondrial complex I, and the putative chloroplast complex I type enzyme.

***Paracoccus* NDH-1 as a Model for the Study of Mitochondrial Diseases.** Human mitochondrial DNA is composed of 16.5 kbp and encodes 13 polypeptides. Seven of these genes (designated as ND1, 2, 3, 4, 4L, 5, and 6) code for polypeptides

NQO8 subunit

P.denitrMAEFWASPYGFALSMLLQGLAVIAFVMSGLIFMVYGDRIWAAVQMRGPNVVGWGLLQTFADALKYIVKEIVIPAGADKFVYFLAPFLSMMLALFAFVVIPIF
 R.capsulMADFWATSLGQTLILLAQGLGIIAFVMIIGLLLVWGDRIWAAVQMRGPNVVGAFGLLQSVADAAYVFEIVVPAGVDKPVYFLAPFLSLVLALLAWVVPIF
 BovineMFMINI.....LMLIIPILLAVAFITLVERKVLGYMQLRKGPVNVGYPYGLLQPIADAIFKEIPLPATSSSMFIFLAPINALGALATMHIPLPM
 Live mitMRIYILGLVAKMLGIMMPLLLGVAFILVAERKVMASHQRKGPVNVGILGLLQPLADGLKLMKKEPILPSSANIFIFLMAFVLTFTLALCAWAVIPIF
 SynechocMTSGIDLQNSFLQSLQGFGLPPLAKLFWIPLPSILMIIGATVGVLVVWLERKISAAQQRIPEYAGPLGLVLPQVADGILKLVFKEDVVPKADPWLFTLGPVLVLPVFLSYLIVPIF
 Live chlMISINILEDKFFSFFTLGFSKEFFNFWIIFSLIILMLGVITGLVLVWLERKISAAIQRIPEYAGPLGIQALADGILKLVKEDIVPAQGDVWLFNIGPLVLVLPVFLSYLIVPIF
 Rice chlMIIDRVQVEA.....INSFNLLELLKEVYGLIW.ILPILTLGLGITIEVLIVWLEREISASIQRIPEYAGPLGLLQAIADGTLKLVKEDILPSRGDPLFSIGPSTAVISILLSFVLPIF

P.denitrDEGWVMANINVGILFIFAASSLEVYGVIMGGWASNSKYPFLASLSRAAQMSIEVSLGLIIGIISTGSMNLTAIVEAHGGDYGLLNWYWLPHLPMVVLFFVSALAEENRPPFDLVEAE
 R.capsulNEGWMADINVAVLFFVAVSLLEVYGVIMGGWASNSKYPFLGSLRSAAQMSIEVSMGLIIVGVIIISTGSMNLTAIVEAQRGDGLLNWYWLPHLPMVVLFFISALAEENRPPFDLPEAE
 BovinePYPLI.....NMNLGVLFLMAMSLAVSILWSGWASNSKYALIGALRAVAQTISEVTLAIILLVLLMSGSFTLSTLITTEQ.....MWLILPAWPLAMMWFISTLAETNRAPFDLPEAE
 Live mitDYGWVFSQNLIGVLYLFAMSSLGVIYITAGWASNSKYAFLGALRSAAQMSIEVSMGLIIMVLCAGSCNFSEIVLAQTR.....MWVFPFLFPVFLMFFISCLAETNRAPFDLPEAE
 SynechocGQNLVITDINVGIFLWIALSSIAPIGLMSGYASNNKYSLLGLLRAAQAISIEIPLSLAVLAIVMMSNSLSTIDIVDQSG.YGILGWNWRQPVGFLIFWIAALAECEERLPDLPEAE
 Live chlEYNVILANFSIGVFFWIAVSSVPLGLLMAGYGSNNKYSFLGLLRAAQAISIEIPLALSVLSIALLSNSLSTVDIVDAQSK.YGFLSWNLWRQPIGFIFFIASLAECERLPDLPEAE
 Rice chlGYRFLVADLSIGVFLWIAISSIAPIGLLMAGYGSNNKYSFSGGLRAAQAISIEIPLTCVLAISLLSNSSTVDIVDAQSK.YGFFGWNWRQPIGFIFFIASLAECERLPDLPEAE

P.denitrSELVAGFMTEYSSTPYLLFMAGEYIAMYLMLCALLSLLFFGGW.....LS.....PVPFIADGWWMVIKMWFYFAMVKAIVPRYRYDQMLRIGWKVFLPLSLGWVV
 R.capsulSELVAGFMVEYSSTPYLLFMAGEYIAVLMCALTSVLLFFGGW.....LS.....PIPGVPDGLVLMVAKMAAVFVFMVKAIVPRYRYDQMLRIGWKVFLPLSLAWVV
 BovineSELVSGFNVEYAAGPFALFMAEYANIMMNIITAILFLGTS.....HNPHMPELY.....TINFITKSLLLTMSFLWIRASYPRFRYDQMLHLLWKNFLPLTLALCM
 Live mitAELVAGYVVEYSSMGFALFFLGEYANMLMSLCTLLFLGGW.....LPILDPIFQVPIGSIWFSMKVLFLLFVVMVRAAFPRYRYDQMLRIGWKVFLPLSLAWVV
 SynechocEELVAGYQTEYSGMKFAFFYLASYLNLVSSLVFTIYLGLGWHSFIPFSLFKNFENLMSNGISEVSIITIGIVITLVKSYLFLFISIMTRWTLPRIRIDQLNLGWKFLPLALGNLL
 Live chlEELVAGYQTEYSGMKFAFFYLASYLNLVSSLVFTIYLGLGWHSFIPFSLFKNFENLMSNGISEVSIITIGIVITLVKSYLFLFISIMTRWTLPRIRIDQLNLGWKFLPLALGNLL
 Rice chlEELVAGYQTEYSGIKYGLFYLVSYLNLVSSLVFTIYLGLGWHSFIPFSLFKNFENLMSNGISEVSIITIGIVITLVKSYLFLFISIMTRWTLPRIRIDQLNLGWKFLPLALGNLL

