Mueckler, M., Caruso, C., Baldwin, S. A., Panico, M., Blench, I., Morris, H. R., Allard, W. J., Lienhard, G. E., & Lodish, H. F. (1985) *Science 229*, 941–945.
Sadler, S. E., & Maller, J. L. (1989) *J. Biol. Chem. 264*, 856–861.
Stefanovic, D., Erikson, E., Pike, L. J., & Maller, J. L. (1986) *EMBO J. 5*, 157–160.
Stith, B. J., & Maller, J. L. (1984) *Dev. Biol. 102*, 79–82.
Suzuki, K., & Kono, T. (1981) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2542–2546.
Thorens, B., Sarkar, H. K., Kaback, H. R., & Lodish, H. F. (1988) *Cell 55*, 281–290.

Mitochondrial NADH:Ubiquinone Reductase: Complementary DNA Sequence of the Import Precursor of the Bovine 75-kDa Subunit†

Michael J. Runswick, Robert B. Gennis,‡ Ian M. Fearnley, and John E. Walker*
The Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.
Received April 20, 1989; Revised Manuscript Received July 24, 1989

ABSTRACT: The 75-kDa subunit of complex I (NADH:ubiquinone oxidoreductase) from bovine heart mitochondria is its largest subunit and is a component of the iron–sulfur (IP) fragment of the enzyme. It is encoded in nuclear DNA and is imported into the organelle. Protein sequences have been determined at the N-terminus of the intact protein and on fragments generated by partial cleavage with cyanogen bromide and with *Staphylococcus aureus* protease V8. Parts of these data have been used to design two mixtures of oligonucleotides 17 bases long, containing 192 and 256 different sequences, which have been synthesized and used as hybridization probes for identification of cognate cDNA clones. Two different but overlapping clones have been isolated, and the sequences of the cloned DNAs have been determined. Together they code for a precursor of the 75-kDa subunit of complex I. The mature protein is 704 amino acids in length, has a calculated molecular mass of 75961 daltons, and contains no segments of sequence that could be folded into hydrophobic α-helixes of sufficient length to span the inner membrane of the mitochondrion. Its precursor has an N-terminal extension of 23 amino acids to specify its import into the mitochondrion from the cytoplasm. Seventeen cysteine residues are dispersed throughout the 75-kDa subunit; some of them are close to each other in the sequence in three separate groups and, by analogy with other iron–sulfur proteins, could be involved in iron–sulfur clusters. The sequence of the 75-kDa subunit of complex I is not related closely to any known protein sequence, and the protein is not, as has been proposed, the major human biliary cirrhosis autoantigen.

Complex I (NADH:ubiquinone oxidoreductase) is the least understood of all the components of the mitochondrial respiratory chain [for a review, see Ragan (1987)]. This is due to the astonishing complexity of the enzyme. It contains one FMN plus at least seven Fe-S clusters, and analyses by SDS–PAGE[1] and two-dimensional PAGE of the bovine heart enzyme resolve 26–30 different subunits, of which seven are encoded in mitochondrial DNA (Chomyn et al., 1985, 1986; Fearnley & Walker, 1987). Complex I is located in the inner mitochondrial membrane where it catalyzes the oxidation of NADH in the matrix and the reduction of ubiquinone-10 in the bilayer. Of particular interest is the fact that the electron transfer reaction catalyzed by complex I is coupled to the generation of a protonmotive force with about two protons translocated per electron (Wikström, 1984). Mechanisms by which this might be accomplished have been discussed recently (Krishnamoorthy & Hinkle, 1988).

With the use of chaotropic agents, bovine complex I has been split up into distinct multisubunit fragments which maintain some structural and/or functional integrity. Two water-soluble fragments can be derived from the detergent-solubilized enzyme; the flavin-containing or FP fragment has three subunits (53, 24, and 10 kDa), contains the FMN plus two Fe-S clusters, and retains some NADH dehydrogenase activity. The FP subunits are probably all located on the matrix side of the inner membrane (Han et al., 1988). The iron-containing fragment, IP, contains six subunits (75, 49, 30, 18, 15, and 13 kDa) and at least three Fe-S clusters that retain their characteristic EPR spectral line shapes. Chemical labeling studies show that the 75-kDa subunit is located predominantly on the matrix side of the mitochondrial inner membrane, whereas the 49-kDa subunit is exposed on the cytoplasmic side. This would suggest that the IP fragment is transmembranous (Patel et al., 1988). However, the IP fragment appears to be water-soluble, though it is not monodisperse (Ragan, 1987). In addition to the three or more iron–sulfur clusters within the IP fragment, the 15-kDa subunit of the IP fragment is reported to bind ubiquinone (Suzuki & Ozawa, 1986). No catalytic activity has been ascribed to isolated IP. The remaining subunits in complex I, following removal of those comprising IP and FP, constitute the HP fraction, the main hydrophobic component of the enzyme. This fraction contains at least two Fe-S clusters and a subunit which binds to rotenone, a potent inhibitor of the enzyme (Earley

† The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02877.
* To whom correspondence should be addressed.
‡ Present address: School of Chemical Sciences, University of Illinois at Urbana–Champaign, Urbana, IL 61801.

[1] Abbreviations: CAPS, 3-(cyclohexylamino)propanesulfonic acid; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; PVDF, poly(vinylidene difluoride).

0006-2960/89/0428-9452$01.50/0

et al., 1987). This subunit is thought to be a mitochondrial gene product, ND1, or the 29-kDa subunit. Dicyclohexylcarbodiimide reacts with the same subunit, but at a different site from rotenone (Yagi & Hatefi, 1988).

Structural information on complex I is limited. Chemical labeling studies (Smith & Ragan, 1980; Patel & Ragan, 1988; Patel et al., 1988) and immunological methods (Han et al., 1988) have identified subunits exposed to the membrane interior or to either side of the bovine inner mitochondrial membrane. In addition, chemical cross-linking experiments have provided some information on contacts between subunits of complex I (Patel & Ragan, 1988; Patel et al., 1988). The overall dimensions and shape of the enzyme have been determined by electron microscopy of complex I from *Neurospora crassa* mitochondria (Leonard et al., 1987). In addition to those of the seven subunits coded in the mitochondrial DNA, the primary structures of only two other subunits are known, namely, those of the 24-kDa subunit from rat (Nishikimi et al., 1988) and man and cow (Pilkington & Walker, 1989) and the bovine 49-kDa subunit (Fearnley et al., 1989). Both of these subunits are coded in nuclear DNA. The former protein appears to contain cysteine residues that by homology with other proteins could be involved in Fe-S clusters (Pilkington & Walker, 1989), but the latter does not appear to contain sequences that are diagnostic of such clusters.

