Newman, A. K., Rubin, R. A., Kim, S.-H., & Modrich, P. (1981) *J. Biol. Chem. 256*, 2131-2139.
Newman, P. C., Nwosu, V. U., Williams, D. M., Cosstick, R., Seela, F., & Connolly, B. A. (1990a) *Biochemistry 29*, 9891-9901.
Newman, P. C., Williams, D. M., Cosstick, R., Seela, F., & Connolly, B. A. (1990b) *Biochemistry 29*, 9902-9910.
Pingoud, A., Alves, J., & Geiger, R. (1991) in *Enzymes in molecular biology* (Burrell, M. M., Ed.) Humana Press, Clifton, NJ (in press).
Posfai, J., Bahwat, H. S., Posfai, G., & Roberts, R. J. (1989) *Nucleic Acids Res. 17*, 2421-2435.
Roberts, R. J. (1990) *Nucleic Acids Res. 18*, Suppl., 2331-2365.
Schildkraut, I., Banner, C. D., Rhodes, C. S., & Parekh, S. (1984) *Gene 27*, 327-329.
Smith, H. O., Annau, T. M., & Chandrasegaran, S. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 826-830.
Smith, T. F., & Waterman, M. S. (1981) *Adv. Appl. Math. 2*, 482-489.
Stephenson, F. H., Ballard, B. T., Boyer, H. W., Rosenberg, J. M., & Greene, P. (1989) *Gene 85*, 1-13.
Vershon, A. K., Youderian, P., Weiss, M. A., Susskind, M. M., & Sauer, R. T. (1985) in *Sequence specificity in transcription and translation*, pp 209-218, Alan R. Liss, New York.
Wolfes, H., Alves, J., Fliess, A., Geiger, R., & Pingoud, A. (1986) *Nucleic Acids Res. 14*, 9063-9080.

The NADH-Binding Subunit of the Energy-Transducing NADH-Ubiquinone Oxidoreductase of *Paracoccus denitrificans*: Gene Cloning and Deduced Primary Structure†,‡

Xuemin Xu, Akemi Matsuno-Yagi, and Takao Yagi*
Division of Biochemistry, Department of Molecular and Experimental Medicine, Research Institute of Scripps Clinic, La Jolla, California 92037

Received February 4, 1991; Revised Manuscript Received April 3, 1991

ABSTRACT: The NADH dehydrogenase complex isolated from *Paracoccus denitrificans* is composed of approximately 10 unlike polypeptides and contains noncovalently bound FMN, non-heme iron, and acid-labile sulfide [Yagi, T. (1986) *Arch. Biochem. Biophys. 250*, 302-311]. The NADH-binding subunit (M_r = 50000) of this enzyme complex was identified by direct photoaffinity labeling with [^{32}P]NADH [Yagi, T., & Dinh, T. M. (1990) *Biochemistry 29*, 5515-5520]. Primers were synthesized on the basis of the N-terminal amino acid sequence of this polypeptide, and these primers were used to synthesize an oligonucleotide probe by the polymerase chain reaction. This probe was utilized to isolate the gene encoding the NADH-binding subunit from a genomic library of *P. denitrificans*. The nucleotide sequence of the gene and the deduced amino acid sequence of the entire NADH-binding subunit were determined. The NADH-binding subunit has 431 amino acid residues and a calculated molecular weight of 47 191. The encoded protein contains a putative NAD(H)-binding and an iron-sulfur cluster-binding consensus sequence. The deduced amino acid sequence of the *Paracoccus* NADH-binding subunit shows remarkable similarity to the α subunit of the NAD-linked hydrogenase of *Alcaligenes eutrophus* H16. When partial DNA sequencing of the regions surrounding the gene encoding the NADH-binding subunit was carried out, sequences homologous to the 24-, 49-, and 75-kDa polypeptides of bovine complex I were detected, suggesting that the structural genes of the *Paracoccus* NADH dehydrogenase complex constitute a gene cluster.

The energy-transducing NADH-quinone (Q)[1] oxidoreductase (complex I or NDH-1) in the respiratory chains of mitochondria and bacteria contains a noncovalently bound FMN and multiple iron-sulfur clusters. The enzyme catalyzes electron transfer between NADH and Q and leads to the concomitant formation of a proton gradient across the membrane (Hatefi et al., 1985; Hatefi, 1985; Ragan, 1987; Yagi, 1989, 1991). The NADH-Q oxidoreductase of mitochondria is composed of more than 25 unlike polypeptides (Tuschen et al., 1990; Hatefi, 1985; Ragan, 1987) and appears to have the most intricate structure of any known membrane-bound protein complex. This complexity has hampered progress in studies on many aspects of the structure and mechanism of action of this enzyme complex (Yagi, 1991). However, such studies have taken on a greater significance in recent years since there has been an increasing number of reports on human mitochondrial diseases involving structural and functional defects at the level of this enzyme complex (Morgan-Hughes et al., 1990; Wallace et al., 1988). It has been recently suggested that Parkinson's and Huntington's diseases might also

† This work was supported by US Public Health Service Grant GM33712 to T.Y. and MO1 RR00833 to General Clinical Research Center. This is publication 6657-MEM from the Research Institute of Scripps Clinic, La Jolla, CA.
‡ The nucleotide sequence of the *Paracoccus* NADH-binding subunit gene given in this paper has been submitted to GenBank under Accession Number M64432.
* To whom correspondence should be addressed.

[1] Abbreviations: Q, quinone; UQ, ubiquinone; NDH-1 or complex I, energy-transducing NADH-quinone oxidoreductase; NDH-2, NADH-quinone oxidoreductase lacking energy coupling site; complex III, ubiquinol-cytochrome *c* oxidoreductase; bp, base pairs; FP, IP, and HP, respectively, the flavoprotein, the iron-sulfur protein, and the hydrophobic protein fractions of complex I; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; PVDF, poly(vinylidene difluoride).

0006-2960/91/0430-6422$02.50/0

be caused by functional defects in complex I of mitochondria (Parker et al., 1989, 1990).