P.denitrLVAILARYEILGGFWARFAVGG
 R.capsulVVAFLAKFEVLGGFWARWSIGA
 BovineWHVSLP.....ILTSGIPPTQ...
 Live mitFVSGVL.....VAFEWLP.....
 SynechocITAALKLTFP.....
 Live chlLTTSFQLFLL.....
 Rice chlLTTSSQLVSL.....

NQO9 subunit

P.denitrMAFDARATKYFLMWDIFKGFGLGMYFVSPKPTLNYPHEKGPLSPRFRGEHALRRYPNGEERCIACKLCEAVCPAQAITIDAERREDGSRRT..TRYDIDMTKCIYC
 R.capsulMAFDYVRAAKYFVLWDFIKGFALGMYFVAPKPTLNYPHEKGPLSPRFRGEHALRRYPNGEERCIACKLCEAICPAQAITIDAEPDRDGSRT..TRYDIDMTKCIYC
 BovineTYKYVNLREPSMDMKSVTDRAAQTLLWTELIRGLMTLSYLFREPATINYPFEKGPLSPRFRGEHALRRYPNGEERCIACKLCEAVCPAQAITIEAGPRADGSRT..TRYDIDMTKCIYC
 SynechocMFNNILKQVGDYAKESLQAAYIG.....QGLAVTFDHMSRRPITVQYPIEKLIPSERFRG.....RIHFEFDKCIACEVCVRVCPINLPVVDWELFNKAVKKELKHYSIDFGVCIFC
 Live chlMFSIINGLNKNYQQAIAAARYIG.....QGLVTLDMHNRPLTTIYQPIEKLIPSERFRG.....RIHFEFDKCIACEVCVRVCPINLPVVDWELFKKTKKKQLKNYSIDFGVCIFC
 Rice chlMFPMTGFM.....GQQTIRAAARYIG.....QSFITLSTNRLPITIHYPYEKISTSERFRG.....RIHFEFDKCIACEVCVRVCPIDLPVLDWRFEKDIKRKQLLNYSIDFGVCIFC

P.denitrGFCQEAACPDAIVEGPNFEYATETREELFYDKQKLLANGERWEAEIARNLQLDAPYR.....
 R.capsulGYCQEAACPDAIVEGPNFEYATETREELFYTKELLENARWEAEIARNIEMDAPYR.....
 BovineGFCQEAACPDAIVEGPNFEFSTHEELLYNKEKLLNNGDKWEAEIAANIADYLYR.....
 SynechocGNCVEYCPNTCLSMTEEYELAAAYDRHLDNYDNVAL.....GRLPYKVTEDPHVTPRELGYLPKGVIENHNLPGKSGRAGQHPEDLVKAE
 Live chlGNCVEYCPNTCLSMTEEYELSTYRHELNQDIAL.....GRLPISIIEDSTIENIFNLTSLPKGKIEGHIYSRNITNIVN.....
 Rice chlGNCVEYCPNTCLSMTEEYELSTYRHELNQDIAL.....SRLPISIMGDTITQIRNST.QSKIDEESKWSNRITIDY.....

NQO10 subunit

P.denitrMMTFAFYLFASACVAGFMVIGRNPVHSLVLIILAFLSAAGLFLVQGAEFVAMLLVVVYVGAVALFLFVMMLDVDFAEKLGELARYLPLALVIGVLLLAQLGIAFSGWTPS
 BovineMMLYIVFVILSVIFVMGVFGFSKSPSPYIGGLGLIVSGGVGCGIIVLNFSGSFLGLMVFLIYLGMMHVVFGYTTAMATEQYPIEWSNKAVALGAFVTLGLMEFFMVVYVLKDKVE
 Live mitMMLFYVVLALVSGAMVMRAKNPVHSLVLIILVFCNTSGLLVLLGLDFAMIFLVVYVGAVALFLFVMMHLMRMEIEHENVRLYLPVGGIIGLIFLLEIFLMVDNDYIP
 SynechocVNLAEQVAYISFLILAFVIGAAAGVLLSNIVYSAFLGGVFLSTSGIYILLNADFVAAAQVLYVYVGAVALFLFAVMLN...KKQYSNFFVYWTIGDGTITLCTSIPLLLNFI
 Live chlMKLPESFYETIFLFLESGLITLGSGLVILLTNIVYSALFLGFVFCISLLYLLNADFVAAAQVLYVYVGAVALFLFAVMLN...KKQYSNFFVYWTIGDGTITLCTSIPLLLNFI
 Rice chlMDLPGPIHEILVLFGGVFLVLLGLGLVLLTNPTFSFSLGLVLVCSILFYLLNSYFVAAQVLYVYVGAVALFLFAVMFVN...GSEWSKDKNFWTIGDGTISLVCITIPFSMLTITPD

P.denitrDQAESLRAAPVDAV.....ENTLGLGLVLYDRYVLMFLQAGLVLVAMIGAILVLMRHRDKVRQNVLEQMWDRPAKTMELKDVKPGQGL
 BovineVVFENGLGDWVIYDTGDSGFFSEEAAGIAALYSYGTWLV...VTGSHLLIGVVIMEITRGN.....
 Live mitMLPTKLSATYLTYYVAGKMSHTNLETGLNLLYTYFFFLVSSILLVALIGAMVLTMHKTKVKRQDVFQNMAMDQONT...IKYVR.....
 SynechocTPW.....QINETGPFVVENTLVITGKHFFSDYLLPFELASVLLMAMVGAIIILARPGLNSRIVRGKQNRHGPDFARASPGVNLRFQIDA
 Live chlTSWSKIFLMTKPNLVKDIILI.NTVRHIGSELLTEFLPLPESLILLVALIGAILARREKKEIELEKNDFFNF.....
 Rice chlTSWYGILWTRSNQIVEQ.GLI.NNVQGIHGLATDYPFLPESLILLVSLIGAITMARQ.....