In this paper the primary structure of the bovine 75-kDa subunit is reported. This was deduced from the sequences of cDNAs which were cloned with oligonucleotides whose design was based upon partial protein sequence information. In addition to being the largest subunit in complex I and part of the IP fragment, this subunit has been shown previously to contain one of the Fe-S clusters, and it has been speculated that it forms part of the active site crevice where NADH is oxidized (Chen & Guillory, 1981; Ragan, 1987). The sequence contains the motif CysXXCysXXXCys, which has been found in ferredoxins in association with 4Fe-4S clusters (Stout, 1982), and two other CysXXCys sequences, which could be involved in 2Fe-2S clusters.

Materials and Methods

Isolation of Complex I. The enzyme was isolated from bovine heart mitochondria (Smith, 1967) according to the procedure of Hatefi et al. (1962) as described by Ragan et al. (1987). Its activity was approximately 50 units/mg of protein, where 1 unit is the amount of enzyme required to reduce 1 μmol of ferricyanide in 1 min at 30 °C (Hatefi et al., 1978). The enzyme was stored frozen at −20 °C. The complexity and sizes of subunits resolved by 10–20% polyacrylamide gradient gels in the presence of SDS were identical with those obtained with a similar preparation from the laboratory of Dr. C. I. Ragan (kind gift of Dr. G. Gibb, Southampton University).

Protein Determination. All protein measurements were based on a procedure that employs bicichoninic acid (the BCA protein assay reagent from Pierce Chemical Co.).

Electroelution of the 75-kDa Subunit. The subunits of complex I were resolved by SDS–PAGE using the Laemmli (1970) system and a 16% acrylamide slab gel (20 cm × 10 cm × 0.12 cm). The bands were visualized without staining and destaining by the procedure of Schägger et al. (1988). This simply requires modification of the cathode buffer by halving the concentration of SDS to 0.05% and inclusion of the dye Page Blue 83 (25 mg/L; BDH Chemicals, Poole, Dorset, U.K.), which is normally used to stain the gel after electrophoresis. This changes the mobilities of several subunits, such that the 53- and 49-kDa bands are coincident as are also the 42- and 39-kDa bands. However, the 75-kDa band is clearly visible while the gel is running. It was excised with a razor and was either stored frozen or eluted immediately with a Bio-Trap (Schleicher & Schuell, Dassel, West Germany) or with an Isco Model 1750 sample concentrator. In each case, the procedure suggested by the manufacturer was followed, except that the elution buffer consisted of 0.1 M ammonium bicarbonate and 0.1% SDS. Recovery was not quantified but appeared to be greater than 70%. About 0.5–1 mg of pure complex I could be resolved with a single 1 in. wide lane on the gel, and so this system was useful for obtaining relatively large amounts of the 75-kDa subunit (>100 μg).

Following electroelution, the 75-kDa subunit was usually sufficiently concentrated to be used directly in subsequent cleavage experiments, but in some cases, it was advantageous to concentrate and to desalt the sample by ethanol precipitation as follows. The sample (0.3–1 mL) was dried in a Speed-Vac (Savant Instruments) and then redissolved in water (50 μL), and cold (−20 °C) absolute ethanol (0.95 mL) was added. This solution was stored at −20 °C overnight, and then the precipitate was pelleted and redissolved in buffer (50 μL) containing 50 mM Tris and 0.1% SDS, pH 7.5.

The subunit obtained from this procedure was shown to be pure by SDS–PAGE using a 10%–20% gradient of polyacrylamide.

Cyanogen Bromide Fragmentation. All cleavages were carried out on a small scale with excess cyanogen bromide. Optimal conditions were determined empirically, and they varied from one preparation to another. Typically, an amount (1–10 μg) of the 75-kDa subunit was used which was sufficient to give an intense stain when visualized with Coomassie Brilliant Blue (or Page Blue 83) after resolution on a minigel (10 cm × 10 cm × 0.05 cm; 10%–20% polyacrylamide gradient). To this sample (ca. 20 μL) was added β-mercaptoethanol (2 μL), and the mixture was incubated at 37 °C for 10 min. Then the sample was dried in a Speed-Vac and dissolved in a freshly prepared solution of cyanogen bromide in 70% formic acid (100 μL). This solution was made by adding one white crystal to 0.5 mL of 70% formic acid. The protein was incubated for 2 h at room temperature and then dried in vacuo in a desiccator. The sample was dissolved in water (50 μL) and 2× Laemmli sample buffer (10 μL), applied to the minigel, and then transferred by electrophoresis to poly(vinylidene difluoride) (PVDF) membrane (Immobilon from Millipore; Matsudaira, 1987). The transfer was carried out in 10 mM CAPS buffer, pH 12.0. The transferred peptides were visualized on the PVDF membrane with 0.2% Page Blue 83 (BDH Chemicals) in 50% methanol–1% acetic acid and destained with 50% methanol.

Proteolysis by S. aureus V8 Protease. To a sample of 75-kDa protein (5–10 μg) in 50 mM Tris–0.1% SDS (pH 7.5; 40 μL) was added a portion (5 μL) of a stock solution of *S. aureus* V8 protease (0.05 mg/mL). Digestion of the protein was carried out for 45 min at room temperature and then was stopped by freezing the solution. Then the sample was dried in a Speed-Vac, and just before SDS–PAGE electrophoresis was to be carried out, it was redissolved in 2× Laemmli sample buffer (10 μL). The gel and blotting procedures employed were those described in the previous section.

Protein Sequence Analysis. Procedures employed in the sequence analysis of proteins and fragments that had been transferred to PVDF membranes have been described elsewhere (Fearnley et al., 1989).

Synthesis of Oligonucleotides. Syntheses were carried out by automated phosphoramidite chemistry with the aid of an

Applied Biosystems 380B synthesizer. The partial protein sequences IEFVDG and GNDVAA determined respectively in the N-terminal region and from a partial cyanogen bromide digest of the 75-kDa subunit of complex I were used to design two pairs of mixtures of oligonucleotides, all of them being 17 bases long. The first pair (O-1 + O-2) had the sequences ATY/AGAAGTNTTYGTNGA and ATY/AGAGGTNTTYGTNGA, and the second pair (O-3 + O-4) were GGNAATGAYGTNGCNGC and GGNAACGAYGTNGCNGC. Each of the former pair contained 96 different sequences, and each of the latter pair contained 128 different sequences. Their minimum dissociation temperatures were estimated by the procedure of Suggs et al. (1981) to be 42, 44, 50, and 52, °C, respectively. In addition, 19 unique oligonucleotides 17 bases in length were synthesized and used as primers in DNA sequencing experiments.

