Gene Cluster of the Energy-Transducing NADH–Quinone Oxidoreductase of *Paracoccus denitrificans*: Characterization of Four Structural Gene Products†,‡

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 March 19, 1992; Revised Manuscript Received May 13, 1992

ABSTRACT: In previous reports from our laboratory, the three structural genes (*NQO1*, *NQO2*, and *NQO3*) of the energy-transducing NADH–quinone oxidoreductase of *Paracoccus denitrificans* were characterized [Xu, X., Matsuno-Yagi, A., & Yagi, T. (1991) *Biochemistry 30*, 6422–6428; (1991) *Biochemistry 30*, 8678–8684; (1992) *Arch. Biochem. Biophys. 296*, 40–48]. In this report, the four structural genes *NQO4*, *NQO5*, *NQO6*, and *NQO7* of the same *Paracoccus denitrificans* oxidoreductase were cloned and sequenced. On the basis of sequence homology and immunological cross-reactivity, these genes encode counterparts of the 49-, 30-, and 20-kDa polypeptides and the mitochondrial DNA ND3 polypeptides of bovine mitochondrial complex I. These seven structural genes were found to be located in the same gene cluster. The order of the seven structural genes of the *Paracoccus* NADH–quinone oxidoreductase in the gene cluster is *NQO7*, *NQO6*, *NQO5*, *NQO4*, *NQO2*, *NQO1*, and *NQO3*. Upstream of the *NQO7* gene, an open reading frame encoding a predicted polypeptide homologous to the UV repair enzyme A of *Escherichia coli* and *Micrococcus lysodeikticus* was detected. The 5′-terminus of the gene cluster carrying the *Paracoccus* NADH–quinone oxidoreductase was studied, and the possible promoter region is discussed. The *NQO4* and *NQO5* genes appear to code for the M_r 48 000 and 21 000 polypeptides of the isolated *Paracoccus* NADH dehydrogenase complex [Yagi, T. (1986) *Arch. Biochem. Biophys. 250*, 302–311] on the basis of amino acid analyses and N-terminal protein sequence analyses. The antisera to the bovine complex I 49- and 30-kDa polypeptides cross-reacted with the *Paracoccus* 48- and 21-kDa subunits, respectively.

The NADH–quinone (Q)[1] oxidoreductases of the respiratory chains from many species of animal and bacteria catalyze electron transfer from NADH to Q (Hatefi, 1985; Hatefi et al., 1985; Weiss et al., 1991; Yagi, 1992). The NADH–Q oxidoreductases of the bacterial respiratory chain can be divided in two groups depending on whether they bear an energy-coupling site (Yagi, 1989, 1991, 1992; Yagi et al., 1988). Those enzymes that bear the coupling site are designated as NDH-1 and those that do not as NDH-2. In terms of prosthetic groups, inhibitor specificity, and polypeptide composition, the NDH-1 appears to be related to the mammalian mitochondrial NADH–UQ oxidoreductase (complex I) (Yagi, 1987, 1990, 1991, 1992; Meinhardt et al., 1987).

Paracoccus denitrificans is a Gram-negative soil bacterium (John & Whatley, 1977). Aerobically grown *Paracoccus denitrificans* expresses a mitochondrial-type respiratory chain (Stouthamer, 1980) which appears to contain NDH-1 but not NDH-2 (Yagi, 1991). The NDH-1 of the *Paracoccus* membranes has been shown to have considerable similarity to its mitochondrial counterpart in terms of EPR-visible iron–sulfur clusters and immuno-cross-reactivity (Albracht et al., 1980; Meinhardt et al., 1987; Yagi, 1986, 1991; Yagi & Dinh, 1990; Xu & Yagi, 1991; George et al., 1986; Xu et al., 1991a, 1992). On the other hand, the NADH dehydrogenase complex purified from *Paracoccus* membranes appears to be structurally simpler than its mitochondrial counterparts (Yagi, 1986, 1991, 1992). Recently, we have cloned the 5.7 kbp *Paracoccus* DNA fragment (designated pXT-1) carrying the structural gene (*NQO1*) which encodes the 50-kDa (NADH-binding) subunit of the *Paracoccus* NDH-1 (Xu et al., 1991a). Furthermore, the structural genes encoding the subunits of the *Paracoccus* NDH-1 have been found to constitute at least a single gene cluster (Xu et al., 1991a,b, 1992; Yagi et al., 1991). At the present time, the *NQO1* (NADH-binding subunit), *NQO2* (25-kDa subunit), *NQO3* (66-kDa subunit), URF1, URF2, and URF3 genes in this cluster have been sequenced (Xu et al., 1991a,b, 1992). The arrangement of these genes is as follows: *NQO2*, URF2, URF1, *NQO1*, URF3, and *NQO3*. This arrangement resembles that of the structural genes of the NAD-linked hydrogenase of *Alcaligenes eutrophus* H16, suggesting that the NAD(H) catalytic fraction of the *Paracoccus* NDH-1 may be evolutionally related to that of the *Alcaligenes* NAD-linked hydrogenase (Tran-Betcke et al., 1990).

Recently, our partial DNA sequencing downstream of the *NQO3* gene indicated the presence of polypeptides homologous to the bovine mitochondrial ND-1, ND-5, and ND-2 gene products which are believed to be components of the hydrophobic protein fraction (HP) of bovine complex I (Xu et al., 1992; Yagi, 1992; Yagi et al., 1992). These data suggest that this gene cluster contains not only structural genes (*NQO1*, *NQO2*, and *NQO3*) encoding hydrophilic subunits but also genes encoding hydrophobic subunits. In addition, we have

† This work was supported by U.S. Public Health Service Grant R01 GM33712. Facilities for the computer are supported by U.S. Public Health Service Grant M01 RR00833 for the General Clinical Research Center. Synthesis of oligonucleotides is in part supported by the Sam & Rose Stein Charitable Trust. This is Publication No. 7253-CH from The Scripps Research Institute, La Jolla, CA.

‡ The nucleotide sequence in this paper has been submitted to GenBank under Accession Number M93015.

* To whom correspondence should be addressed.