NQO11 subunit

P.denitrMIGLTHYLVVGAITLVFTGIFGIVNRKNVILMSIELMLLAVNINFAVSTHLG..DLAQGVFTMFVLTVAAAEAAIGAILVVFRRNGTIAVEDVNMKG
 BovineMSMVYNNIMMAFTVSLVGLMYRSHLMSLLCLEGMM..LSLFVMAALTILNSHTFLASHMPPILLVFAACEAALGLSLLVMVSNTYGTDDYQNLNLLQC
 Live mitMDLVKYLTFSMILFLGIGIFLNRKNILMLMSIELMLLAVNINFLVFSVYLD...DMGQLFALFVLTVAAAEAAIGAILVITFRGTIAVEFICNCKG
 SynechocMHLQYQCLILAAALFCIGIYGLITSR.NAVRVLMSIELMLLAVNINFLMGSNYLDPSNIRGQVFAIFVITIAAAEAAGVLAIVLAIYRNRETTDEMEQFNLLKW
 Live chlMLEHILTSALFLFCIGVFLITSR.NMVRALMCLLEIFNAVINLVAFSNFDSSQIKGEIIFSIIFIAIAAAEATIGLAIVLAIYRNKRSTRIDQFNLLKW
 Rice chlMMFEHVLFLSVYLSIGIYGLITSR.NMVRALICLELILMSINLNLVTFSDLPDSRQLKGDIFAFVIALAAEAAGVLAIVLAIYRNKRSTRIDQFNLLKW

NQO12 subunit

P.denitrMEKF....VLFAPIIASLIAGLGWRAIGEKAAQYLTGVLFLSCLISWYLF..LSF.....DGVPRIHPIVLDVWVTGDFHAEWAIRDLRLTAIMLIVVTVSALVHMYSGLYMAHDDN
 BovineMMHFSLSLVTLTLLTPIMHMSFNTRYKPSNPLYVKTAISYAFITSMIPTMFIH.....SGQELIISNWHLTITQTLKLSLFSKMDYFSMMFIPVALFVTVSIMEFSMWYMSDPN
 Live mitMYLL.....MVILPLIGSFAAGFFRFLGSRGVAVVTTTCVLSLISFSCIAFYEVALL.....CASACYNKIAPWIFSELFDAAWGLFDLSLTVILLVTVIVSSLVHIYSISYMSEDPH
 Live chlMELIFQNWVFPVLPFLASILLGIGLFFFPNSIKFRRLSSSITSMFLNIAMLLSHFHQQITGSPIHRYLWSVLYKNFVLEIGYLLDPLTSHMLVLTVAVMVMIYSDSYMFYDEG
 Rice chlMEHTYQYAWVPIPLPLPVMSMGFLFLVPTATKNLRRIWAFPSVLLLSIAMVFSVHLSIQQINGSSIQYLSWTVWNDSLEFGYLDPLTSHMLILITTVGILVLIYSDDYMSHDEG

P.denitr WTHDEHYKARFFAYLSFFTFAMLMVLTADNLLQMFHGWGVGVASYLLIGFYKKASANAAMAKAFIVNRVGFGLLIGFIYWLTVGSVQFDEIRFQVPQLAQTEIDFLW..RDWNAAN
 BovineINKFFKYLFLITMLILVTANLFLQFLIGWGVGIMSFLIGWYGRADANTAALQAILYNRIGDIGFILAMAWFLTNLTWDLQIFMLNPSDSN.....MP
 Live mit S.....PRFFCYLSITFFMLMLVTGDNFIQLFLGWGVGLASYLLMNFWTRIQANKAAIKAMLMNRVQDFGLALGIMGCFITQTVDFSTIFACASAFSEPHHYFLFCNMGFHAMT
 Live chl YI.....KFFCYLSITFFMLMLVTGDNFIQLFLGWGVGLASYLLMNFWTRIQANKAAIKAMLMNRVQDFGLALGIMGCFITQTVDFSTIFACASAFSEPHHYFLFCNMGFHAMT
 Rice chl YL.....RFFVYISFFNTSMLGLVTSNLIQIYFFWELVGMCSYLLIGWFFTRIPASACQKAFVTRNRVQDFGLLGLIGFVITGSELRDLFKIANNW.....IPNNEINSLTIT