Screening the cDNA Library. A library of cDNAs derived from mRNA isolated from bovine heart and liver in the plasmid vector pUC8 had been made earlier (Gay & Walker, 1985). As described before (Powell et al., 1989; Fearnley et al., 1989), about 0.5×10^6 recombinants were grown up on the surface of Pall Biodyne nylon membranes placed in contact with agar plates. These served as master plates from which replicas were prepared as follows. Up to ten wetted membranes were placed sequentially in contact with each of the master plates; then, each of these copies was placed on agar, and the bacteria on their surface was grown for about 18 h. Then the copy membranes with attached bacterial colonies were lifted off the agar, and the DNA in the colonies was fixed to the membrane under conditions that have been given earlier (Gay & Walker, 1985). Only four sets of copy membranes were required for the experiments described in this paper. Samples (ca. 50 pmol) of each of the oligonucleotide mixtures (O-1, O-2, O-3, and O-4) were radiolabeled separately with polynucleotide kinase and [γ-^{32}P]ATP (specific activity about 3000 Ci/mmol) and purified by electrophoresis through a 20% polyacrylamide gel (Powell et al., 1989). Then O-1 was mixed with O-2, and O-3 with O-4, and each pair of mixtures (O-1 + O-2 and O-3 + O-4) was used separately as a hybridization probe with duplicate sets of cDNA library filters. The experimental conditions for hybridization have been described before (Powell et al., 1989), the hybridizations and subsequent washings being performed at 5 °C below the minimum dissociation temperature of the mixture, namely, at 37 °C (O-1 + O-2) and 45 °C (O-3 + O-4). Usually, 10–20 filters were probed at the same time in hybridization buffer (ca. 100 mL) containing all of the recovered radioactive probe. Filters were autoradiographed with preflashed Fuji RX film at −70 °C.

Subcloning and DNA Sequencing. The inserts in pUC plasmids were excised with a mixture of the restriction enzymes *Eco*RI and *Bam*HI. The resultant fragments were purified by electrophoresis in the presence of ethidium bromide in a 1% agarose minigel. DNA was recovered from the agarose by phenol extraction (Wieslander et al., 1979), and then *Bam*HI–*Eco*RI fragments were cloned into the M13 mp8 and mp19 vectors and *Eco*RI fragments into M13 mp18. DNA sequences were determined by the modified dideoxy procedure (Sanger et al., 1977; Biggin et al., 1983), and in addition to the universal primer LMB2 (Duckworth et al., 1981), 19 synthetic oligonucleotide primers were used to establish the sequence fully in both senses of the DNA. Sequences were compiled with the aid of the computer programs DBAUTO and DBUTIL (Staden, 1982).

Data Analysis. The protein sequence of the 75-kDa subunit of bovine NADH:ubiquinone reductase was compared with those in the PIR database with the program FASTP (Lipman & Pearson, 1985), and its hydrophobic profile was calculated with HYDROPLOT, a version of SOAP (Kyte & Doolittle, 1982).

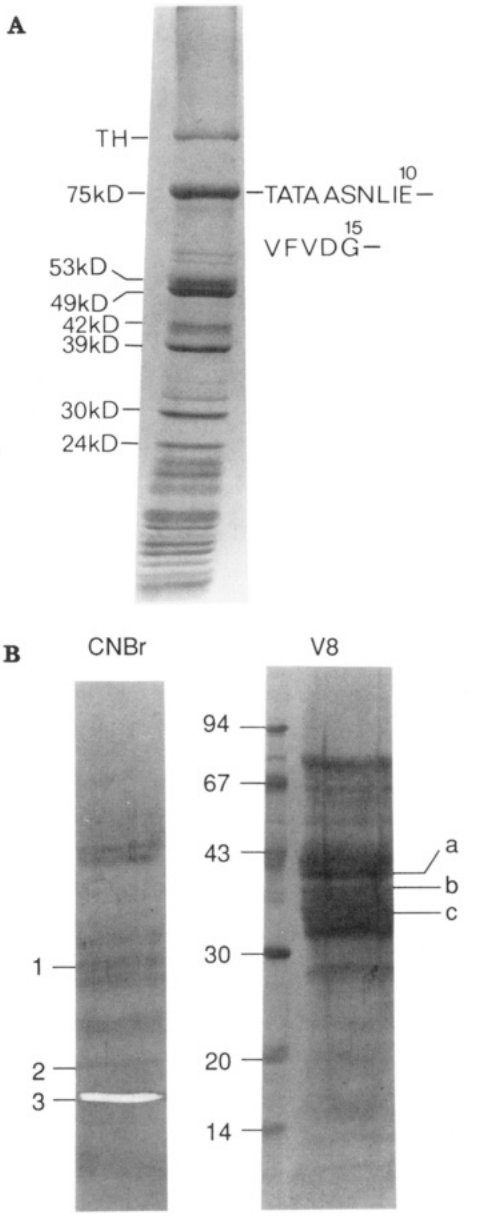


FIGURE 1: Fractionation by polyacrylamide electrophoresis in the presence of SDS of subunits of complex I from bovine heart mitochondria and of fragments derived from its 75-kDa subunit. In (A), several of the subunits of the enzyme are indicated according to their estimated molecular masses, and the sequence determined at the N-terminus of the 75-kDa subunit is given alongside that protein. The position of a contaminant, transhydrogenase (TH), is also marked. In (B), the separation is shown of fragments in partial digests of the 75-kDa subunit with cyanogen bromide and with *S. aureus* V8 protease. Bands that yielded sequence information (see Table I) are labeled 1–3 and a–c, respectively. Band 3 had been excised before the membrane had been photographed.

Results and Discussion

Amino Terminal and Internal Sequences Determined on the 75-kDa Subunit. The subunits of complex I were separated by polyacrylamide gel electrophoresis in the presence of SDS (see Figure 1A). They were then transferred to a poly(vinylidene difluoride) membrane and stained with dye, and the stained protein containing the 75-kDa subunit was excised and degraded by the Edman procedure. This generated the sequence from the N-terminus of residues 1–15, and the same

Table I: Protein Sequences Determined on the 75-kDa Subunit of Complex I and on Partial Digestion Products

digest	band[a]	apparent molecular mass (kDa)	determined sequence[b]	position in the protein	possible position of fragment in sequence	calcd molecular mass (daltons)
intact protein	75-kDa	75	TATAASNLIEVFVDG...	1–15		
cyanogen bromide	1	29	VIL/SGSSAXQ...	451–459	451–704	27967
cyanogen bromide	2	23	LFLLGADGXXI...	522–532	522–704	19967
cyanogen bromide	3	19	LQSFQGNDVAAIAGGLVDAEAL...	307–328	307–521	22988
V8 protease	a	41	TATAASNL...	1–8	1–373 or 1–347	40996 or 38304
V8 protease	b	40	VFPTAGAGTDLRSNYLLN...	348–365	348–704	38674
V8 protease	c	38	ADVVLLVGTNPRFEAPLFNAR...	374–394	374–704	35982

[a] See Figure 1. [b] See Figure 3.

sequence was obtained on the 75-kDa band isolated by the same procedure from the IP fraction of complex I. In order to obtain further protein sequence information from the 75-kDa subunit, additional material was purified from complex I by gel electrophoresis, recovered from the gel, and then partially cleaved by cyanogen bromide in one experiment and in a second experiment with *S. aureus* V8 protease (see Figure 1B). The partial cleavage products from both experiments were fractionated, and their N-terminal sequences were investigated as described for complex I. Thereby, the five additional internal sequences listed in Table I were obtained.