[1] Abbreviations: Q, quinone; UQ, ubiquinone; UQ_1, ubiquinone-1; NDH-1 or complex I, energy-transducing NADH–quinone oxidoreductase; NDH-2, NADH–quinone oxidoreductase lacking an energy coupling site; complex III, ubiquinol–cytochrome *c* oxidoreductase; bp, base pair(s), FP, IP, and HP, flavoprotein, iron–sulfur protein, and hydrophobic protein fractions of complex I, respectively; Tris-buffered saline, a buffer containing 20 mM Tris-HCl (pH 7.5) and 150 mM NaCl; SDS, sodium dodecyl sulfate; EPR, electron paramagnetic resonance.

0006-2960/92/0431-6925$03.00/0

reported that a sequence similar to the bovine complex I 49-kDa subunit is present upstream of *NQO2*. However, this gene (*NQO4*) is truncated at the 5′-terminus of pXT-1 DNA fragment (Xu et al., 1991a). Therefore, it was of interest to isolate a fragment in close proximity to the 5′-terminus of pXT-1 in hope that the 5′-terminus of the gene cluster of the *Paracoccus* NDH-1 as well as other structural genes coding for subunits of this enzyme complex might be discovered.

This paper describes the complete nucleotide sequence of a *Hin*dIII 3.4 kbp fragment designaed pXT-3. This DNA fragment overlaps the 5′-terminal region of pXT-1 and contains four previously unidentified structural genes encoding subunits of the *Paracoccus* NDH-1. These genes have been designated *NQO4* (1239 bp), *NQO5* (624 bp), *NQO6* (522 bp), and *NQO7* (366 bp). The calculated molecular weights of the polypeptides encoded by these four genes are 46 674, 23 731. 19 117, and 13 601, respectively. The *Paracoccus* NDH-1 genes are arranged in the pXT-3 fragment in the following order: *NQO7*, *NQO6*, *NQO5*, and *NQO4*. In addition, this *Hin*dIII 3.4 kbp DNA fragment bears a gene upstream from the *NQO7* gene which encodes a homologue of *uvrA* polypeptide which is a subunit of ATP-dependent excinuclease (Husain et al., 1986; Shiota & Nakayama, 1989). The possible 5′-terminus of the gene cluster bearing the *Paracoccus* NDH-1 has been discussed.

MATERIALS AND METHODS

Isolation of the 48- and 21-kDa Subunits of the Paracoccus NADH Dehydrogenase Complex. Paracoccus denitrificans (ATCC13543) cells were grown aerobically by BIOPURE Fine Chemicals, Inc., Boston, in a 550-L culture with glucose as substrate as described previously (Yagi, 1986). *Paracoccus* membranes and the *Paracoccus* NADH dehydrogenase complex were prepared according to Yagi (1986). The 48- and 21-kDa subunits of the *Paracoccus* NADH dehydrogenase complex were purified by preparative SDS–polyacrylamide gel electrophoresis, on 10% polyacrylamide gels, followed by electroelution (Yagi & Hatefi, 1988; Yagi, 1989; Yagi & Dinh, 1990).

Amino Acid Analysis of the 48- and 21-kDa Subunits. Amino acid analyses of the isolated 48- and 21-kDa subunits were conducted by subjecting the purified 48- and 21-kDa subunits (10 μg of each) to lyophilization followed by hydrolysis in an evaporated and sealed tube in 6 N HCl at 100 °C for 24 h. Amino acid analyses were then performed on a Beckman 7300 amino acid analyzer.

Sequence Analyses of the 48- and 21-kDa Subunits. The N-terminal amino acid sequences of the 48- and 25-kDa subunits were determined according to Matsudaira (1987). Samples of the 48- and 21-kDa subunits (10 μg of each) were lyophilized, dissolved in SDS–PAGE sample buffer containing 80 mM Tris-HCl (pH 6.8), 6% SDS, 5% 2-mercaptoethanol, 0.005% bromophenol blue, and 20% glycerol, applied to an SDS–polyacrylamide slab gel (55 × 95 × 0.75 mm) composed of 10% acrylamide (Laemmli, 1970), and electrophoresed for 1 h at 200 V. The subunits were then electrophoretically transferred to poly(vinylidene difluoride) (PVDF) membranes (Immobilon, Millipore) as described by Matsudaira (1987). The membranes were stained with 0.1% Coomassie brillant blue R-250 in 50% methanol for 5 min, destained for 10 min with a solution containing 10% acetic acid and 50% methanol, rinsed with H_2O, and air-dried. The transferred protein on the PVDF membrane was subjected to sequence analyses using an Applied Biosystem 470A gas-phase protein sequencer.

Immunoblotting. Immunoblotting experiments were carried out by a modification of the procedure of Hekman and Hatefi (1991). Subsequent to SDS–PAGE and transfer of the proteins to nitrocellulose membranes, the membranes were blocked with 2% skim milk in Tris-buffered saline for 1 h at 37 °C (Yagi & Dinh, 1990). Affinity-purified primary antibodies (Xu & Yagi, 1991; Han et al., 1989) were incubated with the nitrocellulose membranes for 2.5 h at room temperature. After three 5-min washes with Tris-buffered saline, the nitrocellulose membranes were incubated for 1 h at 37 °C with a 1 to 5000 dilution of goat anti-rabbit IgG conjugated to alkaline phosphatase. The color development was performed as described (Hekman & Hatefi, 1991) using nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate.

Construction and Screening of the Paracoccus Genomic Library. General cloning techniques were carried out essentially according to Sambrook et al. (1989). The *Paracoccus* genomic DNA was isolated and digested with *Hin*dIII. The *P. denitrificans* genomic library was prepared for ligation into the *Hin*dIII site of the pBluescript KS phagemid vector and recombinated in *E. coli* JM109. The *Eco*RI–*Hin*dIII 222 bp fragment located at 5′-terminal region of the pXT-1 was excised, and isolated from agarose gels. This DNA fragment was used as the probe for screening. Screening was performed as described previously (Xu et al., 1991a, 1992).

DNA Sequencing Strategy. The isolated DNA insert was digested with *Apa*I, *Sac*II, *Pst*I, *Sal*I, *Bss*HII, and *Eag*I. All the fragments thus produced were subcloned into the pBluescript KS phagemid vector and amplified in *E. coli* JM109.