P.denitr LLGFLLVFVGMGKSAQLLHTWLPDAMEGTPVPSALIHATMTVAGVFLVCRMSPLYEFAPDAKNFIVIGATTAFFAATVGLVQNDIKRVIAYSTCSQLGMYFAAGVGVYSAAMFHL
 BovineAATGKSAQFGLHWPVPSAMEGTPVPSALLHSSMTMVAGIFLLIRFYPLTENNKIYQSIITLCLGAIITLFTAMCALTQNDIKIIAFSTSSQLGLMMVTIGINQPYLAFLHIC
 Live mit VICILVFIAGVKGSAQIGLHTWLPDAMEGTPVPSALIHATMTVAGVFMARCSPLFEYSNALIVITFVGAMTSFFAATTGMLQNDLKRVMAYSTCSQLGYMI FACGISNYSVSFVHLM
 Live chl LCALFLFLGPVAKSAQFPLHWPVPSAMEGTPISALIHATMTVAGVFLVCRMSPLYEFAPDAKNFIVIGATTAFFAATVGLVQNDIKRVIAYSTCSQLGMYFAAGVGVYSAAMFHL
 Rice chl LCALFLFLGAVAKSAQFPLHWPVPSAMEGTPISALIHATMTVAGVFLVCRMSPLYEFAPDAKNFIVIGATTAFFAATVGLVQNDIKRVIAYSTCSQLGMYFAAGVGVYSAAMFHL

P.denitr THAFFKAMFLGAGSVIHAM.....HHEQDMRNYGGLRKKIPLTFWAMMIGTFAITGVGIPLTHLGFAGFLSKDAIIIESAYA...GSGYAFWLLVIAACFTSFYSWRILFTFYGK
 Bovine THAFFKAMFLMCSGSIHSL.....NDEQDIRKMGGFLKAMPFTTALIVGSLATGMFF.....LTGFYSKDLIIIEANKSYT..NAWALLMTLITATSTAIYSTRIFALLGQ
 Live mit NHACFKALLFLSAGSVIHAM.....SDEQDMRKMGGFLASLPFTYAMHLMGSLSL..IGFPF.....LTGFYSKDVILELAYTKYITISGNFAFWLGSVSFFTSYYSFRLLFTFLAP
 Live chl THAYSKALLFLGSGSVIHMEPIVGYHPNKSQNMIFMGGRLQYMPITAITFLGTLSLCGIP.....PFACFWSKDEILVNSWHLFPILGSIASF...TAGLTAFYMFRIYFLTFEGD
 Rice chl THAYSKALLFLGSGSVIHMEPIVGYHPNKSQNMIFMGGRLQYMPITAITFLGTLSLCGIP.....PLACFWSKDEILVNSWHLFPILGSIASF...TAGLTAFYMFRIYFLTFEGD

P.denitr PRGDHHAHDHAHESPPVMTIPLGLVLAIGAVFAGHVVYGPFGDHHKVTEFYHIAAGHAAEEGEEAHATAEAPVEHAVADT.....ATAEGEAAAEAEHAEIAPVGGAIYMH
 Bovine PR..FPTLVNINENPLINIKRLLIGSLFAGYI.....ISNNIPPTTI.....
 Live mit TNSFKRDLRSRCHDAPILMAMPLILAFGSIFFVGLAKDMMIG...LGTNFW.....
 Live chl FRGHF..FDDVKK..LSSISIWG.....SLEFNKEQFK.....LDKXSTL.YPKEANNIMLPLIILITPTVIFIGFGLFDEKNMNVDSLSVWLT
 Rice chl LRVHFQNYSTKEDSLYSISLWGRISKGVNRDFVLSTAKSGVSFFSQNLSKIHVNTGNRISGFSSTLGTKNFTVYPHEPGNTMLFPLILLCTLFGISGIFHDNEIGELTILSKWT

P.denitr PDNHIMDEAHHA.PAWVK.....VSPFVAMVGLITAWTF.YIANP.SLPRRLA.....AHEPALYRFLLNKWFDEIYEFIVRPAKWLGR.VLWKGDDGAV
 BovinePQMTMPYYLK.....TTALIVTILGFILELISNMTKMLKYHYPSN.....AF..KFSTLLG...YFPTIMHRLAPYMNLSMSKSSASSLLDLIW
 Live mit PKNEILAESEFATPIIK.....LIPILFSTLGSFVAYSVNVVNPPLIFALKTS.....TFGNRLYCFNKRWFDFKVFNDFLARSFLRFGEVSFKALDKGA
 Live chl LSNISFNYS..NSEKFLFELFNAIPSVSIAFFGILIAFYI...YGNPFSFLKKEKKLQKSEIDVLKFSNFIYNWSYRAYIDGFSYFFIKGLRFLIKIVSF...IDRWI
 Rice chl PSINFFQESSNSSINSYEFITNAISSVSLAIFGLFIGIYVLWICLLFFSEFDLINSFVKGKPKY.....FFHQLKKKIYSWSYNRGYIDIFYTRTFTLIGRLTELTFQF..FDKGV

P.denitr IDGTINGVAMGLIPRLTRAARVQSGYLFHYAFAMVLG.....IVGLLIWMMRGH
 Bovine LEAILPKTIS.LAGMAKSTLVNKGGLIKLYFLSFLITILISMLFNHE
 Live mit IEILGPYGISYIRKMAQMSKIQSGFYVYHAFVMLGLTMTFISVWGLVDIFSWVDNRLYFIYMVSLFINI
 Live chl IDGITH.GIGISFSGESLKYIEGGRISSYLLFFII..FCMFLFLYSYII
 Rice chl IDGITH.GVGLASFCIGEEIKYVGGRISSYLLFFLCYVSFLFFFLS