Aerobically grown *Paracoccus denitrificans* expresses a mitochondrial-type respiratory chain (Stouthamer, 1980). The NADH-Q oxidoreductase segment of *Paracoccus* membranes has been shown to bear an energy coupling site and to have considerable similarity to its mitochondrial counterpart in terms of electron carriers (Meinhardt et al., 1987; Stouthamer, 1980). However, in contrast to the mitochondrial enzyme, the NADH dehydrogenase complex purified from *Paracoccus* membranes is composed of approximately 10 unlike polypeptides (Yagi, 1986), suggesting that the *Paracoccus* NADH-Q oxidoreductase is structurally simpler than its mitochondrial counterpart. Therefore, the *Paracoccus* NADH dehydrogenase complex appears to provide a useful model system for studying the structure and mechanism of action of the mitochondrial NADH-Q oxidoreductase.

Another advantage of using a bacterial system is the high probability that the structural genes of a particular enzyme complex constitute an operon. This has been illustrated for other respiratory chain enzyme complexes (Kurowski & Ludwig, 1987; Ishizuka et al., 1990) and also for the ATP synthases (Walker et al., 1984). In the mitochondrial system, it is difficult to demonstrate conclusively that a polypeptide is a component of a particular enzyme complex rather than a copurified contaminant unless functional and/or structural justification is provided. However, this difficulty is eliminated when the genes of the enzyme complex constitute an operon, since a polypeptide encoded by a structural gene in an operon is generally found to be an essential component of the enzyme complex coded by this operon (Tzagoloff et al., 1990; Nobrega et al., 1990).

The identification of the NADH-binding subunit of the *Paracoccus* NDH-1 was recently accomplished in this laboratory by direct photoaffinity labeling with [^{32}P]NADH (Yagi & Dinh, 1990). The amino acid composition was found to be very similar to that of the bovine enzyme (Galante & Hatefi, 1979). Furthermore, the cross-immunoreactivity between the *Paracoccus* and the bovine NADH-binding proteins and their similarity in apparent molecular weight, as estimated from SDS–PAGE, suggested that the primary structures of these two enzymes might show considerable sequence identity (Galante & Hatefi, 1979; Xu & Yagi, 1991). A prerequisite for understanding the structure and mechanism of such an intricate enzyme complex as the NADH-Q oxidoreductase is a knowledge of the primary structure of each subunit. Therefore, we attempted to clone and sequence the gene encoding the NADH-binding subunit of *Paracoccus* NDH-1 in the hope that it might provide information regarding both the structure of this enzyme complex and also the possible presence of an operon carrying the genes of the complex.

This paper describes the cloning of the structural gene of the NADH-binding subunit of the *Paracoccus* NADH-UQ oxidoreductase. The complete nucleotide sequence of the gene and the deduced primary amino acid sequence were determined. The deduced primary structure contains the motif GxGxxGxxxG (where x is any amino acid), which has been found in NAD(H)-binding proteins. The consensus sequence CxxCxxC, which is involved in tetranuclear Fe-S clusters from chloroplast and bacterial ferredoxin, was also present (Yasunobu & Tanaka, 1980). The predicted amino acid sequence was found to be similar to that of the α subunit of the NAD-reducing hydrogenase of *Alcaligenes eutrophus* H16 (Tran-Betcke et al., 1990), and the evolutionary relationship of these enzymes is discussed. Furthermore, when partial DNA sequencing of the nearby regions was performed, sequences homologous to the 24-, 49-, and 75-kDa polypeptides of bovine complex I were found, suggesting that the structural genes of the *Paracoccus* NADH dehydrogenase complex construct at least one gene cluster.

Materials and Methods

Isolation of the NADH-Binding Subunit of the Paracoccus NADH Dehydrogenase Complex. *P. denitrificans* (ATCC13548) cells were grown aerobically with glucose as a substrate, as described previously (Yagi, 1986). The *Paracoccus* NADH dehydrogenase complex was purified according to Yagi (1986). A total of 50 μg of the purified *Paracoccus* NADH dehydrogenase complex was loaded into each of 10 wells (1-cm width) of a sodium dodecyl sulfate–polyacrylamide gel (15 × 14 × 0.15 cm) composed of 10% acrylamide. Electrophoresis was conducted as described by Laemmli (1970) except the running buffer contained 0.1 mM thioglycolic acid. The gel was electrophoresed for 3 h at 30 mA and then stained for 30 min at room temperature in a solution containing 0.05% Coomassie Brilliant Blue R-250, 25% 2-propanol, and 10% acetic acid. Destaining was carried out for 2–3 h at 4 °C in 5% acetic acid containing 16.5% methanol. The band corresponding to the NADH-binding subunit (M_r = 50 000) was assigned as reported previously (Yagi & Dinh, 1990), excised from the gel, and washed several times with water. The subunit was then electroeluted as described by Hunkapiller et al. (1983). The purity of the isolated subunit was then enhanced by a second cycle of electrophoresis followed by electroelution conducted as described above.

Protein Sequence Analysis. The N-terminal amino acid sequence of the *Paracoccus* NADH-binding subunit was determined as follows. Samples of the *Paracoccus* NADH-binding subunit (10 μg) were lyophilized, dissolved in Laemmli sample buffer containing 80 mM Tris-HCl (pH 6.8), 6% SDS, 5% 2-mercaptoethanol, 0.005% bromphenol blue, and 20% glycerol and applied to a 10% polyacrylamide slab (15 × 14 × 0.15 cm) gel. Following electrophoresis (Laemmli, 1970) the subunit was electrophoretically transferred to poly(vinylidene difluoride) (PVDF) membranes (Immobilon, Millipore) as described by Matsudaira (1987). The transferred protein on the PVDF membranes was sequenced in an Applied Biosystems 470A gas-phase protein sequencer.