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 ExoIII/mung bean deletion kit from Stratagene. The universal primers T3 and T7 and unique internal primers, 18 bases in length, were used in these experiments. When unsolved regions were encountered, the dITP sequencing kit (Pharmacia) was used instead of the 7-deaza-dGTP. Both complementary DNA strands were sequenced at least 3 times. The 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, 1992), the University of Wisconsin Genetic Computer Group's software programs were used to analyze the sequence data (Devereux et al., 1984). The open reading frames and the terminators were searched by the CODONPREFERENCE and the TERMINATOR programs, respectively. A comparison of the polypeptides was carried out with the BESTFIT and PILEUP programs. The FASTA program was used to search the GenBank/EMBL Sequence databases for sequences homologous to the 48-kDa subunit, the 21-kDa subunit, the *NQO6* product, the *NQO7* product, and the open reading frames.

Other Analytical Procedures. Protein was estimated by the method of Lowry et al. (1951) or by the biuret method in the presence of sodium deoxycholate at 1 mg/mL (Gornall et al., 1949). Any variations from these procedures and other details are described in the figure legends.

Materials. Acrylamide, SDS, and Coomassie brilliant blue R-250 were from Bio-Rad; 7-deaza-dGTP and dITP sequencing kits and universal primers were from Pharmica LKB; [α-^{35}S]dATP was from Amersham; Sequenase version 2 was from United States Biochemical Corp.; *Eag*I was from New England Biolab; alkaline phosphatase-conjugated affinity-purified antibodies to rabbit IgG were from Calbiochem; deletion mutant kits, pBluescript, and conventional restriction enzymes were from Stratagene. The monospecific antisera to the bovine 49-kDa subunit and the bovine 30-kDa subunit

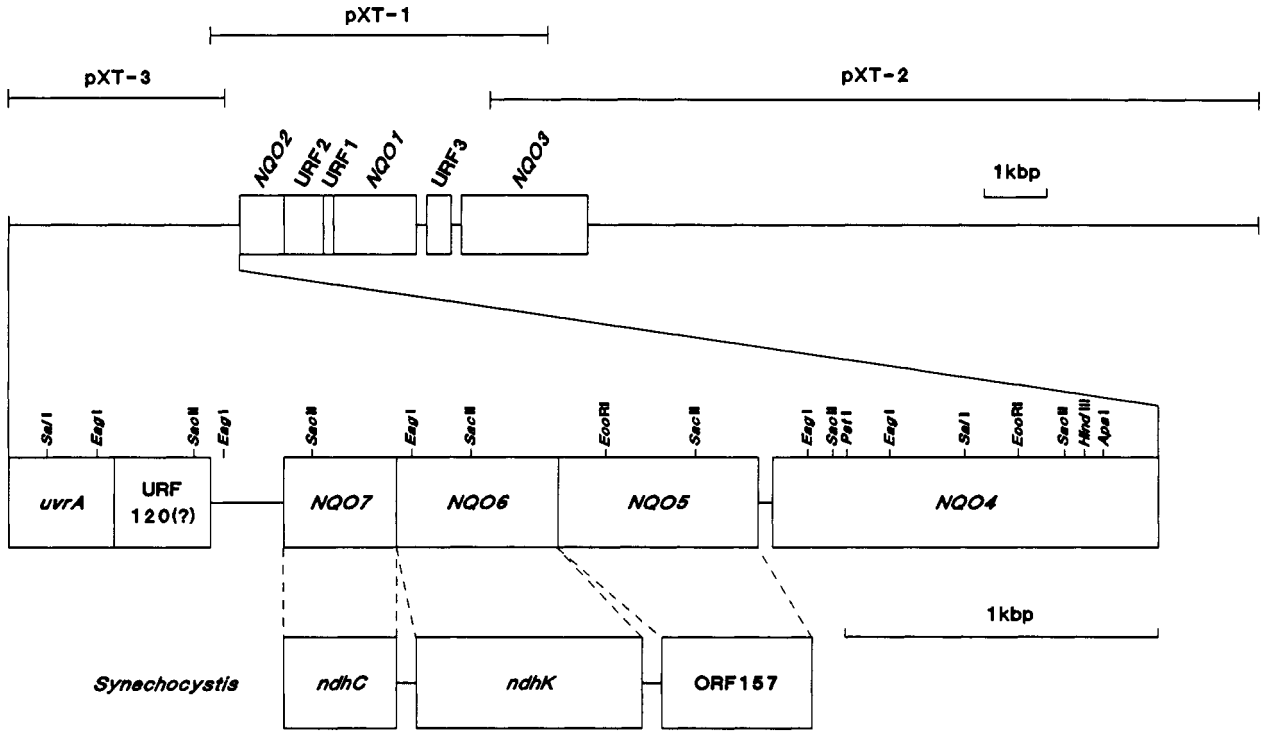

FIGURE 1: Gene map of the NDH-1 gene cluster. The gene map of the pXT-1, pXT-2, and pXT-3 DNA fragments is illustrated together with the *NQO4*, *NQO5*, *NQO6*, *NQO7*, URF120, and *uvrA* genes. The *NQO1*, *NQO2*, and *NQO3* genes, respectively, are the structural genes of the NADH-binding subunit, the 25-kDa subunit, and the 66-kDa subunit of the *Paracoccus* NDH-1. For comparison, the gene arrangements of cyanobacterium *Synechocystis* PCC6803 DNA are shown. Similar gene arrangements are reported in chloroplast DNA (Ohyama et al., 1986; Steinmüller et al., 1989).

Chart I

	1	55
Paracoccus	KLSKELSRRATGKTLYILDEPTTGLHFEDVRKLLEVLHSLVDQGNTVVVIEHNLD	
E. coil	KLARELSKRGTGQTLYILDEPTTGLHFADIQQLLDVLHKLRDQGNTIVVIEHNLD	
	839	893
	56	110
Paracoccus	VIKTADWIIDIGPEGGDGGGAIVATGTPEQVAEVPESHTGRYLAPMLDAARSAAE	
E. coil	VIKTADWIVDLGPEGGSGGGEILVSGTPETVAECEASHTARFLKPML	
	894	940

and the bovine complex I were a generous gift from Dr. Youssef Hatefi (The Scripps Research Institute, La Jolla).