NQO13 subunit

P.denitrMTNLLSIITFLPIVAAIHAFVLRGQDEAAARNAKWALLTTTATFVLSLVL.FRFPDANTGFQFVEDHAWIM..GVCKYMGVDGIVSVLVLTTFMPLTILSTW.QVQD.
 BovineMLKYIPTIMLPLTWLTKNMMIWNSTAHSLIISFTSLLLMN.....QFGDNS.....LNFSLFFSDSLSTPLLTITMMLPLMLMASQHLHLS
 Live mit MLQVLAPFYNSLGLILLPLLSLIIIVIPNSR...VLMRGTITWTLITFLYSLFFW.MRFENDTAKFQFVETIRWLPYSNINIFYMGMDGILSWFVILTTFTLPCILVGVYSVKS
 SynechocMNTFPWLTITILLPIVAAIFPIIPDKDG...KTVRYSYLAAGVDFALIVYAFYSGFDLSEPGQLQVSEYTWLPQIDLQVSGVADGLSMLPILITGFIITLATMAAW.PVTL.
 Live chlMNHFPWLTITIVLPISAGLVIPFLPSTGN...KIRWYTLGVCLLEFLITTYICFYHQFNDHLIQLKEDYNWISFINFHWRLGIDGFSIGLITLGTITLATLAAW.PVTR.
 Rice chlMSSFPWLTITLVLPILFAGSLIFFLPHRG...KIRWYTMSCILLEFLITMTYAFCYHFQLEDPLIKEDSKWIDVFNFWRLGIDGLSGLSILITGFMFTLATLAAW.PVTR.

P.denitr ...KVKEYMIAFLVLEGLMGVFTALDVLVYLFEEAGLIPMFLIIGIUGKKDR.IYASFKFYFLFGLSVMLVAMIAMYRMAG...TTDPTLTTDFDPSENFRLLGMTVVGGMQMLL
 Bovine NLTRKFLITMLISLQLFIMTFAMLEILFYILFEATLVPTLIIITRWGQTERLNAGLYFLFYTLAGSLPLLVALIYQNTVGSNLFMLQVWVQPVHNS.....WSNVFM
 Live mit ...KVKEYMAFFICESFLTAIVFCSLDLLIFVYFVESVLPIMFINGVWGRQRKIKAAAYQFFLYTLMGSLFMLLAILFIFFGTG...TTDLQMLLTTEFS.....ERRQILL
 Synechoc ...KPKLFYFLMLMGGLIAVFAVDILLFFLWLELVPVYLLISLWGGKRR.LYAATKFILYTAGGSLFILLAGLTAFYGDVN.TFMSAIAAKDIPVN.....LQILL
 Live chl ...NPRLFYFLMLAMYSQIGLFAQDILLFFMWELELIPVYLLAMWGGKRR.LYAATKFILYTAGGSLFILLAGLTAFYGDVN.TFMSAIAAKDIPVN.....LQILL
 Rice chl ...NSRLFYFLMLAMYSQIGLFAQDILLFFMWELELIPVYLLAMWGGKRR.LYAATKFILYTAGGSLFILLAGLTAFYGDVN.TFMSAIAAKDIPVN.....LQILL

P.denitr FLAFAFSAFVKMPMPVHTWLPDAHVQAPTAGSVLLAAVLKMGYGFRLFSPLMPFVASGVAQPVVFLWSAIAIVYTSVLAQSDMKKVIAYSSVAHMGVYTMGVFAANQIGVDGAIF
 Bovine WLACHMAFMVKMPLYGLHLPLKAVHAPAGSMVLAVALLLKGGYGLMRLITLIPNMDTMAFPFIM.LSLWGMINTSSICLRQTDKSLIAYSSVSHMALVIVAILITQTPWSYMGATA
 Live mit WLAFAFSAFVKMPMPVHTWLPDAHVQAPTAGSVLLAAVLKMGYGFRLFSPLMPFVASGVAQPVVFLWSAIAIVYTSVLAQSDMKKVIAYSSVAHMGVYTMGVFAANQIGVDGAIF
 Synechoc YAGFLIAYGVKLPILFHTWLPDAHGEATAPAHMLLAGILLKMGYGLRLMNVGMLPDAAHAFVAPVVLIVGVNIIYAAFTSFAQRNLRKRIAYSSVSHMGFVLIGLASFTDLGMSGAML
 Live chl YLSFLIAYAVKLPILFHTWLPDAHGEATAPAHMLLAGILLKMGYGLRLMNVGMLPDAAHAFVAPVVLIVGVNIIYAAFTSFAQRNLRKRIAYSSVSHMGFVLIGLASFTDLGMSGAML
 Rice chl YFGFLIAYAVKLPILFHTWLPDAHGEATAPAHMLLAGILLKMGYGLRLMNVGMLPDAAHAFVAPVVLIVGVNIIYAAFTSFAQRNLRKRIAYSSVSHMGFVLIGLASFTDLGMSGAML

P.denitr QMLSHGFIISGALFLCVGIVYDRMHTREIDAYGGLVNRMPAYAAVFFMTMANVGLPGTSGFVGFELTLMGVFRVDTW.....VALVATSGVILSAAYALWLYRRVTLGQLIKESLKS.
 Bovine LMIHAGLTSSMLFCLANSYERIHSTRMILARGLQTLPLMNTWLLASLTNLALPPTINILIGELFVVMSTFSWSN.....ITILMGVNMVITALYSLYM...LIMTQRKYTYHIN
 Live mit LMLSHGLVSSALFLCVGALYDRHKTIRVYGGGLVSTPMFSTIFLFTLANMSLPGTSSGIFELILVGAQFQRNSL.....VATLAAGMILGAAYSLWLYNRVVFNGFKNPFLIK.
 Synechoc QMISHLIGASLFFMGVATYDRTHLMDMGGLIGKMKKGFAMWTACSLASLALPGMSGFVAELMVVGFATSDAYNLVFTIIVVLMGVGVILTPIYLLSMLREMLYGPENEELVNH
 Live chl QMISHLIGASLFFLAGISYDRTRTLVLDQMGGLIGKMKKGFAMWTACSLASLALPGMSGFVAELMVVGFATSDAYNLVFTIIVVLMGVGVILTPIYLLSMLREMLYGPENEELVNH
 Rice chl QILSHGFIGATLFFLAGTACDRMLVYLEELGGASIPMPKIFTMSSFSMASLALPGMSGFVAELMVVGFATSDAYNLVFTIIVVLMGVGVILTPIYLLSMLREMLYGPENEELVNH