Cloning and DNA Sequence Analysis. Residues 9–15 from the N-terminus and the sequence GNDVAA present in one of the partial cyanogen bromide fragments were used to design two mixtures of oligonucleotides 17 bases in length containing 192 and 256 different sequences, respectively. They were radiolabeled and used as hybridization probes for the isolation of cognate cDNA clones. Some recombinants gave a strong signal with both probes, and others gave a strong signal with the second mixture only. The former were rescreened with both probes and the latter in duplicate with the second mixture only. Thus, one recombinant, pBovCI-75.1, was identified that hybridized with both probes, and four recombinants, pBovCI-75.2 to pBovCI-75.5, that hybridized with the single probe only. Isolates pBovCI-75.4 and pBovCI-75.5 were found not to encode sequences related to the 75-kD component of complex I. Restriction digestion released *Bam*HI–*Eco*RI fragments of about 1300 bp from isolates pBovCI-75.1, pBovCI-75.2, and pBovCI-75.3 and *Eco*RI fragments of about 300 bp from the first of these and of about 600 bp from the last two. These fragments were cloned into appropriate M13 vectors, and sequence analysis showed that isolates pBovCI-75.2 and pBovCI-75.3 were identical and that the 1300-bp fragments released from isolates pBovCI-75.1 and pBovCI-75.2 were not (see Figure 2). All three of these isolates encoded protein sequences that could be recognized by comparison with the directly determined partial protein sequences as being parts of the 75-kDa subunit of complex I. Therefore, their sequences were fully determined as outlined in Figure 2, and the internal *Eco*RI sites in these isolates were overlapped by sequencing appropriate *Hae*III fragments. The inserted DNAs in isolates pBovCI-75.1 and pBovCI-75.2 were found to overlap by 749 nucleotides and were completely identical over this region of overlap. Therefore, they appear to be derived from the same mRNA, and together they make up a sequence of 2422 nucleotides (see Figure 3). The 3′-terminal sequence A_{18} is separated by 13 bases from the sequence AATAAA, a typical signal for polyadenylation of mRNA (Proudfoot & Brownlee, 1976).

Sequence of the 75-kDa Subunit of Complex I. The protein encoded in the cDNA contains segments of sequence that are identical with those determined at the N-terminal of the 75-kDa subunit and at internal sites (Table I), thus identifying

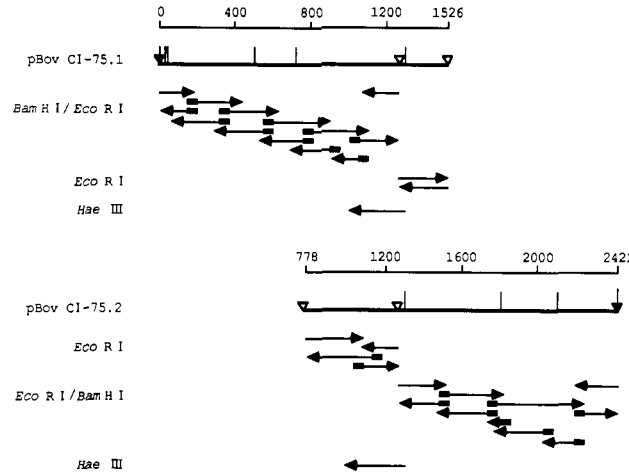
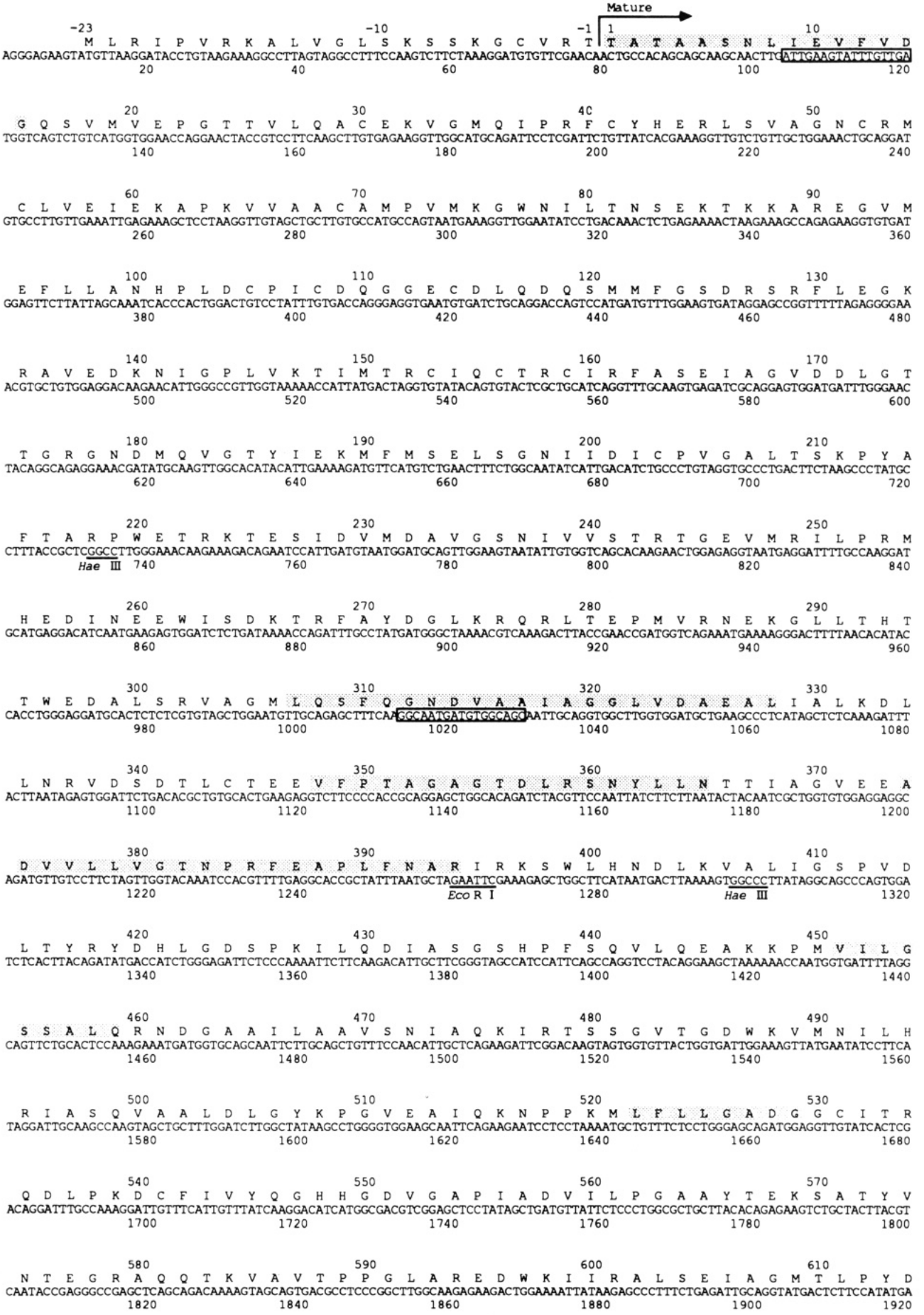


FIGURE 2: Sequence analysis of cDNA clones encoding the 75-kDa subunit of complex I from bovine heart mitochondria. The sequence was determined on two overlapping clones, pBovCI-75.1 and pBovCI-75.2, and the thick horizontal lines represent the inserted DNA in these recombinants. They are flanked in the vector in each case by a *Bam*HI (↓) and an *Eco*RI (▽) site, and in addition each recombinant has equivalent internal *Eco*RI sites. *Hae*III sites are marked by vertical lines. The horizontal arrows indicate the extents and directions of the sequences that were determined; the positions of hybridization of synthetic oligonucleotide primers that were employed in the sequencing reactions are denoted by black rectangles. The scale is in bases.

the sequence as that of the 75-kDa subunit. In further support, the molecular mass of the protein calculated from the sequence is 76961 daltons, in good agreement with the value estimated by polyacrylamide gel electrophoresis.