The amino acid sequence corresponding to an internal segment of the NADH-binding subunit was determined as follows. The isolated subunit (25 μg) was incubated with 1 mg/mL CNBr in 100 μL of 70% formic acid for 24 h at room temperature. Following this incubation, the sample was mixed with 900 μL of H_2O and lyophilized. The CNBr-digested peptides were separated by SDS–PAGE on a 16% polyacrylamide gel (Schägger & von Jagow, 1987). Sequencing of the individual peptides was carried out as described above.

Synthesis of the Oligonucleotide Probe. On the basis of the amino acid sequences 1–7 and 27–33 of the NADH-binding subunit, degenerate sense and antisense primers, respectively, were synthesized as presented under Results. The primers were purified by oligonucleotide purification cartridges (Applied Biosystems). PCR amplification of the portion of the gene encoding the N-terminal sequence of the NADH-binding subunit from *Paracoccus* DNA was performed with a Perkin-Elmer Cetus Instruments thermal cycler by modification of a procedure described previously (Sharma et al., 1990). Briefly, DNA sequences were amplified with *Taq* polymerase in a 100-μL reaction mixture containing 700 ng of *Paracoccus* DNA, 34 mM Tris-HCl (pH 8.8), 8.3 mM $(NH_4)_2SO_4$, 3.4 mM $MgCl_2$, 5% dimethyl sulfoxide, 5 mM

1,4-dithiothreitol, 100 μg/mL gelatin, 500 μM dNTPs (dATP, dCTP, dGTP, and dTTP), 1 μg of both oligonucleotide primers, and 5 units of *Taq* polymerase. The samples were subjected to 35 amplification cycles each involving denaturation at 92 °C for 60 s, annealing at 58 °C for 60 s, and extension from the primers at 72 °C for 90 s. The amplification products were separated on 6.25% polyacrylamide gels. Three major bands (40, 100, and 400 bp) were observed following electrophoresis and staining with ethidium bromide. The amplification product of the predicted size (100 bp) was isolated from the gels and subcloned into the *Eco*RV site of pBluescript II KS(+) (Stratagene). The sequence of the PCR product was determined with use of a T7 DNA sequence kit (Pharmacia). Seven clones inserted by the PCR product were chosen, and their DNA sequences were determined. All of these clones contained the same DNA sequence. The sequence encoded a protein sequence that matched that expected from the N-terminal amino acid sequence (residues 1–33) of the purified NADH-binding subunit (see Results). A fragment of one of the clones was used as a probe for further studies. This oligonucleotide probe was radiolabeled with use of an oligolabeling kit (Pharmacia) and [α-^{32}P]dATP.

Construction and Screening of the Paracoccus Genomic Library. General cloning techniques were carried out essentially as described by Sambrook et al. (1989). Genomic DNA from *P. denitrificans* was isolated, digested with *Eco*RI, and subjected to agarose electrophoresis. The region (4–8 kbp) of the agarose gel that reacted with the PCR product probe was excised, and the DNA fragments were isolated. The *P. denitrificans* genomic library was prepared for ligation into the *Eco*RI site of the λ ZAPII cloning vector and titered with use of the XL1-Blue strain of *Escherichia coli*, yielding a recombinant frequency of 90%.

Plaques were lifted onto Hybond-N filters (Amersham Corp.), denatured in 0.5 N NaOH/1.5 M NaCl for 30 s, neutralized in 0.5 M Tris-HCl (pH 8.0)/1.5 M NaCl for 5 min, and rinsed with 2× SSPE (where 20× SSPE is 3.6 M NaCl/200 mM sodium phosphate (pH 7.7)/20 mM EDTA). The filters were then prehybridized in a solution consisting of 5× Denhardt's solution (where 50× Denhardt's solution is 1.0% (w/v) Ficoll/1.0% poly(vinylpyrrolidone)/1.0% bovine serum albumin), containing 5× SSPE, 0.5% SDS, and 20 μg/mL sonicated salmon sperm DNA at 61 °C for 1 h. Hybridization was allowed to proceed overnight at 61 °C in the prehybridizing solution. Following hybridization, the filters were washed in 1× SSPE containing 0.1% SDS for 10 min and then were washed for 30 min at room temperature and were washed twice in 0.1× SSPE containing 0.1% SDS at 61 °C for 30 min each time. Autoradiography was carried out overnight with Kodak XAR-5 X-ray films and intensifying screens. Positive plaques were selected as detailed under Results.

DNA Sequencing Strategy. The pBluescript II phagemid containing the cloned DNA insert was excised from cloned λ ZAPII by coinfection with the helper phage R408. The DNA insert was digested separately with *Pst*I, *Sal*I, *Sac*II, and *Apa*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) with the T7 sequencing kit from Pharmacia. Deletion mutants for nucleotide sequencing were created by use of the *Exo*III/mung bean deletion kit from Stratagene. The universal primers, T3 and T7, and internal unique primers 18 bases in length were used in these experiments. When compression regions were encountered, 7-deaza-dGTP was used instead of dGTP (Sambrook et al., 1989).

Nucleotide and Protein Sequence Analysis. The University of Wisconsin Genetics Computer Group's computer software programs were used to analyze the sequence data (Devereux et al., 1984). The comparison of polypeptides was performed by the BESTFIT software program. The FASTA program was used to search the GenBank/EMBL Sequence databases for sequences homologous to the *Paracoccus* NADH-binding subunit.

Results

Determination of N-Terminal and Internal Sequences of the NADH-Binding Subunit. The NADH-binding subunit was purified from SDS–polyacrylamide gels of the *Paracoccus* NADH dehydrogenase complex by two successive electroelutions (Yagi & Hatefi, 1988). The purified subunit was subjected to Laemmli SDS–polyacrylamide gel electrophoresis and transferred to PVDF membranes for protein sequencing according to the method of Matsudaira (1987). The 37 N-terminal amino acid residues shown were determined as described below.

```
MLNDQDRIFTNLYGMGDRSLAGAKKRGHWDGTAAIIQ
  (I)                          (II)
```

In order to obtain additional sequence information on the NADH-binding subunit, the purified subunit was cleaved by CNBr. The CNBr peptides were separated by SDS–polyacrylamide gel electrophoresis and transferred to PVDF membranes as described above. From these peptides, one internal sequence (PPFPAGAGLY) was obtained.