P.denitr .ITDMTPRERWFIPLAMTILGVYPLVTDVTPAVAAALVQDYNQSQAPAVATAQASH.....
 Bovine NISPSFTRENALMSLHILPLLLTLNPKIILGLY.....
 Live mit .FSDLNRREVLIPLFVIGVIMGVYVEFLCMTSVSNLVHGKFD.....
 Synechoc NLVDVPEPREVFIIGCLLVPIIGIFYPKILITQIYDPTINQLVQARRSVPSLVQQAANLSPLEVTALRPPTIGF
 Live chl YFMDAGPREIFILICLFFPIISITIGIYPNFVLSIWSKVNFLSNFF.....
 Rice chl NFEDSGPRELFLICIFLPVIGIGIYDPFVLSVDRVEALLSNYPK.....

NQO14 subunit

P.denitrMTSLDFSTILPEVVLGAGYALALMAGAYLGKDR.....LARTLLWTVAAFLVVAAMVGLGNHVDGAAF.HGMFIDGFSRFAKVVTLVAAA
 Bovine
 Live mitMFEHDFLALFPEIFLINATIIILLIYGVVSTSKKYDPPVLRNVGWLGLSVLMTILLVAVGSPLAVANLVYNNLMMDNFYFCQIFLLISTA
 SynechocMDFLTLAGLQAGNAGVILPEGIVITLTLVTLTDLILGRQS.....LRLTPALAITGLSAATAVTLQWNTSQNLAFGLGFGNDLSTVFRGIVLLSAA
 Live chlMKLELDMFFLYG...STILPECILIFSLIIILIDLTFFPKD.....TIWLYFISLTSLLISIIILLFOYKTDPIISFLGSFQDTSFNRIQSFIVFCFSI
 Rice chl MIWHVQNFENFILDSTRIFMKAHFHLLFQG...SFIFPECILIFGLIILLIDLTSPQKD.....RPWFFYISSTSLVISITALLFRWREPIISFGNGFQTNFNEIFQFILLCST



FIGURE 3: Comparison of the amino acid sequences of the *Paracoccus NQ08-14* subunits with their homologues from various organisms. The comparison was conducted by the PILEUP program. Asterisks indicates amino acid residues conserved in all listed organisms. ♦ indicates those amino acid residues which are conserved within the respiratory NDH-1 of chloroplasts but not between the respiratory and chloroplast sequences. The amino acid sequences cited are from the indicated references: *R. capsulatus* homologues (Dupuis, 1992), bovine mitochondrial homologues (Anderson et al., 1982; Dupuis et al., 1991), liverwort mitochondrial homologues (Oda et al., 1992), *Synechocystis* homologues (K. Steinmüller, unpublished results, GenBank/EMBL Accession Number X62517; U. Eilersiek & K. Steinmüller, unpublished results, GenBank/EMBL Accession Number X65170), *Synechococcus* homologue (A. Kaplan, unpublished results, GenBank/EMBL Accession Number X65027), liverwort chloroplast homologues (Ohyama et al., 1986), and rice chloroplast homologues (Hiratsuka et al., 1989).

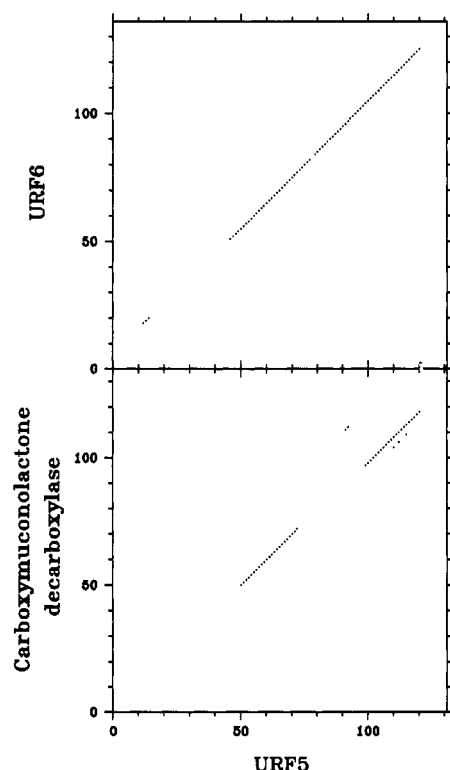


FIGURE 4: Homology between the putative URF5 and URF6 polypeptides (upper panel) and between the putative URF5 polypeptide and the γ -carboxy muconolactone decarboxylase of *A. calcoacetatus* (lower panel) as illustrated by dot matrix analyses. The calculations were conducted using the COMPARE program with a window of 21 and a stringency of 14.