The sites of cleavage in the partial digests of the 75-kDa protein with cyanogen bromide and V8 protease from *S. aureus* deserve some comment. Cleavages with cyanogen bromide were found at Met–Val and Met–Leu bonds only, although not all bands that were investigated gave clearly interpretable sequence. This perhaps indicates a preference for cleavage at methionine residues that have a hydrophobic amino acid on their C-terminal side. Also, both of the identified sites of hydrolysis by V8 protease occurred after the second glutamic acid in the sequence GluGluX, consistent with a preference of the enzyme for sites containing two consecutive glutamic acids. The sizes of the fragments that yielded N-terminal sequence can be rationalized in terms of the observed cleavages only (see Table I), although in the absence of additional data it is not possible to be certain that the origins of fragments proposed in Table I are precisely correct.

The N-terminal sequence of the mature 75-kDa subunit is preceded by a protein sequence of 23 amino acids commencing with a methionine residue, presumed to be the translational initiator. This N-terminal extension is absent from the mature protein and has characteristic features of a mitochondrial import sequence: it has a net positive charge, acidic amino


Mature
-23 -10 -1 1 10
M L R I P V R K A L V G L S K S S K G C V R T T A T A A S N L I E V F V D
AGGGAGAAGTATGTTAAGGATACCTGTAAGAAAGGCCTTAGTAGGCCTTTCCAAGTCTTCTAAAGGATGTGTTCGAACAACTGCCACAGCAGCAAGCAACTTGATTGAAGTATTTGTTGA
20 40 60 80 100 120
20 30 40 50
G Q S V M V E P G T T V L Q A C E K V G M Q I P R F C Y H E R L S V A G N C R M
TGGTCAGTCTGTCATGGTGGAACCAGGAACTACCGTCCTTCAAGCTTGTGAGAAGGTTGGCATGCAGATTCCTCGATTCTGTTATCACGAAAGGTTGTCTGTTGCTGGAAACTGCAGGAT
140 160 180 200 220 240
60 70 80 90
C L V E I E K A P K V V A A C A M P V M K G W N I L T N S E K T K K A R E G V M
GTGCCTTGTTGAAATTGAGAAAGCTCCTAAGGTTGTAGCTGCTTGTGCCATGCCAGTAATGAAAGGTTGGAATATCCTGACAAACTCTGAGAAAACTAAGAAAGCCAGAGAAGGTGTGAT
260 280 300 320 340 360
100 110 120 130
E F L L A N H P L D C P I C D Q G G E C D L Q D Q S M M F G S D R S R F L E G K
GGAGTTCTTATTAGCAAATCACCCACTGGACTGTCCTATTTGTGACCAGGGAGGTGAATGTGATCTGCAGGACCAGTCCATGATGTTTGGAAGTGATAGGAGCCGGTTTTTAGAGGGGAA
380 400 420 440 460 480
140 150 160 170
R A V E D K N I G P L V K T I M T R C I Q C T R C I R F A S E I A G V D D L G T
ACGTGCTGTGGAGGACAAGAACATTGGGCCGTTGGTAAAAACCATTATGACTAGGTGTATACAGTGTACTCGCTGCATCAGGTTTGCAAGTGAGATCGCAGGAGTGGATGATTTGGGAAC
500 520 540 560 580 600
180 190 200 210
T G R G N D M Q V G T Y I E K M F M S E L S G N I I D I C P V G A L T S K P Y A
TACAGGCAGAGGAAACGATATGCAAGTTGGCACATACATTGAAAAGATGTTCATGTCTGAACTTTCTGGCAATATCATTGACATCTGCCCTGTAGGTGCCCTGACTTCTAAGCCCTATGC
620 640 660 680 700 720
220 230 240 250
F T A R P W E T R K T E S I D V M D A V G S N I V V S T R T G E V M R I L P R M
CTTTACCGCTCGGCCTTGGGAAACAAGAAAGACAGAATCCATTGATGTAATGGATGCAGTTGGAAGTAATATTGTGGTCAGCACAAGAACTGGAGAGGTAATGAGGATTTTGCCAAGGAT
Hae III 740 760 780 800 820 840
260 270 280 290
H E D I N E E W I S D K T R F A Y D G L K R Q R L T E P M V R N E K G L L T H T
GCATGAGGACATCAATGAAGAGTGGATCTCTGATAAAACCAGATTTGCCTATGATGGGCTAAAACGTCAAAGACTTACCGAACCGATGGTCAGAAATGAAAAGGGACTTTTAACACATAC
860 880 900 920 940 960
300 310 320 330
T W E D A L S R V A G M L Q S F Q G N D V A A I A G G L V D A E A L I A L K D L
CACCTGGGAGGATGCACTCTCTCGTGTAGCTGGAATGTTGCAGAGCTTTCAAGGCAATGATGTGGCAGCAATTGCAGGTGGCTTGGTGGATGCTGAAGCCCTCATAGCTCTCAAAGATTT
980 1000 1020 1040 1060 1080
340 350 360 370
L N R V D S D T L C T E E V F P T A G A G T D L R S N Y L L N T T I A G V E E A
ACTTAATAGAGTGGATTCTGACACGCTGTGCACTGAAGAGGTCTTCCCCACCGCAGGAGCTGGCACAGATCTACGTTCCAATTATCTTCTTAATACTACAATCGCTGGTGTGGAGGAGGC
1100 1120 1140 1160 1180 1200
380 390 400 410
D V V L L V G T N P R F E A P L F N A R I R K S W L H N D L K V A L I G S P V D
AGATGTTGTCCTTCTAGTTGGTACAAATCCACGTTTTGAGGCACCGCTATTTAATGCTAGAATTCGAAAGAGCTGGCTTCATAATGACTTAAAAGTGGCCCTTATAGGCAGCCCAGTGGA
1220 1240 *Eco* R I 1280 *Hae* III 1320
420 430 440 450
L T Y R Y D H L G D S P K I L Q D I A S G S H P F S Q V L Q E A K K P M V I L G
TCTCACTTACAGATATGACCATCTGGGAGATTCTCCCAAAATTCTTCAAGACATTGCTTCGGGTAGCCATCCATTCAGCCAGGTCCTACAGGAAGCTAAAAAACCAATGGTGATTTTAGG
1340 1360 1380 1400 1420 1440
460 470 480 490
S S A L Q R N D G A A I L A A V S N I A Q K I R T S S G V T G D W K V M N I L H
CAGTTCTGCACTCCAAAGAAATGATGGTGCAGCAATTCTTGCAGCTGTTTCCAACATTGCTCAGAAGATTCGGACAAGTAGTGGTGTTACTGGTGATTGGAAAGTTATGAATATCCTTCA
1460 1480 1500 1520 1540 1560
500 510 520 530
R I A S Q V A A L D L G Y K P G V E A I Q K N P P K M L F L L G A D G G C I T R
TAGGATTGCAAGCCAAGTAGCTGCTTTGGATCTTGGCTATAAGCCTGGGGTGGAAGCAATTCAGAAGAATCCTCCTAAAATGCTGTTTCTCCTGGGAGCAGATGGAGGTTGTATCACTCG
1580 1600 1620 1640 1660 1680
540 550 560 570
Q D L P K D C F I V Y Q G H H G D V G A P I A D V I L P G A A Y T E K S A T Y V
ACAGGATTTGCCAAAGGATTGTTTCATTGTTTATCAAGGACATCATGGCGACGTCGGAGCTCCTATAGCTGATGTTATTCTCCCTGGCGCTGCTTACACAGAGAAGTCTGCTACTTACGT
1700 1720 1740 1760 1780 1800
580 590 600 610
N T E G R A Q Q T K V A V T P P G L A R E D W K I I R A L S E I A G M T L P Y D
CAATACCGAGGGCCGAGCTCAGCAGACAAAAGTAGCAGTGACGCCTCCCGGCTTGGCAAGAGAAGACTGGAAAATTATAAGAGCCCTTTCTGAGATTGCAGGTATGACTCTTCCATATGA
1820 1840 1860 1880 1900 1920

```
                620                 630                 640                 650
  T L D Q V R N R L E E V S P N L V R Y D D V E G A N Y F Q Q A S E L S K L V N Q
TACTCTGGATCAAGTGAGAAACAGATTGGAAGAAGTCTCTCCTAATCTTGTTCGATATGATGATGTTGAAGGAGCTAATTATTTCCAGCAAGCAAGTGAGCTTTCCAAGCTAGTGAACCA
                1940                1960                1980                2000                2020                2040

                660                 670                 680                 690
  Q L L A D P L V P P Q L T I K D F Y M T D S I S R A S Q T M A K C V K A V T E G
ACAGCTTCTTGCTGATCCACTTGTTCCACCTCAGCTAACAATAAAAGACTTCTACATGACAGATTCAATTAGCAGAGCCTCACAGACAATGGCTAAATGTGTCAAAGCTGTTACAGAGGG
                2060                2080                2100                2120                2140                2160

                700
  A H A V E E P S I C *
TGCCCATGCAGTAGAGGAACCATCCATATGCTGAAACATCTACTATGGGCCCAGTTAATGGACAGTTATGTTGGCACGATCTCTTGAAGGTGTATCTCTTTCAAAAATAATCTTGATGTA
                2180                2200                2220                2240                2260                2280

ATATTTAAGGTTCTACCATGCTTTATTTGAAAATGCTATTGCAATTATCAGAGAACTTTGAAGTTTCAAAACAATTATGTGTTGAATGTAAATGTTAAATAGTTTAATAAACATTATATA
                2300                2320                2340                2360                2380                2400

GAGGAAAAAAAAAAAAAAAAA
                2420
```