Cloning of the Structural Gene for the NADH-Binding Subunit. The amino acid sequences corresponding to residues 1–7 (MLNDQDR) and 27–33 (HWDGTA) were used to design two mixtures of oligonucleotides 20 and 17 bases in length containing 256 and 64 different sequences, respectively.

```
   Sense primer                Antisense primer
5'ATGTTAAATGATCAAGATCG 3'   3'GTAACCCTACCATGACG 5'
    C T  C  C  G  CA            G     G  T  T
      G                                G  G
      C                                C  C
         (I)                        (II)
```

The degenerate sense and antisense oligonucleotides shown were synthesized and used as primers for PCR amplification. The PCR product was subcloned into the *Eco*RV site of pBluescript II KS. Seven independent clones harboring a PCR product of approximately 100 bp were identified and sequenced. Each clone was found to encode the N-terminal amino acid sequence of the NADH-binding subunit as follows.

```
ATGCTGAATGATCAGGATCGGATCTTTACGAACCTTTACGGCATGGGCGAC
                     I  F  T  N  L  Y  G  M  G  D

CGCAGCCTGGCCGGCGCCAAAAAGCGCGGCCATTGGGACGGCACTGC
 R  S  L  A  G  A  K  K  R  G
```

One of the clones was radiolabeled and used as a hybridization probe for a *Paracoccus* genomic λ ZAPII library (approximately 10^4 phages/plate were screened). We obtained 200 positive clones from the 2.5×10^5 plaques of the library. Ten of these were subjected to restriction enzyme mapping. All 10 clones were found to be identical and to contain an *Eco*RI fragment of 5.7 kbp. Therefore, one clone (designated pXT-1) was selected and used in all subsequent analyses.

DNA Sequence Analysis. Figure 1 displays the restriction map of the *Eco*RI 5.7-kbp fragment as well as the sequencing strategy used. Digestion of the *Eco*RI 5.7-kbp DNA fragment from pXT-1 with *Sal*I released three *Sal*I fragments. Of these three fragments, one (1.3 kbp) reacted with the PCR product

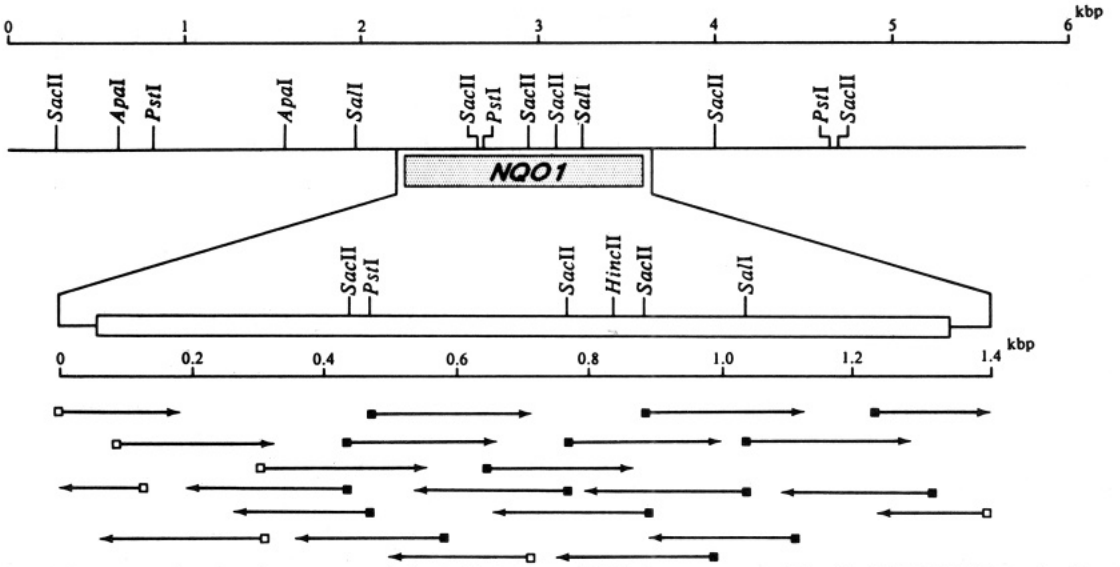

FIGURE 1: Restriction map and sequencing strategy of the *P. denitrificans* DNA fragment containing the NADH-binding subunit coding region. The *NQO1* indicates the structural gene of the NADH-binding subunit of the *Paracoccus* NDH-1. The direction and extent of the nucleotide sequence are indicated by the length of the arrows. The universal and synthetic oligonucleotide primers (18mer) used for sequencing are indicated by ■ and □, respectively. Restriction fragments were ligated into the linker region of pBluescript KS plasmid. *Exo*III deletion was used to generate several deletion mutants from the *Sal*I and *Sac*II fragments.

probe and was isolated. Within the *Sal*I 1.3-kbp fragment, an open reading frame of 330 codons appeared to encode the NADH-binding subunit but was truncated at the 3′ end as shown in Figures 1 and 2. The remaining were found in the overlapping *Sac*II fragment. As shown in Figure 2, the sequence (1293 bases) codes for 431 amino acid residues with a calculated molecular weight of 47 191, which approximates the *Paracoccus* NADH-binding subunit with M_r = 50 000 estimated by SDS–polyacrylamide gel electrophoresis (Yagi & Dinh, 1990). The putative Shine–Dalgarno sequence (Shine & Dalgarno, 1975) for ribosome binding (AAGGA) was found 8 base pairs upstream from the initiation codon, as shown in Figure 2. This sequence is the same as the putative ribosome-binding sequence of the subunit I gene of *Paracoccus* cytochrome oxidase (Raito et al., 1987).