RESULTS AND DISCUSSION

Sequence Analysis. As described previously, the pXT-1 which carries the 5.7 kbp *Eco*RI DNA fragment of the *Paracoccus denitrificans* NDH-1 contains the *NQO1*, *NQO2*, and *NQO3* genes which encode the 50- (NADH-binding), 25-, and 66-kDa subunits of the *Paracoccus* NADH–UQ oxidoreductase, respectively. In addition, pXT-1 also contains a gene, designated *NQO4*, which appears to encode a subunit of the *Paracoccus* NDH-1 which may be a counterpart of the 49-kDa subunit of the bovine complex I. However, the *NQO4* gene is truncated at the 5′-terminus of pXT-1. In order to isolate the remaining region of the *NQO4* gene, we screened a *Paracoccus* genomic library constructed from *Hin*dIII DNA fragments. The probe used to screen the library was the *Eco*RI–*Hin*dIII DNA fragment (222 bp) corresponding to the 5′-terminal region of pXT-1 (see Figure 1). Three positive clones were obtained and subjected to restriction mapping. All three clones were found to be identical and to contain a 3.4 kbp DNA fragment. Furthermore, DNA sequence determination of the 3′-terminal region of these clones confirmed that they all overlap the 5′-terminal region of pXT-1. Therefore, one clone (designated pXT-3) was selected and used in all subsequent analyses.

As seen in Figure 1, the *NQO4* gene and the entire DNA of pXT-3 have been sequenced. These data confirm that the *NQO4* gene, which lies upstream from the *NQO2* (25-kDa subunit) gene in pXT-1, encodes a polypeptide homologous to the 49-kDa polypeptide of bovine complex I. In addition to *NQO4*, pXT-3 also contains three other structural genes which have been found to be homologous to polypeptides of mitochondrial complex I. The genes have been designated *NQO5*, *NQO6*, and *NQO7* (see Figure 1) and encode homologues of the 30-kDa, 20-kDa, and ND3-encoded polypeptides of the bovine complex I, respectively (see Table I).

Figure 1 further indicates that two open reading frames are located upstream from the *NQO7* gene. When these amino acid sequences were compared with other reported amino acid sequences by the FASTA program, it became apparent that a putative polypeptide encoded by one of these open reading frames has significant homology (74% identity) to the C-terminal 100 residues of the UV repair enzyme A encoded by the *uvrA* gene of *E. coli* (Husain et al., 1986) as described in Chart I. Furthermore, this *Paracoccus* open reading frame

Table I: Characteristics of the *NQO4*, *NQO5*, *NQO6*, and *NQO7* Genes and Products

gene	DNA length (bp)	characteristics of polypeptide product: no. of amino acids	MW[a]	pI	bovine homologue
NQO4	1239	413	46674	5.41	49 kDa (IP)[b]
NQO5	624	208	23731	4.96	30 kDa (IP)
NQO6	522	174	19117	7.70	20 kDa (IP)
NQO7	366	122	13601	4.19	ND3 product (HP)

[a] The values of the molecular weights of subunits have been calculated on the basis of the amino acid sequences deduced from the structural genes. [b] Bovine complex I can be resolved into a water-soluble fraction and a water-insoluble fraction (hydrophobic protein fraction, HP) (Hatefi, 1985; Hatefi et al., 1985). The water-soluble fraction can then be separated into a flavoprotein fraction (FP) and an iron–sulfur protein fraction (IP) by ammonium sulfate fractionation. The subfraction containing each bovine homologue is indicated.

Table II: Amino Acid Composition of the 48-kDa (*NQO4*) and the 21-kDa (*NQO5*) Subunits of the *Paracoccus* NDH-1

amino acid	48 kDa (*NQO4*): by amino acid analysis	48 kDa (*NQO4*): from DNA sequence	21 kDa (*NQO5*): by amino acid analysis	21 kDa (*NQO5*): from DNA sequence
alanine	39.4	34	15.1	13
arginine	29.0	35	16.1	11
aspartic acid	39.9	42	22.4	22
glutamic acid	46.8	42	25.5	24
glycine	39.4	35	14.8	12
histidine	8.0	12	2.0	4
isoleucine	19.8	22	10.5	11
leucine	42.1	45	21.8	21
lysine	13.1	13	6.7	7
methionine	9.6	16	0.6	2
phenylalanine	13.9	14	11.4	11
proline	22.4	22	10.3	11
serine	16.0	12	10.8	11
threonine	20.0	16	8.4	8
tyrosine	13.3	14	8.0	8
valine	24.7	24	18.0	21

has 72% identity to the C-terminal region of the predicted UV repair enzyme A of *Micrococcus lysodeikticus* (Shiota & Nakayama, 1989). With the exception of these two bacterial *uvrA* enzymes, no polypeptides having any significant similarity to any of the other sequences in this open reading frame have been found in the GenBank/EMBL database. Therefore, this open reading frame, which is truncated at the 5′-terminus of pXT-3, has been designated *uvrA*. A predicted polypeptide encoded by the other open reading frame (URF120) has no significant homology to any polypeptide in the GenBank/EMBL database. If this URF expresses a polypeptide, the polypeptide will most likely be related to the *uvrA* enzyme because the initiation codon of URF120 is located only 5 base pairs downstream from the termination codon of the *uvrA* gene. However, when the search of the putative terminator was carried out, such a terminator sequence was found in the central region of URF120 (see Figure 2). Therefore, it remains to be seen whether URF120 is coding for a protein. In either case, the *NQO7* gene appears to be the 5′-terminal structural gene of the gene cluster encoding the *Paracoccus* NDH-1.

As shown in Figure 2, the initiation codons of the *NQO4*, *NQO5*, *NQO6*, and *NQO7* genes were all found to be preceded by sequences resembling the Shine–Dalgarno ribosome-binding site (Shine & Dalgarno, 1975). Although the initiation codons of the *NQO4*, *NQO5*, and *NQO6* genes are ATG, that of the *NQO7* gene is probably GTG. GTG has been established as the initiation codon of several genes (Gold et al., 1981) including a number of complex III operons of various bacteria

FIGURE 2: Nucleotide and predicted amino acid sequences of *NQO2*, *NQO4*, *NQO5*, *NQO6*, *NQO7*, URF120, and *uvrA*. The amino acid sequences confirmed by protein N-terminal sequence analysis are underlined. The putative Shine–Dalgarno sequences are doubly lined. The putative terminator sequences are marked by a dashed line. The putative promoter (the GATA sequence) and inverted sequences are shown, respectively, by a box and arrows.

including *P. denitrificans* (Kurowski & Ludwig, 1987; Yun et al., 1990). The G-C content of the *NQO4* gene (65.6%), the *NQO5* gene (62.5%), the *NQO6* gene (66.7%), and the *NQO7* gene (59.2%) was high as was also found to be the case for the other genes of *P. denitrificans* (Xu et al., 1991a,b, 1992).