FIGURE 3: cDNA sequence encoding the 75-kDa subunit of complex I from bovine heart mitochondria. The position of the N-terminus of the mature protein is shown, and protein sequences determined directly on the isolated protein and fragments of it are shaded. Boxes denote the sites in the DNA sequence of hybridization with mixed oligonucleotides. Also shown are the positions in the DNA sequence of restriction enzyme sites that were important in the sequence analysis. Near the 3'-end of the DNA, a potential signal for polyadenylation of mRNA is doubly underlined.

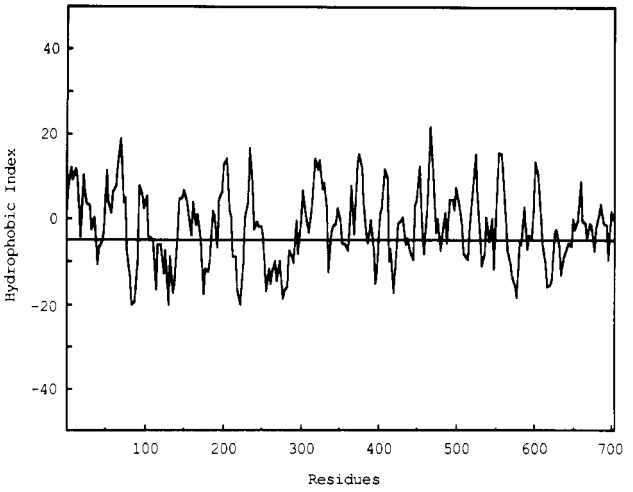


FIGURE 4: Hydrophobic profile of the 75-kDa subunit of complex I from bovine heart mitochondria. The calculation was made with HYDROPLOT using an 11 amino acid span.

acids are absent, it is not noticeably hydrophobic, and an arginine residue is found close to the junction with the mature protein (von Heijne, 1986; Roise et al., 1986). The secondary structure of this sequence is not strongly predicted to be α-helical by the procedure of Garnier et al. (1978), but if it had such a secondary structure, the helix would be amphipathic.

The hydrophobic profile of the 75-kDa subunit of complex I (Figure 4) indicates that it contains no segments of sequence that could form hydrophobic α-helixes of sufficient length to traverse the inner membrane of the mitochondrion, and the profile is characteristic of a globular protein. A similar conclusion has been reached with the 49-kDa subunit of the IP fragment of complex I (Fearnley et al., 1989). Chemical labeling experiments indicate that the 75-kDa subunit is exposed largely on the matrix side of the membrane but that it may also be exposed on the cytoplasmic side (Ragan, 1987). However, it is not labeled by hydrophobic probes, and so any membrane-spanning segment would have to be shielded from reaction with such reagents by other proteins. Both the lack of hydrophobic segments in the protein sequence and the failure of the protein to react with hydrophobic probes are accommodated by models of complex I in which the 75-kDa subunit crosses the membrane, but is shielded from the lipid bilayer by an annulus of more hydrophobic proteins (Ragan, 1987; Patel et al., 1988). However, the evidence for reaction of the 75-kDa protein with hydrophilic reagents on the cytoplasmic face of the inner membrane is not convincing (Ragan, 1987), and at present it is equally probable that the 75-kDa subunit is a peripheral membrane protein associated with the surface of the membrane on the matrix side.