The G-C content of the NADH-binding subunit gene is high (65.7%). This is mainly due to codon bias, especially bias at the third position of the codons. High G-C content is commonly observed in the *Paracoccus* genomes (Horváth et al., 1990; Van Spanning et al., 1990a,b). Among the Gram-negative bacteria, *Rhodopseudomonas*, *Rhodospirillum*, *Agrobacterium*, and *Acromobacter* genomes also show a high G-C content (60–70%) (Osawa et al., 1990). Osawa et al. (1990) have suggested that the G-C content of bacteria is related to phylogeny rather than thermophilicity. If this is the case, *Paracoccus* may have an evolutionary relationship to photosynthetic bacteria (*Rhodopseudomonas* and *Rhodospirillum*).

Structure of the NADH-Binding Subunit. The amino acid composition deduced from the structural gene agreed with the results of the amino acid analysis (data not shown). Furthermore, the isoelectric point of 6.5 calculated from the deduced amino acid composition appears to correspond with the results of the isoelectric focusing gel (Yagi & Dinh, 1990). In agreement with the fact that the NADH-binding subunit isolated from bovine complex I is water-soluble (Ragan et al., 1982), the hydropathy plot of the *Paracoccus* NADH-binding subunit does not show a hydrophobic stretch typical for membrane-spanning regions (data not shown).

Putative Prosthetic Group-Binding Sites. Common to the NAD(H)-binding domains of many enzymes is a $\beta\alpha\beta$ fold, centered around the highly conserved sequence GxGxxG that constitutes a tight turn at the end of the first strand of a β sheet and marks the beginning of the suceeding α helix. This GxGxxG sequence occurs at residues 54–59, 199–204, and 271–276 of the *Paracoccus* NADH-binding subunit. On the basis of the predicted occurrence of the ADP-binding $\beta\alpha\beta$ fold in proteins by Wierenga et al. (1986), the scores of these three regions were calculated. Unfortunately, there is no full mark sequence in the deduced amino acid sequence of the *Paracoccus* NADH-binding subunit. However, comparison of the scores of the three sequences suggested that the region containing residues 54–59 is most likely the NADH-binding site. In addition, the analysis of consensus sequences of the NADH-binding site by Scrutton et al. (1990) also suggests that residues 54–59 constitute the most likely candidate for the NADH-binding site since the sequence G_{54}xGxxGxxxG_{63} of the *Paracoccus* NADH-binding subunit was also observed in the NADH-binding site of the dihydrolipoamide dehydrogenases from various organisms (Scrutton et al., 1990). Furthermore, as discussed below (see Discussion), the sequence G_{54}xGxxGxxxG_{63} is conserved in the α subunit (NADH-binding subunit) of the NAD-linked dehydrogenase from *A. eutrophus* H16. However, neither G_{199}xGxxG_{204} nor G_{271}xGxxG_{276} is conserved in the α subunit.

The predicted primary structure of the *Paracoccus* NADH-binding subunit contains 12 cysteine residues and the motif C_{347}xxC_{350}xxC_{353}, which is associated with tetranuclear Fe-S clusters from ferredoxins of bacteria and chloroplasts (Yasunobu & Tanaka, 1980). The NADH-Q oxidoreductase segment of *Paracoccus* contains four EPR-visible Fe–S centers (Meinhardt et al., 1987). The question arises as to which Fe-S center is associated with this consensus sequence in the NADH-binding subunit. Bovine complex I can be resolved into a water-soluble fraction and a water-insoluble fraction (HP) (Hatefi et al., 1985; Hatefi, 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. EPR studies of FP reveal two signals whose relaxation behavior is consistent with a tetranuclear cluster and a binuclear cluster, in agreement with the chemically determined iron content. FP is composed of three subunits

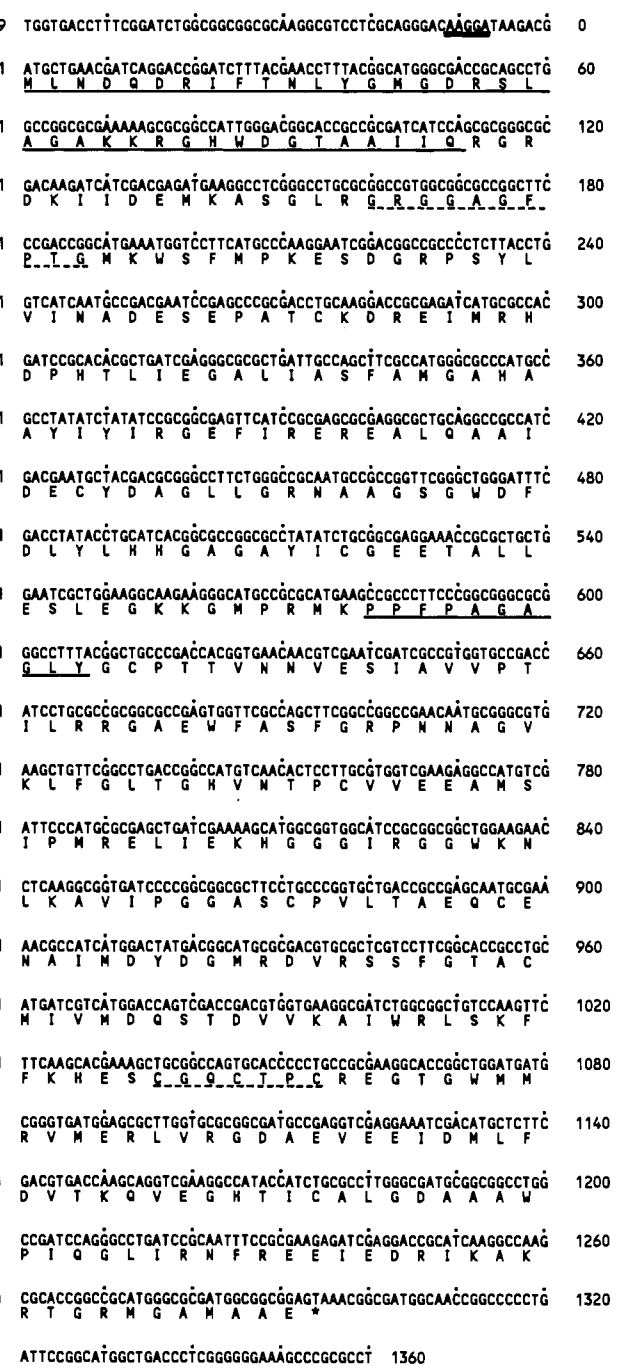

FIGURE 2: Nucleotide sequence of the *Paracoccus* NADH-binding subunit gene. The amino acid sequences determined by protein analysis are underlined. The putative Shine–Dalgarno sequence is double-underlined. The putative NADH-binding and iron–sulfur cluster binding sites are marked by broken lines.