of complex I (Yagi et al., 1992). Morgan-Hughes' group has reported that patients with the Kearn-Sayre syndrome have mitochondria containing large mtDNA deletions (Holt et al., 1988). The *Paracoccus* NDH-1, not only because it contains

homologies of all seven human ND genes but also because of the ease with which bacterial genes can be manipulated, would appear to be a useful model system for conducting biochemical studies of mitochondrial diseases resulting from deletions of these ND genes. Leber's hereditary optic neuropathy is a maternally inherited disease associated with the late onset of bilateral loss of central vision. Wallace et al. (1988) found that in most Leber's disease patients a single nucleotide mutation at position 11778 of mtDNA is present. This mutation converts the 340th amino acid of the ND4 gene product from an arginine to a histidine. As shown below, this arginine is conserved in the *Paracoccus NQ013* gene product (ND4 homologue). In addition, the following mutations have been reported in some Leber's disease patients: ND1 product (alanine₅₂ → threonine and leucine₂₈₅ → proline), ND2 product (glycine₂₅₉ → serine), and ND5 product (alanine₄₅₈ → threonine) (Huoponen et al., 1991; Howell et al., 1991; Johns & Berman, 1991; Brown et al., 1992; Wallace et al., 1991).

Human ND1	47	QPFADAMKLF	56	281	RYDQLMHLW	290
<i>Paracoccus NQ08</i>	60	QTFADALKYI	69	302	RYDQLMRIGW	311
Human ND2	254	LPPLTGFLPK	263			
<i>Paracoccus NQ014</i>	382	VPPTLGFFAK	391			
Human ND4	335	DRTHSRIMIL	344			
<i>Paracoccus NQ013</i>	362	ERMHTREIDA	371			
Human ND5	453	PIKGLAAGSLF	463			
<i>Paracoccus NQ012</i>	474	PLGVLAIGAVF	484			

These wild-type residues are also conserved in the ND homologues of the *Paracoccus* NDH-1 as shown in the alignments above. In each alignment, the asterisk indicates the position of mutation in Leber's hereditary optic neuropathy. The limited amount of material available from patient tissues coupled with the mtDNA heteroplasmy observed in these

tissues presents a serious impediment to biochemical research in mitochondrial diseases. The use of bacterial model systems would alleviate this difficulty. Therefore, the *Paracoccus* NDH-1 represents an attractive model system for the investigation of mitochondrial diseases caused not only by the deletion of ND genes but also by point mutations within the ND gene products.

ACKNOWLEDGMENT

We thank Daniel G. Owens for his excellent technical assistance and Dr. Carla Hekman (R. W. Johnson PRI, New Jersey) for critical reading of the manuscript.