Table II: Amino Acid Composition of the 75-kDa Subunit of Complex I from Bovine Heart Mitochondria

amino acid	no. of residues	amino acid	no. of residues
aspartic acid	45	methionine	23
asparagine	27	isoleucine	43
threonine	45	leucine	63
serine	40	tyrosine	15
glutamic acid	43	phenylalanine	18
glutamine	29	histidine	11
proline	33	lysine	35
glycine	51	arginine	38
alanine	65	tryptophan	7
cysteine	17		
valine	56	total	704

Iron–Sulfur Clusters in the 75-kDa Subunit. The IP fragment of complex I contains at least three iron–sulfur clusters, and a binuclear cluster has been proposed to be associated with the 75-kDa subunit (Ragan, 1987). Comparisons of its sequence with those of other iron–sulfur proteins failed to detect any extensive homology. However, its amino acid composition calculated from the sequence (see Table II) shows that it contains 17 cysteine residues, and some of these are in clusters that resemble those that are found in ferredoxins. Notably, the sequence CysXXCysXXCys (amino acids 153–159) has been found in 4Fe-4S ferredoxins (Stout, 1982), although usually but not invariably [see Yasonobu and Tanaka (1980)] in association with a fourth cysteine located four amino acids in a C-terminal direction. The sequence motif CysX-XXXXCysXXCys found in many 2Fe-2S ferredoxins (Yasonobu & Tanaka, 1980; Cammack, 1983) is not present in the 75-kDa subunit, but the sequence CysXXCys, in addition to the two examples between residues 153–159, also occurs at residues 52–55 and 105–108. This sequence is found twice in rubredoxins (Bruschi et al., 1976) where it coordinates a single iron atom and also is associated with the 2Fe-2S cluster of the Rieske iron–sulfur protein in *N. crassa* (Harnisch et al., 1985). Further experiments are required to establish the precise nature of the Fe-S centers in the 75-kDa subunit (and in other subunits of complex I). Sequences of homologous

proteins in other species might help to narrow down the cysteine residues that are involved in these clusters by demonstrating which are conserved and which are not.

Lack of Sequence Relationship to Known Proteins. The sequence of the 75-kDa protein is not evidently related significantly to any protein sequence in the PIR database. Of particular interest in this respect is the lack of relationship to any of the proteins coded in the chloroplast DNAs of *Nicotiana tabacum* and *Marchantia polymorpha* since it has been shown that they do encode homologues of another nuclear-encoded subunit of the IP fragment of complex I, namely, the 49-kDa component (Fearnley et al., 1989), as well as homologues of the mitochondrially encoded components of complex I (Shinozaki et al., 1986; Ohyama et al., 1986). So it appears that chloroplasts contain an enzyme that is similar to mitochondrial complex I, with some subunits encoded in the nucleus and others in the plastid DNA. Homologues of the 75-kDa subunit and also of the 24-kDa subunit of complex I, if present in the chloroplast enzyme, would be in the former category.

Also of interest is the lack of relationship of the sequence of the 75-kDa subunit of mitochondrial complex I with that of a major autoantigen closely associated with primary biliary cirrhosis (Gershwin et al., 1987). It had been proposed on the basis of immunological studies with antibodies against bovine complex I that the antigen was the 75-kDa subunit of complex I (Frostell et al., 1988), but subsequently, it was shown to be a different mitochondrial protein, the E2 component of pyruvate dehydrogenase (Yeaman et al., 1988; Coppell et al., 1988; Fussey et al., 1988).

ACKNOWLEDGMENTS

We thank Dr. J. M. Cooper for a gift of a sample of the IP fragment of complex I, J. M. Skehel for his assistance with the operation of the protein sequencer, and R. Jakes for help with protein cleavage experiments. R.B.G. was a Fulbright Scholar during this work and was supported, in part, as a Fellow of the John Simon Guggenheim Foundation.

REFERENCES

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

Bruschi, M., Bonicel, J., Bovier-Lapierre, G., & Couchoud, P. (1976) *Biochim. Biophys. Acta 434*, 4–17.

Cammack, R. (1983) *Chem. Scr. 21*, 87–95.

Chen, S., & Guillory, R. J. (1984) *J. Biol. Chem. 259*, 5124–5131.

Chomyn, A., Mariottini, P., Cleeter, M. W. J., Ragan, C. I., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R. F., & Attardi, G. (1985) *Nature (London) 314*, 592–597.

Chomyn, A., Cleeter, M. W. J., Ragan, C. I., Riley, M., Doolittle, R. F., & Attardi, G. (1986) *Science (Washington, D.C.) 234*, 614–618.

Coppel, R. L., McNeilage, L. J., Surh, C. D., van de Water, J., Spithill, T. W., Whittingham, S., & Gershwin, M. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7317–7321.

Duckworth, M. L., Gait, M. J., Goelet, P., Hong, G. F., Singh, M., & Titmas, R. (1981) *Nucleic Acids Res. 9*, 1691–1706.

Earley, F. G. P., Patel, S. D., Ragan, C. I., & Attardi, G. (1987) *FEBS Lett. 219*, 108–113.

Fearnley, I. M., & Walker, J. E. (1987) *Biochemistry 26*, 8247–8251.

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

Frostell, Å., Mendel-Hartvig, I., Nelson, B. D., Tötterman, T. H., Björkland, A., & Ragan, C. I. (1988) *Scand. J. Immunol. 28*, 157–165.

Fussey, S. P. M., Guest, J. R., James, O. F. W., Bassendine, M. F., & Yeaman, S. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8654–8658.

Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97–120.

Gay, N. J., & Walker, J. E. (1985) *Biochem. J. 225*, 707–712.

Gershwin, M. E., Mackay, I. R., Sturgess, A., & Coppel, R. L. (1987) *J. Immunol. 138*, 3525–3531.

Han, A.-L., Yagi, T., & Hatefi, Y. (1988) *Arch. Biochem. Biophys. 267*, 490–496.

Harnisch, U., Weiss, H., & Sebald, W. (1985) *Eur. J. Biochem. 149*, 95–99.

Hatefi, Y., Haavik, A. G., & Griffiths, D. E. (1962) *J. Biol. Chem. 237*, 1676–1680.

Krishnamoorthy, G., & Hinkle, P. C. (1988) *J. Biol. Chem. 263*, 17566–17575.

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

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

Leonard, K., Haiker, H., & Weiss, H. (1987) *J. Mol. Biol. 194*, 277–286.

Lipman, D. J., & Pearson, W. R. (1985) *Science (Washington, D.C.) 227*, 1435–1441.

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

Nishikimi, M., Hosokawa, Y., Toda, H., Suzuki, H., & Ozawa, T. (1988) *Biochem. Biophys. Res. Commun. 157*, 914–920.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., & Ozeki, H. (1986) *Nature (London) 322*, 571–574.

Patel, S. D., & Ragan, C. I. (1988) *Biochem. J. 256*, 521–528.

Patel, S. D., Cleeter, M.W. J., & Ragan, C. I. (1988) *Biochem. J. 256*, 529–535.