(51, 24, and 9 kDa). The 51-kDa subunit has been reported to be the NADH-binding subunit of bovine complex I by Chen and Guillory (1981). Separation of the 51-kDa subunit from the other two (24 and 9 kDa) can be achieved by sodium trichloroacetate (Ragan et al., 1982). Both FP subfragments contain iron and acid-labile sulfides. In view of the suggested interaction between FMN and cluster N3, and the likelihood that FMN is the primary oxidant of NADH, both FMN and cluster N3 of bovine complex I can be provisionally assigned to the 51-kDa subunit (Ragan, 1987). Meinhardt et al. (1987) have reported that cluster N3 of *Paracoccus* is tetranuclear and similar to its bovine counterpart in terms of g and E_m values. Therefore, it seems likely that the consensus sequence $C_{347}xxC_{350}xxC_{353}$ may be the site of cluster N3. However, this assignment must be more thoroughly investigated in the future.

Discussion

This paper describes the cloning and sequencing of the gene of the NADH-binding subunit of the *P. denitrificans* NADH-UQ oxidoreductase. As described in the introduction, the cross-immunoreactivity and the similarity in the amino acid composition and molecular mass in the *Paracoccus* and mammalian NADH-binding proteins imply that considerable sequence identity may be present in the gene sequences of this subunit in these species (see below). Although bacterial energy-transducing NADH-Q oxidoreductases have been isolated from three sources other than *Paracoccus* [*Thermus thermophilus* HB-8 (Yagi et al., 1988), *E. coli* (Hayashi et al., 1990), and *Vibrio alginolyticus* (Hayashi et al., 1989)], the relationship between the species cannot yet be determined because nothing is known about their primary sequences.

As described previously (Yagi, 1989, 1991), in bacteria and plant and fungal mitochondria, another type of NADH-Q oxidoreductase known as NDH-2 is present. The NDH-2 lacks an energy coupling site, is composed of a single polypeptide, contains FAD as a prosthetic group, and is devoid of iron–sulfur clusters. The DNA sequences of the NDH-2 gene of *E. coli* (Young et al., 1981) and alkalophilic *Bacillus* YN-1 (Xu et al., 1991) and their deduced primary structures have been determined. However, no significant sequence homology was found between either of these NDH-2 sequences and the *Paracoccus* NADH-binding subunit (data not shown), suggesting a distant, if any, evolutionary relationship between the NDH-2 and the NADH-binding subunit of the NDH-1.

When the amino acid sequence of the *Paracoccus* NADH-binding subunit was used to search the GenBank/EMBL database, the only sequence with significant similarity to this polypeptide was the α subunit of the NAD-linked hydrogenase of *A. eutrophus* H16 (Figure 3). This NAD-linked hydrogenase, located in the cytoplasm of the bacterium, catalyzes the reaction

$$H_2 + NAD^+ \rightleftarrows NADH + H^+$$

The α subunit of this enzyme contains FMN and an iron–sulfur cluster similar to the NADH-binding subunit of the NDH-1 (Tran-Betcke et al., 1990). The high degree of sequence homology between the NADH-binding subunit of *Paracoccus* NDH-1 and the α subunit of *A. eutrophus* NAD-linked hydrogenase is also clear from the dot-matrix plot (data not shown). It should be noted that the highest degree of similarity between the *Paracoccus* NADH-binding subunit and the NAD-linked hydrogenase α subunit is observed in the putative NADH and Fe-S cluster binding sites. The sequence similarity between the *Paracoccus* NADH-binding subunit and the α subunit of the *A. eutrophus* NAD-linked hydrogenase may suggest that both subunits belong to the same family and share the same ancestor. In addition, our homology search indicated that the primary structure of the 75-kDa subunit of bovine IP is similar to that of the γ subunit of the *A. eutrophus* NAD-linked hydrogenase. These results support the above hypothesis of an evolutionary relationship between the water-soluble fraction of the NADH dehydrogenase complex and the NAD-linked hydrogenase.

The structural genes of the three subunits of the *Paracoccus* complex III have been reported to constitute a single operon (Kurowski & Ludwig, 1987). On the other hand, the structural genes of the three subunits of cytochrome oxidase in this organism are separated into two gene clusters (possibly operons; one contains the *COXI* gene and the other bears the *COXII* and *COXIII* genes) (Raitio et al., 1987). In an attempt

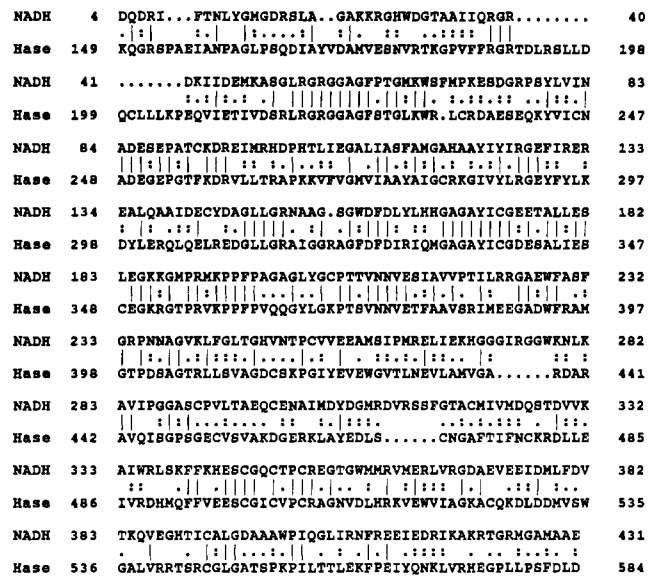

FIGURE 3: Comparison of the amino acid sequences of the NADH-binding subunit of the *Paracoccus* NADH dehydrogenase complex and the α subunit of the NAD-linked hydrogenase of *A. eutrophus* H16. The upper sequence is the deduced amino acid sequence of the *Paracoccus* NADH-binding subunit; the lower sequence is the deduced amino acid sequence of the α subunit of the NAD-linked hydrogenase. The comparison was performed by the BESTFIT program. Vertical bars are put between residues that are the same, two dots and one dot are put between residues whose comparison value is greater or equal to 0.5 and 0.1, respectively.