REFERENCES

- Anderson, S., De Bruijn, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F., & Young, I. G. (1982) *J. Mol. Biol.* 156, 683–717.
- Brown, M. D., Voljavec, A. S., Lott, M. T., Torroni, A., Yang, C.-C., & Wallace, D. C. (1992) *Genetics* 130, 163–173.
- Chomyn, A., Marionetti, P., Cleeter, M. W. J., Ragan, C. I., Doolittle, R. F., Matsuno-Yagi, A., Hatefi, Y., & Attardi, G. (1985a) in *Achievements and Perspectives of Mitochondrial Research* (Quagliariello, E., Ed.) pp 259–275, Elsevier, Amsterdam.
- Chomyn, A., Mariottini, P., Cleeter, M. W. J., Ragan, C. I., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R. F., & Attardi, G. (1985b) *Nature* 314, 591–597.
- Chomyn, A., Cleeter, M. W. J., Ragan, C. I., Riley, M., Doolittle, R. F., & Attardi, G. (1986) *Science* 234, 614–618.
- Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res.* 12, 387–395.
- Dupuis, A. (1992) *FEBS Lett.* 301, 215–218.
- Dupuis, A., Skehel, J. M., & Walker, J. E. (1991) *Biochemistry* 30, 2954–2960.
- Earley, F. G. P., Patel, S. D., Ragan, C. I., & Attardi, G. (1987) *FEBS Lett.* 219, 108–113.
- George, C. L., Ferguson, S. J., Cleeter, M. W. J., & Ragan, C. I. (1986) *FEBS Lett.* 198, 135–139.
- Harms, N., deVries, G. E., Maurer, K., Veltkamp, E., & Stouthamer, A. H. (1985) *J. Bacteriol.* 164, 1064–1070.
- Harms, N., De Vries, G. E., Maurer, K., Hoogendijk, J., & Stouthamer, A. H. (1987) *J. Bacteriol.* 169, 3969–3975.
- Hartnett, C., Neidle, E., Ngai, K.-L., & Ornston, L. N. (1990) *J. Bacteriol.* 172, 956–966.
- Hatefi, Y. (1985) *Annu. Rev. Biochem.* 54, 1015–1069.
- Hatefi, Y., Ragan, C. I., & Galante, Y. M. (1985) in *The Enzymes of Biological Membranes* (Martonosi, A. N., Ed.) pp 1–70, Plenum Publishing, New York.
- Hayashi, M., Hasegawa, K., Oguni, Y., & Unemoto, T. (1990) *Biochim. Biophys. Acta* 1035, 230–236.
- Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honjo, Y., Sun, C. R., Meng, B. Y., Li, Y. Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., & Sugiura, M. (1989) *Mol. Gen. Genet.* 217, 185–194.
- Holt, I. J., Harding, A. E., & Morgan-Hughes, J. A. (1988) *Nature* 331, 717–719.
- Howard, P. K., Shaw, J., & Otsuka, A. J. (1985) *Gene* 35, 321–331.
- Howell, N., Kubacka, I., Xu, M., & McCullough, D. A. (1991) *Am. J. Hum. Genet.* 48, 935–942.
- Huoponen, K., Vilkkii, J., Aula, P., Nikoskelainen, E. K., & Savontaus, M.-L. (1991) *Am. J. Hum. Genet.* 48, 1147–1153.
- John, P., & Whatley, F. R. (1975) *Nature* 254, 495–498.
- John, P., & Whatley, F. R. (1977) *Biochim. Biophys. Acta* 463, 129–153.
- Johns, D. R., & Berman, J. (1991) *Biochem. Biophys. Res. Commun.* 174, 1324–1330.
- Kikuno, R., & Miyata, T. (1985) *FEBS Lett.* 189, 85–88.
- Kurowski, B., & Ludwig, B. (1987) *J. Biol. Chem.* 262, 13805–13811.
- Meinhardt, S. W., Kula, T., Yagi, T., Lillich, T., & Ohnishi, T. (1987) *J. Biol. Chem.* 262, 9147–9153.
- Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y., Kohchi, T., & Ohyama, K. (1992) *J. Mol. Biol.* 223, 1–7.
- Ohnishi, T., Ragan, C. I., & Hatefi, Y. (1985) *J. Biol. Chem.* 260, 2782–2788.
- Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., & Ozeki, H. (1986) *Nature* 322, 571–574.
- Parke, D. (1979) *Biochim. Biophys. Acta* 578, 145–154.
- Ragan, C. I. (1987) *Curr. Top. Bioenerg.* 15, 1–36.
- Raitio, M., Jalli, T., & Saraste, M. (1987) *EMBO J.* 6, 2825–2833.
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463–5467.
- Shine, J., & Dalgarno, L. (1975) *Nature* 254, 34–38.
- Steinmüller, K., Ley, A. C., Steinmetz, A. A., Sayre, R. T., & Bogorad, L. (1989) *Mol. Gen. Genet.* 216, 60–69.
- Stouthamer, A. H. (1980) *Trends Biochem. Sci.* 5, 164–166.
- Stouthamer, A. H. (1992) *Antonie Van Leeuwenhoek* 61, 1–33.
- Takahashi, Y., Shonai, F., Fujita, Y., Kohchi, T., Ohyama, K., & Matsubara, H. (1991) *Plant Cell Physiol.* 32, 969–981.
- Tuschen, G., Sackmann, U., Nehls, U., Haiker, H., Buse, G., & Weiss, H. (1990) *J. Mol. Biol.* 213, 845–857.
- Unemoto, T., Shimada, H., & Hayashi, M. (1992) *Biochim. Biophys. Acta* 1099, 170–174.
- Van Spanning, R. J. M., Wansell, C., Harms, N., Oltmann, L. F., & Stouthamer, A. H. (1990a) *J. Bacteriol.* 172, 986–996.
- Van Spanning, R. J. M., Warnsell, C. W., Reijnders, W. N. M., Oltmann, L. F., & Stouthamer, A. H. (1990b) *FEBS Lett.* 275, 217–220.
- Wallace, D. C., Singh, G., Lott, M. T., Hodge, J. A., Schurr, T. G., Lezza, A. M. S., Elsas, L. J., & Nikoskelainen, E. K. (1988) *Science* 242, 1427–1430.
- Wallace, D. C., Lott, M. T., Torroni, A., & Shoffener, J. M. (1991) *Cytogenet. Cell Genet.* 58, 1103–1123.
- Weidner, U., Nehls, U., Schneider, R., Fecke, W., Leif, H., Schmiede, A., Friedrich, T., Zensen, R., Schulte, U., Ohnishi, T., & Weiss, H. (1992) *Biochim. Biophys. Acta* 1101, 177–180.
- Weiss, H., Friedrich, T., Hofhaus, G., & Preis, D. (1991) *Eur. J. Biochem.* 197, 563–576.
- Xu, X., & Yagi, T. (1991) *Biochem. Biophys. Res. Commun.* 174, 667–672.
- Xu, X., Matsuno-Yagi, A., & Yagi, T. (1991a) *Biochemistry* 30, 8678–8684.
- Xu, X., Matsuno-Yagi, A., & Yagi, T. (1991b) *Biochemistry* 30, 6422–6428.
- Xu, X., Matsuno-Yagi, A., & Yagi, T. (1992a) *Biochemistry* 31, 6925–6932.
- Xu, X., Matsuno-Yagi, A., & Yagi, T. (1992b) *Arch. Biochem. Biophys.* 296, 40–48.
- Yagi, T. (1986) *Arch. Biochem. Biophys.* 250, 302–311.
- Yagi, T. (1987) *Biochemistry* 26, 2822–2828.
- Yagi, T. (1989) *Protein Nucleic Acid Enzyme* 34, 351–363.
- Yagi, T. (1990) *Arch. Biochem. Biophys.* 281, 305–311.
- Yagi, T. (1991) *J. Bioenerg. Biomembr.* 23, 211–225.
- Yagi, T. (1993) *Biochim. Biophys. Acta* 1141, 1–17.
- Yagi, T., & Hatefi, Y. (1988) *J. Biol. Chem.* 263, 16150–16155.
- Yagi, T., & Dinh, T. M. (1990) *Biochemistry* 29, 5515–5520.
- Yagi, T., Hon-nami, K., Ohnishi, T. (1988) *Biochemistry* 27, 2008–2013.
- Yagi, T., Xu, X., & Matsuno-Yagi, A. (1992) *Biochim. Biophys. Acta* 1101, 181–183.
- Yasunobu, K. T., & Tanaka, M. (1980) *Methods Enzymol.* 69, 228–238.
- Yeh, W.-K., Feltcher, P., & Ornston, L. N. (1980) *J. Biol. Chem.* 255, 6342–6346.