The promoters for these genes have not yet been identified. However, in the front of the *NQO7* gene, there is GATA sequence that may represent a promoter element; TATA boxes

Table III: Comparison Matrices for the *NQO4*, *NQO5*, *NQO6*, and *NQO7* Subunits of the Energy-Transducing NADH–Ubiquinone Oxidoreductases

		percent identity					
subunit	source	(1)	(2)	(3)	(4)	(5)	(6)
NQO4	(1) *P. denitrificans*	100	58.5	59.0	41.1	41.1	40.7
	(2) bovine		100	66.5	41.4	41.6	41.0
	(3) *N. crassa*			100	38.1	37.8	37.4
	(4) *Synechocystis*				100	77.3	69.7
	(5) liverwort					100	82.7
	(6) rice						100
NQO5	(1) *P. denitrificans*	100	46.2	44.7	26.1	20.7	20.8
	(2) bovine		100	53.3	26.1	23.7	20.1
	(3) *N. crassa*			100	23.6	23.1	22.0
	(4) *Synechocystis*				100	55.4	54.1
	(5) liverwort					100	70.4
	(6) rice						100
NQO6	(1) *P. denitrificans*	100	59.6	40.2	44.3	40.8	40.2
	(2) *Paramecium*		100	44.9	43.0	47.4	44.2
	(3) liverwort			100	65.3	55.6	53.7
	(4) rice				100	51.1	48.9
	(5) *Synechocystis* G1					100	62.6
	(6) *Synechocystis* G2						100
NQO7	(1) *P. denitrificans*	100	33.0	28.7	35.8	38.3	39.2
	(2) bovine		100	30.4	27.0	27.8	27.8
	(3) *Podospora*			100	24.2	30.8	29.2
	(4) *Synechocystis*				100	62.5	66.7
	(5) liverwort					100	71.7
	(6) rice						100

and other sequences typical of *E. coli* promoters have not been found at the anticipated sites. In the front of the *NQO7* gene, inverted repeats flank the GATA sequence. Raitio et al. (1987) have proposed that GATA sequences may be involved in the control of expression of the *Paracoccus* cytochrome oxidase operons. Therefore, it may also be possible that the GATA sequence regulates expression of the NDH-1 gene cluster. However, this hypothesis needs additional direct verification.

Characterization of the NQO4, NQO5, NQO6, and NQO7 Gene Products. On the basis of the predicted molecular weights for the *NQO4* and *NQO5* gene products, the 48- and 21-kDa subunits of the *Paracoccus* NDH-1 complex seemed to be likely candidates for these gene products. In an attempt to demonstrate this identity, the 48- and 21-kDa subunits were purified from the *Paracoccus* NADH dehydrogenase complex by electroelution from SDS–polyacrylamide gels. As seen in Table II, the results of amino acid analyses conducted on the purified 48- and 21-kDa subunits agreed with the amino acid composition deduced from the *NQO4* and the *NQO5* genes, respectively. In addition, the isoelectric points of 5.4 and 5.0 calculated from the deduced amino acid compositon of the *NQO4* and the *NQO5* gene products appear to correspond with the values obtained from isoelectric focusing (Yagi & Dinh, 1990). To further confirm this identity, the purified subunits were subjected to N-terminal amino acid sequencing according to Matsudaira (1987). The sequences of the first 15 amino acids were MDGDIRKNSLDDGSM for the 48-kDa subunit and SEALSDEALLELAEX for the 21-kDa subunit (Figure 2). These sequences are identical to the amino acid sequences deduced from the respective genes. Together, these data provide strong evidence in support of the hypothesis that the *NQO4* and the *NQO5* genes encode, respectively, the 48- and 21-kDa subunits of the *Paracoccus* NDH-1. When the sequences from the *NQO4* (48 kDa) and the *NQO5* (21 kDa) subunits were compared with the known sequences from the bovine complex I (Table III and Figure 3), it was found that the *Paracoccus* 48- and 21-kDa subunits have significant amino acid sequence similarity, respectively, to the 49- and 30-kDa polypeptides of bovine complex I (Fearnley et al., 1989; Pilkington et al., 1991). Furthermore, the antibodies to the bovine 49- and 30-kDa polypeptides cross-reacted with the corresponding *Paracoccus* subunits (see Figure 4).

When the deduced amino acid sequence of the *NQO6* gene product was used to search the GenBank/EMBL databases, significant similarity was found between this gene product and the putative polypeptides encoded by the *psbG* gene of *Paramecium* mitochondrial DNA and by the *ndhK* gene (previously designated *psbG*) of chloroplast DNA (see Table III and Figure 3). Similar genes are also present in *Synechocystis*, and it has been shown immunochemically that these genes are expressed by this microorganism (Nixon et al., 1989; Berger et al., 1991). Recently, a polypeptide (M_r = 20 000) which is highly homologous to these polypeptides was found to be present in bovine complex I (Masui et al., 1991a). Although the amino acid sequence of this bovine 20K polypeptide has been only partially determined, it has striking similarity to the *Paracoccus NQO6* gene product (Figure 3). With respect to the *NQO7* gene, when its sequence was compared with the known sequences of the bovine complex I polypeptides, the data clearly indicated that the *NQO7* gene product is a counterpart of the mitochondrial ND3 gene product and its chloroplast homologues (Figure 3).