to determine whether the structural gene of the NADH-Q oxidoreductase may also lie in one or more clusters, DNA sequencing was performed on the regions surrounding the NADH-binding subunit gene in the *Eco*RI 5.7-kbp fragment. Sequences homologous to the 24-kDa polypeptide of bovine FP and the 49-kDa polypeptide of bovine IP (Pilkington & Walker, 1989; Fearnley et al., 1989; Chomyn & Lai, 1989) were detected, respectively, 2.2 kbp and 1.4 kbp upstream of the NADH-binding subunit gene.

```
Paracoccus   PVGRNGDCYDRYLCRMAEMRESCKIMQQ
             | |  ||||||||||  ||| |  |  |
Bovine 49k   PIGSRGDCYDRYLCRVEEMRQSTRIISQ

Paracoccus   EVECLGACTNAPMAQIGKDFYEDLTVEKLAALIDR
             |||||||| |||| ||     |||||        ||
Bovine 24k   EVECLGACVNAPMVQINDNYYEDLTPKDIEEIIDE
```

Furthermore, 1 kbp downstream of this gene, there is a sequence homologous to the 75-kDa polypeptide of bovine IP (Runswick et al., 1989).

```
Paracoccus   KKAREGVMEFLLINHPLDCPICDQGGECDL
             |||||||||||| ||||||||||||||||||
Bovine 75k   KKAREGVMEFLLANHPLDCPICDQGGECDL
```

These results indicate that the structural genes of the *Paracoccus* NADH dehydrogenase complex constitute a gene cluster. It remains to be seen whether the gene cluster of the *Paracoccus* NADH dehydrogenase complex is an operon. This is the first report of a gene cluster encoding the subunits of the energy-transducing NADH-Q oxidoreductase.

After this paper was submitted, the cDNA sequence of the NADH-binding subunit of bovine complex I has been reported by two groups (Patel et al., 1991; Pilkington et al., 1991). Comparison of the *Paracoccus* NADH-binding subunit with its bovine counterpart shows 64.3% sequence identity. As described above, such high sequence identity confirms our expectation described in the previous paper (Yagi & Dinh, 1990) on the basis of chemical and immunochemical studies. In addition, a putative NADH-binding sequence and an Fe-S cluster binding consensus sequence present in the *Paracoccus* NADH-binding subunit are conserved in the bovine enzyme. Both papers showed homology between the bovine NADH-binding subunit and residues 187–602 of the α subunit of the NAD-linked hydrogenase (Patel et al., 1991; Pilkington et al., 1991). Furthermore, residues 1–250 of the bovine 75-kDa subunit and the bovine 24-kDa subunit have been reported to be similar, respectively, to the γ subunit and residues 1–186 of the α subunit of the NAD-linked hydrogenase (Pilkington et al., 1991).

ACKNOWLEDGMENTS

We thank Dr. Alan McLachlan and Ms. Anneke Raney for the use of their PCR instrument and Drs. Youssef Hatefi, Eric F. Johnson, and Carla Hekman for critical reading of the manuscript.

REFERENCES

Chen, S., & Guillory, R. J. (1981) *J. Biol. Chem. 256*, 8318–8323.

Chomyn, A., & Lai, S. S.-A. T. (1989) *Curr. Genet. 16*, 117–125.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Fearnley, I. M., Runswick, M. J., & Walker, J. E. (1989) *EMBO J. 8*, 665–672.

Galante, Y. M., & Hatefi, Y. (1979) *Arch. Biochem. Biophys. 192*, 559–568.

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., Miyoshi, T., Takashina, S., & Unemoto, T. (1989) *Biochim. Biophys. Acta 977*, 62–69.

Hayashi, M., Hasegawa, K., Oguni, Y., & Unemoto, T. (1990) *Biochim. Biophys. Acta 1035*, 230–236.

Horváth, L. I., Drees, M., Beyer, K., Klingenberg, M., & Marsh, D. (1990) *Biochemistry 29*, 10664–10669.

Hunkapiller, M. W., Lujan, E., Ostrander, F., & Hood, L. E. (1983) *Methods Enzymol. 91*, 227–236.

Ishizuka, M., Machida, K., Shimada, S., Mogi, A., Tsuchiya, T., Ohmori, T., Souma, Y., Gonda, M., & Sone, N. (1990) *J. Biochem. 108*, 866–873.

Kurowski, B., & Ludwig, B. (1987) *J. Biol. Chem. 262*, 13805–13811.

Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035–10038.

Meinhardt, S. W., Kula, T., Yagi, T., Lillich, T., & Ohnishi, T. (1987) *J. Biol. Chem. 262*, 9147–9153.

Morgan-Hughes, J. A., Cooper, J. M., Holt, I. J., Harding, A. E., Schapira, A. H. V., & Clark, J. B. (1990) *Biochem. Soc. Trans. 18*, 523–526.

Nobrega, M. P., Nobrega, F. G., & Tzagoloff, A. (1990) *J. Biol. Chem. 265*, 14220–14226.

Osawa, S., Muto, A., Ohama, T., Andachi, Y., Tanaka, R., & Yamao, F. (1990) *Experientia 46*, 1097–1106.