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

Powell, S. J., Medd, S. M., Runswick, M. J., & Walker, J. E. (1989) *Biochemistry 28*, 866–873.

Proudfoot, N. J., & Brownlee, G. G. (1976) *Nature (London) 263*, 211–214.

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

Ragan, C. I., Wilson, M. T., Darley-Usmar, V. M., & Lowe, P. N. (1987) in *Mitochondria: A Practical Approach* (Darley-Usmar, V. M., Rickwood, D., & Wilson, M. T., Eds.) pp 79–112, IRL Press, Oxford and Washington.

Roise, D., Horvath, S. J., Tomich, J. M., Richards, J. H., & Schatz, G. (1986) *EMBO J. 5*, 1327–1334.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Schägger, H., Aquila, H., & von Jagow, G. (1988) *Eur. Bioenerg. Conf. Short Rep. 5*, 259.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., & Sugiura, M. (1986) *EMBO J. 5*, 2043–2049.

Smith, A. L. (1967) *Methods Enzymol. 10*, 81–86.

Smith, S., & Ragan, C. I. (1980) *Biochem. J. 185*, 315–326.

Staden, R. (1982) *Nucleic Acids Res. 10*, 4731–4751.

Stout, C. D. (1982) in *Iron Sulfur Proteins* (Spiro, T. G., Ed.) pp 97–146, Wiley, New York.

Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Johnson, M. J., Itakura, K., & Wallace, R. B. (1981) in *Developmental Biology Using Purified Genes* (Brown, D.

D., Ed.) pp 683-693, Academic Press, New York and London.
Suzuki, H., & Ozawa, T. (1986) *Biochem. Biophys. Res. Commun. 138*, 1237-1242.
von Heijne, G. (1986) *EMBO J. 5*, 1335-1342.
Wieslander, L. (1979) *Anal. Biochem. 98*, 305-309.
Wikström, M. (1984) *FEBS Lett. 169*, 300-304.
Yagi, T., & Hatefi, Y. (1988) *J. Biol. Chem. 263*, 16150-16155.
Yasunobu, K., & Tanaka, M. (1980) *Methods Enzymol. 69*, 228-238.
Yeaman, S. J., Fussey, S. P. M., Danner, D. J., James, O. F. W., Mutimer, D. J., & Bassendine, M. F. (1988) *Lancet 1*, 1067-1070.

A Calcium-Specific Site Influences the Structure and Activity of the Manganese Cluster Responsible for Photosynthetic Water Oxidation†

M. Sivaraja, J. Tso, and G. C. Dismukes*
Department of Chemistry, Princeton University, Princeton, New Jersey 08544
Received August 28, 1989; Revised Manuscript Received September 26, 1989

ABSTRACT: EPR studies have revealed that removal of calcium using citric acid from the site in spinach photosystem II which is coupled to the photosynthetic O_2-evolving process produces a structural change in the manganese cluster responsible for water oxidation. If done in the dark, this yields a modified S_1' oxidation state which can be photooxidized above 250 K to form a structurally altered S_2' state, as seen by formation of a "modified" multiline EPR signal. Compared to the "normal" S_2 state, this new S_2'-state EPR signal has more lines (at least 25) and 25% narrower ^{55}Mn hyperfine splittings, indicative of disruption of the ligands to manganese. The calcium-depleted S_2' oxidation state is greatly stabilized compared to the native S_2 oxidation state, as seen by a large increase in the lifetime of the S_2' EPR signal. Calcium reconstitution results in the reduction of the oxidized tyrosine residue $^{161}Y_D^+$ (E_m ~0.7-0.8 V, NHE) within the reaction center D_1 protein in both the S_1' and S_2' states, as monitored by its EPR signal intensity. We attribute this to reduction by Mn. Thus a possible structural role which calcium plays is to bring Y_D^+ into redox equilibrium with the Mn cluster. Photooxidation of S_2' above 250 K produces a higher S state (S_3 or S_4) having a new EPR signal at $g = 2.004 \pm 0.003$ and a symmetric line width of 163 ± 3 G, suggestive of oxidation of an organic donor, possibly an amino acid, in magnetic contact with the Mn cluster. This EPR signal forms in a stoichiometry of 1-2 relative to Y_D^+. This state is photoaccumulated, does not evolve O_2, and decays in the dark to the stable S_2' state. The enhanced stability and apparent lowered redox potential of the S states can be explained if calcium depletion exposes the Mn cluster to an increased solvent activity, resulting in the binding and hydrolysis of additional water ligands (hydroxo and oxo). The possibility that this causes disproportionation of Mn^{III} to Mn^{II} + Mn^{IV} is considered on the basis of analogy to the hydrolysis-induced disproportionation observed for synthetic dimanganese complexes. A "gatekeeper" role for calcium in limiting access of substrate water to the catalytic Mn cluster is indicated.

Calcium is required for normal functioning of the photosynthetic water-oxidizing complex. Its precise role has been difficult to elucidate owing to it poor spectroscopic detectability, unlike the active site of this complex which is comprised of four closely arranged Mn ions. This complex can exist in five oxidation states, so-called S states, produced by photooxidation of photosystem II (PSII). The highest oxidation state, S_4, oxidizes water to O_2. The effects of calcium depletion by salt washing with or without chelators have been controversial. Calcium depletion has been observed to slow the kinetics of reduction of the photooxidized tyrosine radical Y_Z^+ by the Mn cluster in PSII (Ghanotakis et al., 1984; Dekker et al., 1984; Cole & Sauer, 1987), to eliminate the characteristic multiline EPR signal associated with the S_2 state (de Paula et al., 1986), and to block photooxidation of manganese beyond the S_3 without the ability to form the S_2 multiline EPR signal (Boussac et al., 1985; Boussac & Rutherford, 1988). On the other hand, recent studies employing low-pH incubation with citric acid to quantitatively remove one of two calciums per PSII have shown from thermoluminescence that a block in photooxidation beyond the S_2 state is produced (Ono & Inoue, 1989a).

At the 8th International Congress on Photosynthesis held in Stockholm, Aug 6-11, three groups presented results demonstrating that a "modified" multiline EPR signal can now be observed in PSII membranes treated to deplete calcium (Ono & Inoue, 1989b; Boussac & Rutherford, 1989; Baumgarten et al., 1989). Results from all three groups, including ours, suggested that the new EPR signal was stable over hours in the dark and could be attributed to a modified S_2' oxidation state, formed by dark adaptation following room temperature illumination. It was also noted to form partially in the dark, possibly by oxidation from an unknown species (Boussac & Rutherford). Ono and Inoue observed that its formation by illumination required a higher temperature than the normal S_2 state and that calcium-depleted PSII was unable to undergo further stable charge separation, owing to a block in the $S_2' \rightarrow S_3'$ reaction. They also found the modified S_2' state to be thermodynamically more stable than the normal S_2 state, as seen by an increased temperature for thermoluminescence

† Supported by NIH Grant GM 39932.

0006-2960/89/0428-9459$01.50/0