In general, a polypeptide encoded by a structural gene lying within a gene cluster (operon) is found to be an essential component of the enzyme complex coded for by this gene cluster. On this basis, the work reported here, demonstrating that the *NQO4*, *NQO5*, *NQO6*, and *NQO7* genes lies within the *Paracoccus* NDH-1 gene cluster, provides evidence that 48-kDa, 21-kDa, *NQO6*-encoded, and *NQO7*-encoded polypeptides are essential components of the *Paracoccus* NDH-1. This in turn suggests that their mitochondrial counterparts (49-kDa, 30-kDa, 20-kDa, and ND3 polypeptides in the case of bovine complex I) are also the essential components of the

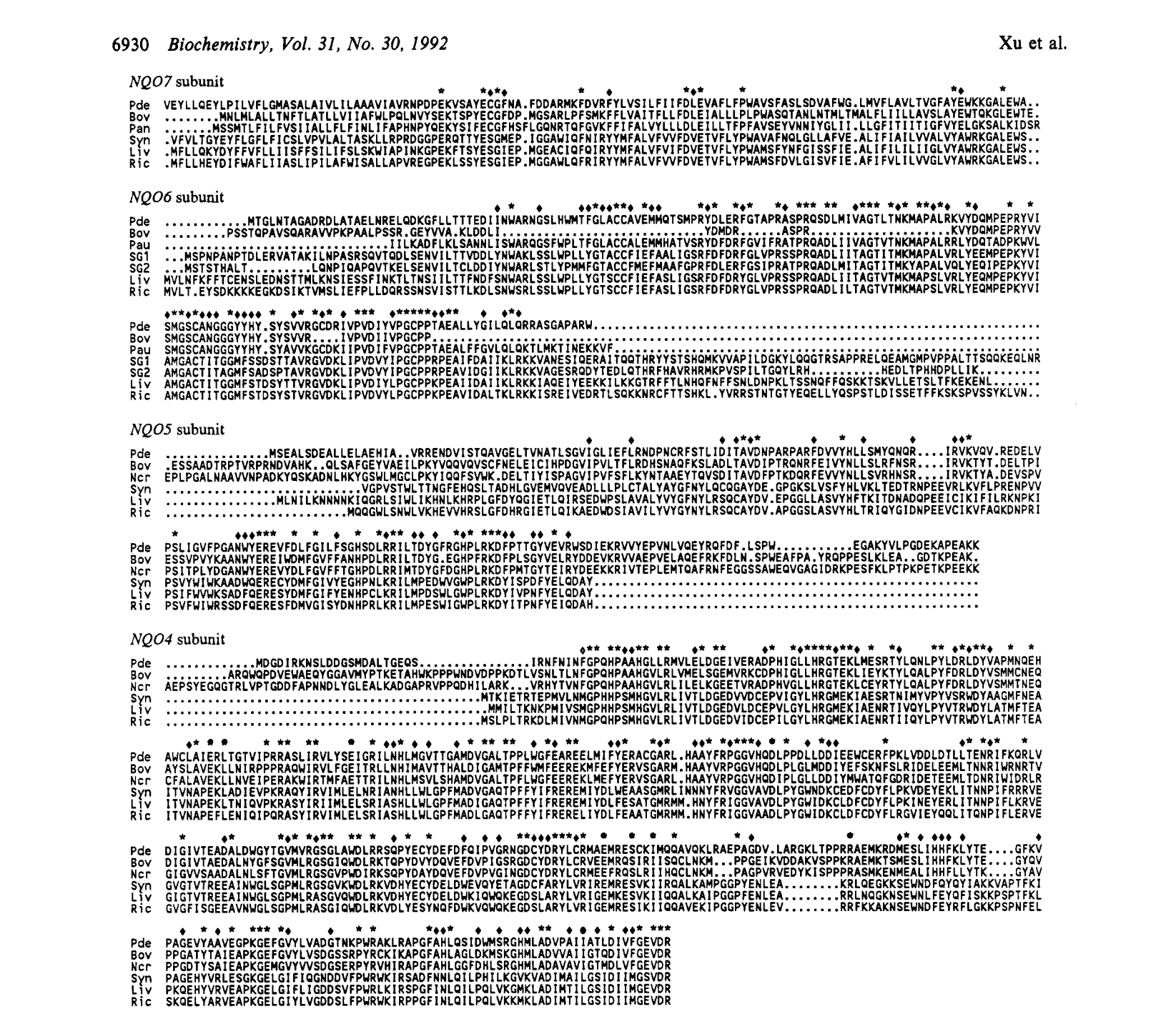

FIGURE 3: Comparison of the amino acid sequences of the *NQO7*, *NQO6*, *NQO5* (21 kDa), and *NQO4* (48-kDa) subunits with their homologues from various organisms. The comparison was conducted by the PILEUP program. The asterisks indicate amino acid residues conserved in all listed organisms. The diamonds indicate those amino acid residues which are conserved in the respiratory NDH-1 and in the predicted NDH-1 of chloroplasts. However, conservation is not seen between the respiratory and chloroplast sequences. The amino acid sequences cited are from the following references. The *NQO4* homologues: bovine (Fearnley & Walker, 1986); *N. crassa* (Preis et al., 1990); *Synechocystis* (Steinmüller, unpublished results, Accession No. x60650); liverwort (Ohyama et al., 1986); and rice (Hiratsuka et al., 1989). The *NQO5* homologues: bovine (Pilkington et al., 1991); *N. crassa* (Videira et al., 1990); *Synechocystis* (Steinmüller et al., 1989); liverwort (Ohyama et al., 1986); and rice (Hiratsuka et al., 1989). The *NQO6* homologues: bovine (Masui et al., 1991a); *Paramecium* (Pritchard et al., 1989); liverwort (Ohyama et al., 1986); rice (Hiratsuka et al., 1989); *Synechocystis* G1 (Steinmüller et al., 1989); and *Synechocystis* G2 (Mayes et al., 1990). In the case of *Synechocystis* PCC6803, two *psbG* genes have been isolated and sequenced (Steinmüller et al., 1989; Mayes et al., 1990). One is located in the operon containing the *ndhC* gene (designated *psbG1*), and the other is not present in or surrounding this operon (designated *psbG2*). The *NQO7* homologues: bovine (Anderson et al., 1982); *Podospora* (Cummings & Domenico, 1988); *Synechocystis* (Steinmüller et al., 1989); liverwort (Ohyama et al., 1986); and rice (Hiratsuka et al., 1989). In terms of the bovine IP 20-kDa polypeptide, only a partial amino acid sequence is available as described in text. In this figure, these partial primary structures are aligned according to Masui et al. (1991a).

mitochondrial NADH–Q oxidoreductase. Experiments using gene deletion mutants may confirm these matters.