Parker, W. D., Jr., Boyson, S. J., & Parks, J. K. (1989) *Ann. Neurol. 26*, 719–723.

Parker, W. D., Jr., Boyson, S. J., Luder, A. S., & Parks, J. K. (1990) *Neurology 40*, 1231–1234.

Patel, S. D., Aebersold, R., & Attardi, G. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4225–4229.

Pilkington, S. J., & Walker, J. E. (1989) *Biochemistry 28*, 3257–3264.

Pilkington, S. J., Skehel, J. M., Gennis, R. B., & Walker, J. E. (1991) *Biochemistry 30*, 2166–2175.

Ragan, C. I. (1987) *Curr. Top. Bioenerg. 15*, 1–36.

Ragan, C. I., Galante, Y. M., Hatefi, Y., & Ohnishi, T. (1982) *J. Biol. Chem. 21*, 590-594.
Raitio, M., Jalli, T., & Saraste, M. (1987) *EMBO J. 6*, 2825-2833.
Runswick, M. J., Gennis, R. B., Fearnley, I. M., & Walker, J. E. (1989) *Biochemistry 28*, 9452-9459.
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.
Schägger, H., & von Jagow, G. (1987) *Anal. Biochem. 166*, 368-379.
Scrutton, N. S., Berry, A., & Perham, R. N. (1990) *Nature 343*, 38-43.
Sharma, P. M., Reddy, G. R., Babior, B. M., & McLachlan, A. (1990) *J. Biol. Chem. 265*, 9006-9010.
Shine, J., & Dalgarno, L. (1975) *Nature 254*, 34-38.
Stouthamer, A. H. (1980) *Trends Biochem. Sci. 5*, 164-166.
Tran-Betcke, A., Warnecke, U., Bocker, C., Zaborosch, C., & Friedrich, B. (1990) *J. Bacteriol. 172*, 2920-2929.
Tuschen, G., Sackmann, U., Nehls, U., Haiker, H., Buse, G., & Weiss, H. (1990) *J. Mol. Biol. 213*, 845-857.
Tzagoloff, A., Capitanio, N., Nobrega, M. P., & Gatti, D. (1990) *EMBO J., 9*, 2759-2764.
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.
Walker, J. E., Saraste, M., & Gay, N. J. (1984) *Biochim. Biophys. Acta 768*, 164-200.
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.
Wierenga, R. K., Terpstra, P., & Hol, W. G. (1986) *J. Mol. Biol. 187*, 101-107.
Xu, X., & Yagi, T. (1991) *Biochem. Biophys. Res. Commun. 174*, 667-672.
Xu, X., Koyama, N., Cui, M., Yamagishi, A., Nosoh, Y., & Oshima, T. (1991) *J. Biochem.* (in press).
Yagi, T. (1986) *Arch. Biochem. Biophys. 250*, 302-311.
Yagi, T. (1989) *Protein, Nucleic Acid, Enzyme 34*, 351-363.
Yagi, T. (1991) *J. Bioenerg. Biomembr. 23*, 211-225.
Yagi, T., & Dinh, T. M. (1990) *Biochemistry 29*, 5515-5520.
Yagi, T., & Hatefi, Y. (1988) *J. Biol. Chem. 263*, 16150-16155.
Yagi, T., Hon-nami, K., & Ohnishi, T. (1988) *Biochemistry 27*, 2008-2013.
Yasunobu, K. T., & Tanaka, M. (1980) *Methods Enzymol. 69*, 228-238.
Young, I. G., Rogers, B. L., Campbell, H. D., Jaworowski, A., & Shaw, D. C. (1981) *Eur. J. Biochem. 116*, 165-170.

Thermodynamic Study of Internal Loops in Oligoribonucleotides: Symmetric Loops Are More Stable Than Asymmetric Loops†

Adam E. Peritz,‡ Ryszard Kierzek,§ Naoki Sugimoto,‡,∥ and Douglas H. Turner*,‡

Department of Chemistry, University of Rochester, Rochester, New York 14627, and Institute of Bioorganic Chemistry, Polish Academy of Sciences, 60-704 Poznan, Noskowskiego 12/14, Poland

Received November 8, 1990; Revised Manuscript Received April 1, 1991

ABSTRACT: Thermodynamic parameters for internal loops of unpaired adenosines in oligoribonucleotides have been measured by optical melting studies. Comparisons are made between helices containing symmetric and asymmetric loops. Asymmetric loops destabilize a helix more than symmetric loops. The differences in free energy between symmetric and asymmetric loops are roughly half the magnitude suggested from a study of parameters required to give accurate predictions of RNA secondary structure [Papanicolaou, C., Gouy, M., & Ninio, J. (1984) *Nucleic Acids Res. 12*, 31-44]. Circular dichroism spectra indicate no major structural difference between helices containing symmetric and asymmetric loops. The measured sequence dependence of internal loop stability is not consistent with approximations used in current algorithms for predicting RNA secondary structure.

Secondary structures of RNA contain motifs other than double helices. Single-stranded regions such as hairpin, bulge, internal, and bifurcated loops are prominent features. Despite their importance, little data exists on how single-strand motifs contribute to the overall stability of RNA. Incorporation of improved nearest-neighbor parameters (Freier et al., 1986a; Turner et al., 1988) into computer algorithms for structure prediction have recently increased the accuracy of RNA secondary structure predictions (Turner et al., 1987; Jaeger et al., 1989). Additional knowledge of thermodynamic parameters for single-strand motifs should further improve such predictions.

Internal loops form in double-helical RNA when the helix is interrupted by nucleotides on both strands that are not Watson–Crick paired. Symmetric or asymmetric loops can form depending on whether an equal or an unequal number of nucleotides are on opposing strands, respectively. Internal loops are thought to play important roles in biological functions

† This work was supported by National Institutes of Health Grant GM22939.
‡ University of Rochester.
§ Polish Academy of Sciences.
∥ Current address: Department of Chemistry, Konan University, Kobe 658, Japan.

0006-2960/91/0430-6428$02.50/0