Iron–Sulfur Clusters. Analysis of the *NQO4*, *NQO5*, *NQO6*, and *NQO7* sequences revealed that none of these gene products contain typical structural motifs associated with the iron–sulfur clusters found in ferredoxins of either bacteria or chloroplasts (Yasunobu & Tanaka, 1980). However, the isolated fraction composed of the bovine counterparts of these subunits (complex I IP 49-, 30-, and 13-kDa polypeptides) has been reported on the basis of the EPR studies (Ohnishi et al., 1985; Hatefi et al., 1985) to contain one tetranuclear and one binuclear iron–sulfur cluster. This bovine IP fraction can be further resolved into two subfractions: one consisting of the 49-kDa polypeptide and the other of the 30-kDa plus 13-kDa

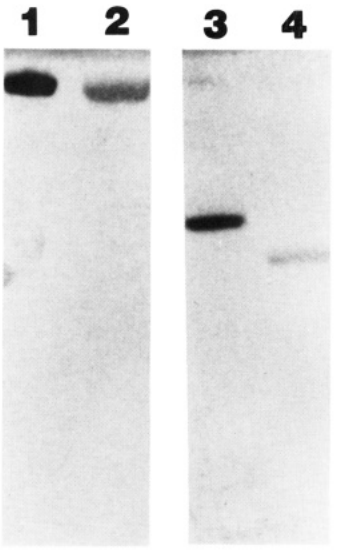

FIGURE 4: Cross-immunoreactivity between the *Paracoccus* 48- and 21-kDa subunits and the bovine 49- and 30-kDa polypeptides, respectively. A 12% SDS–polyacrylamide gel was prepared and loaded as follows: lanes 1–4 were loaded by 2 μg of bovine complex I, 0.5 μg of the purified *Paracoccus* 48-kDa subunit, 2 μg of bovine complex I, and 0.5 μg of the purified *Paracoccus* 21-kDa subunit, respectively. Subsequent to electrophoresis and transfer of the proteins to nitrocellulose membranes, the membranes were blotted with affinity-purified antibodies against the bovine 49-kDa (lanes 1 and 2) and 30-kDa (lanes 3 and 4) polypeptides as described under Materials and Methods. The detecting antibody used was an alkaline phosphatase conjugated anti-rabbit IgG.

polypeptides. Each subfraction has been reported to contain non-heme irons and acid-labile sulfides. In light of these data, the bovine 49-kDa polypeptide is expected to bear an iron–sulfur cluster. As seen in Figure 3, the bovine 49-kDa subunit contains six cysteines, of which four are needed to bind an iron–sulfur cluster. Only two of these cysteines (Cys-291 and Cys-297) are conserved in *Paracoccus* and *N. crassa*. However, although this number of cysteine residues is not enough to ligate iron–sulfur clusters, it should be taken into consideration that histidines may also be involved in binding of iron–sulfur clusters (Britt et al., 1991; Gurbiel et al., 1991). Sequence alignment showed that there are 11 histidines conserved among the bovine, *Paracoccus*, and *N. crassa* polypeptides (Figure 3). With respect to the other subfraction, the bovine 30-kDa polypeptide contains two cysteines and five histidines. Of these, no cysteines and only two histidines are conserved in *Paracoccus* and *N. crassa*, suggesting that this polypeptide by itself cannot bear the iron–sulfur cluster. In the case of the nitrogenase enzyme, it has been proposed that a tetranuclear iron–sulfur cluster may be conjugated utilizing residues from each of two polypeptides (Orme-Johnson, 1985; Hausinger & Howard, 1982). The bovine 13-kDa(A) and 13-kDa(B) polypeptides have been reported to have three and one cysteines and four and two histidines, respectively (Masui et al., 1991b). Therefore, we consider it a possibility that two or more of the 30-kDa, the 13-kDa(A), and the 13-kDa(B) polypeptides contribute ligands to the iron–sulfur cluster jointly.

Evidence for an NDH-1-Type Enzyme in Chloroplasts. Chloroplast DNA contains open reading frames encoding polypeptides homologous to the seven ND gene products of the mitochondrial complex I (Anderson et al., 1982; Chomyn et al., 1985, 1986; Ohyama et al., 1986). In addition, Dupuis et al. (1991) have reported that the bovine HP 23-kDa polypeptide has significant amino acid sequence identity to the predicted polypeptide encoded by the chloroplast *ndhI* (previously named *frxB*). Furthermore, as described above, the bovine IP 20-kDa polypeptide is homologous to the chloroplast *ndhK* (previously designated *psbG*) product (Masui et al., 1991a). On this basis, it has been suggested that an NDH-1 type enzyme may be located in the thylakoid membranes of chloroplasts. Our nucleotide sequence analyses showing that the arrangement of the *NQO7*, *NQO6*, and *NQO5* genes of the *Paracoccus* NDH-1 (Figure 1) corresponds with those of both chloroplast DNA and cyanobacteria (Steinmüller et al., 1989) support the possible presence of NDH-1-type enzymes in chloroplasts and cyanobacteria. These data also suggest that the *Paracoccus* NDH-1 might have an evolutionary relationship with the predicted enzyme complex of chloroplasts and cyanobacteria.

In light of the above, it is of interest to compare the mitochondrial complex I with the predicted NDH-1-like enzyme of chloroplasts in terms of amino acid sequences. As seen in Figure 3, unique differences in individual amino acid residues of the *NQO4*, *NQO5*, *NQO6*, and *NQO7* homologues have been detected between mitochondrial complex I and the putative NDH-1 of chloroplasts. In addition, there are a number of differences between complex I and chloroplast NDH-1 subunits in terms of stretches of amino acid sequence. Thus, for example, among the *NQO4* homologues, the mitochondrial subunits contain the eight amino acid stretch (x_{308}-DxKxSPP$_{315}$) which is absent in the chloroplast subunits. On the other hand, the four amino acid sequence (KPSP) present in the chloroplast subunits is not found in their mitochondrial counterparts. The *NQO5* homologues of mitochondria lack the sequence (xPEE) found in chloroplast homologues, whereas the mitochondrial homologues contain a stretch of approximately 40 amino acid residues at the C-terminal region which are not found in the chloroplasts. Finally, in terms of the *NQO6* homologues, the chloroplast subunits have a stretch of approximately 40 amino acids at the C-terminal region which are not found in the mitochondrial subunits. On the basis of these unique characteristics, it is clear that the subunits of *Paracoccus* NDH-1 are more closely related to the mitochondrial complex I than to the putative chloroplast NDH-1. On the contrary, subunits of *Synechocystis* NDH-1-type enzyme appear to be more akin to their chloroplast NDH-1 homologues than to their mitochondrial complex I counterparts (see Table III and Figure 3).

ACKNOWLEDGMENT

We thank Dr. Youssef Hatefi (The Scripps Research Institute, La Jolla) for the antisera to the bovine 49- and 30-kDa subunits and the bovine complex I and Daniel G. Owens, Nanette C. Wong, and Tri M. Dinh for their excellent technical assistance and Dr. Carla Hekman (ENZON, New Jersey) for critical reading of the manuscript.

REFERENCES

Albracht, S. P. J., van Verseveld, H. W., Hagen, W. R., & Kalkman, M. L. (1980) *Biochim. Biophys. Acta 593*, 173–186.

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.

Berger, S., Ellersiek, U., & Steinmüller, K. (1991) *FEBS Lett. 286*, 129–132.

Britt, R. D., Sauer, K., Klein, M. P., Knaff, D. B., Kriauciunas, A., Yu, C.-A., Yu, L., & Malkin, R. (1991) *Biochemistry 30*, 1892–1901.
Chomyn, A., Mariottini, P., Cleeter, M. W. J., Ragan, C. I., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R. F., & Attardi, G. (1985) *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.
Cummings, D. J., & Domenico, J. M. (1988) *J. Mol. Biol. 204*, 815–839.
Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.
Dupuis, A., Skehel, J. M., & Walker, J. E. (1991) *Biochemistry 30*, 2954–2960.
Fearnley, I. M., & Walker, J. E. (1986) *EMBO J. 5*, 2003–2008.
Fearnley, I. M., Runswick, M. J., & Walker, J. E. (1989) *EMBO J. 8*, 665–672.
George, C. L., Ferguson, S. J., Cleeter, M. W. J., & Ragan, C. I. (1986) *FEBS Lett. 198*, 135–139.
Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S., & Stormo, G. (1981) *Annu. Rev. Microbiol. 35*, 365–403.
Gornall, A. G., Bardawill, C. J., & David, M. M. (1949) *J. Biol. Chem. 177*, 751–766.
Gurbiel, R. J., Ohnishi, T., Robertson, D. E., Daldal, F., & Hoffman, B. M. (1991) *Biochemistry 30*, 11579–11584.
Han, A.-L., Yagi, T., & Hatefi, Y. (1989) *Arch. Biochem. Biophys. 275*, 166–173.
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.
Hausinger, R. P., & Howard, J. B. (1982) *J. Biol. Chem. 257*, 2438–2490.
Hekman, C., & Hatefi, Y. (1991) *Arch. Biochem. Biophys. 284*, 90–97.
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.
Husain, I., Van Houten, B., Thomas, D. C., & Sancar, A. (1986) *J. Biol. Chem. 261*, 4895–4901.
John, P., & Whatley, F. R. (1977) *Biochim. Biophys. Acta 463*, 129–153.
Kurowski, B., & Ludwig, B. (1987) *J. Biol. Chem. 262*, 13805–13811.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265–275.
Masui, R., Wakabayashi, S., Matsubara, H., & Hatefi, Y. (1991a) *J. Biochem. (Tokyo) 110*, 575–582.
Masui, R., Wakabayashi, S., Matsubara, H., & Hatefi, Y. (1991b) *J. Biochem. (Tokyo) 109*, 534–543.
Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035–10038.
Mayes, S. R., Cook, K. M., & Barber, J. (1990) *FEBS Lett. 262*, 49–54.
Meinhardt, S. W., Kula, T., Yagi, T., Lillich, T., & Ohnishi, T. (1987) *J. Biol. Chem. 262*, 9147–9153.
Nixon, P. J., Gounaris, K., Coomber, S. A., Hunter, C. N., Dyer, T. A. & Barber, J. (1989) *J. Biol. Chem. 264*, 14129–14135.
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.
Orme-Johnson, W. H. (1985) *Annu. Rev. Biophys. Biophys. Chem. 14*, 419–459.
Pilkington, S. J., Skehel, J. M., & Walker, J. E. (1991) *Biochemistry 30*, 1901–1908.
Preis, D., Van der Pas, J. C., Nehls, U., Röhlen, D.-A., Sackmann, U., Jahnke, U., & Weiss, H. (1990) *Curr. Genet. 18*, 59–64.
Pritchard, A. E., Venuti, S. E., Ghalambor, M. A., Sable, C. L., & Cummings, D. J. (1989) *Gene 78*, 121–134.
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.
Shiota, S., & Nakayama, H. (1989) *Mol. Gen. Genet. 217*, 332–340.
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. (Pers. Ed.) 5*, 164–166.
Tran-Betcke, A., Warnecke, U., Bocker, C., Zaborosch, C., & Friedrich, B. (1990) *J. Bacteriol. 172*, 2920–2929.
Videira, A., Tropschug, M., & Werner, S. (1990) *Biochem. Biophys. Res. Commun. 171*, 1168–1174.
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*, 6422–6428.
Xu, X., Matsuno-Yagi, A., & Yagi, T. (1991b) *Biochemistry 30*, 8678–8684.
Xu, X., Matsuno-Yagi, A., & Yagi, T. (1992) *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) *Tanpakushitsu Kakusan Koso 34*, 351–363.
Yagi, T. (1990) *Arch. Biochem. Biophys. 281*, 305–311.
Yagi, T. (1991) *J. Bioenerg. Biomembr. 23*, 211–225.
Yagi, T. (1992) *Biochim. Biophys. Acta* (in press).
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. (1991) *Biol. Chem. Hoppe Seyler 372*, 555.
Yagi, T., Xu, X., & Matsuno-Yagi, A. (1992) *Biochim. Biophys. Acta* (in press).
Yasunobu, K. T., & Tanaka, M. (1980) *Methods Enzymol. 69*, 228–238.
Yun, C.-H., Beci, R., Crofts, A. R., Kaplan, S., & Gennis, R. B. (1990) *Eur. J. Biochem. 194*, 399–411.