Analysis of the Subunit Composition of Complex I from Bovine Heart Mitochondria*

Joe Carroll, Ian M. Fearnley, Richard J. Shannon, Judy Hirst, and John E. Walker‡

Complex I purified from bovine heart mitochondria is a multisubunit membrane-bound assembly. In the past, seven of its subunits were shown to be products of the mitochondrial genome, and 35 nuclear encoded subunits were identified. The complex is L-shaped with one arm in the plane of the membrane and the other lying orthogonal to it in the mitochondrial matrix. With mildly chaotropic detergents, the intact complex has been resolved into various subcomplexes. Subcomplex Iα represents the extrinsic arm, subcomplex Iκ represents subcomplex Iα plus part of the membrane arm, and subcomplex Iβ is another substantial part of the membrane arm. The intact complex and these three subcomplexes have been subjected to extensive reanalysis. Their subunits have been separated by three independent methods (one-dimensional SDS-PAGE, two-dimensional isoelectric focusing/SDS-PAGE, and reverse phase high pressure liquid chromatography (HPLC)) and analyzed by tryptic peptide mass fingerprinting and tandem mass spectrometry. The masses of many of the intact subunits have also been measured by electrospray ionization mass spectrometry and have provided valuable information about post-translational modifications. The presence of the known 35 nuclear encoded subunits in complex I has been confirmed, and four additional nuclear encoded subunits have been detected. Subunits B16.6, B14.7, and ESSS were discovered in the SDS-PAGE analysis of subcomplex Iα, in the two-dimensional gel analysis of the intact complex, and in the HPLC analysis of subcomplex Iβ, respectively. Despite many attempts, no sequence information has been obtained yet on a fourth new subunit (mass 10,566 ± 2 Da) also detected in the HPLC analysis of subcomplex Iβ. It is unlikely that any more subunits of the bovine complex remain undiscovered. Therefore, the intact enzyme is a complex of 46 subunits, and, assuming there is one copy of each subunit in the complex, its mass is 980 kDa. Molecular & Cellular Proteomics 2:117–126, 2003.

NADH:ubiquinone oxidoreductase (complex I) (1, 2) catalyzes the first step of the electron transport chain in mitochondria (3, 4). It transfers electrons from NADH to a non-covalently bound FMN and then via a series of iron-sulfur clusters to the terminal acceptor, ubiquinone. The transfer of two electrons is coupled to the translocation of four protons across the inner membrane (5). The enzyme from bovine heart mitochondria is the best characterized, and it serves as a valuable model for the human enzyme where, because of its involvement in human disease, there is growing interest (6, 7). It is an L-shaped assembly of more than 40 different proteins. Seven hydrophobic components are products of the mitochondrial genome (8, 9), and the remainder are nuclear gene products that are imported into the organelle. One arm of the L-shaped complex is in the plane of the membrane, and the other protrudes into the mitochondrial matrix (10, 11). The intact complex has been resolved with chaotropic agents into a number of subcomplexes, and one of them, subcomplex Iκ, represents the extrinsic globular domain of the intact complex (12, 13). Subcomplex Iκ contains both subcomplex Iα and part of the membrane arm, and subcomplex Iβ is another independent portion of the membrane arm (14, 15).

A long term objective is to determine the atomic structure of bovine complex I, and the definition of the subunit compositions of the intact complex and its subcomplexes is an essential step in this process. In the early 1990s, 35 nuclear encoded subunits were characterized. Since then the purity of the complex has improved, and more sensitive methods for protein analysis have been developed. Therefore, as described below, the subunit compositions of the complex and its subcomplexes have been reanalyzed comprehensively by a combination of fractionation of subunits on 1D and 2D gels and by HPLC coupled with modern methods of protein analysis by mass spectrometry. The presence in the complex of the 35 previously described subunits has been confirmed, and four hitherto unknown subunits have been detected. The sequences of three of them are described elsewhere (12, 14), and the fourth subunit has not been sequenced yet. It is unlikely that any more subunits of the complex remain to be discovered. The total number of different subunits in the bovine heart complex is 46.

1 The abbreviations used are: 1D, one-dimensional; 2D, two-dimensional; ASB-14, amidosulfobetaine-14; IPG, immobilized pH gradient; ESI, electrospray ionization; MS, mass spectrometry; MALDI, matrix-assisted laser desorption ionization; TOF, time of flight; HPLC, high pressure liquid chromatography; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; CHAPS, 3-[3-cholamidopropyl]dimethylammonio]-1-propanesulfonic acid.
EXPERIMENTAL PROCEDURES

Purification of Complex I and Its Subcomplexes from Bovine Heart Mitochondria—The isolation of mitochondria from bovine hearts and the preparation of mitochondrial membranes have been described before (16). Complex I was solubilized with n-dodecyl-β-d-maltoside (Anatrace, Maumee, OH) and purified on a Q-Sepharose HP column (Amersham Biosciences) followed by ammonium sulfate precipitation and gel filtration as before (17) except that Superose 6 HR was replaced by Sephacryl S-300 HR (Amersham Biosciences). The S-300 column provided an effective way of removing residual cytochrome-c oxidase. All purification steps were carried out at 4 °C.

Subcomplexes Iα and Iβ were prepared from complex I by chromatography on Q-Sepharose in 0.1% N,N-lauryldimethylamine oxide with a salt gradient (14). Subcomplexes Iα and Iβ eluted at 260 and 325 mM NaCl, respectively. The breakthrough fractions contained material referred to previously as subcomplex Iγ (see “Results”). The 42-kDa subunit, contaminated with lower levels of other subunits, eluted at 125 mM NaCl. The purification of subcomplex Iα has been described elsewhere (12).

Fractionation of Protein Complexes in 1D and 2D Gels—The subunits of protein complexes were fractionated by SDS-PAGE in 12–22% gels (12) and on 2D gels (isoelectric focusing followed by SDS-PAGE). For the latter purpose, samples of complex I and its subcomplexes were prepared either by dialysis against a buffer, pH 7.4, containing 20 mM Tris-HCl and 0.05% n-dodecyl-β-maltoside followed by concentration to 10–20 mg/ml using Ultraflo-0.5 filter units (Millipore, Bedford, MA) or by precipitation with chloroform/methanol containing 5 mM CaCl2 or Asp-N protease (Roche Molecular Bio sciences) in 20 mM Tris-HCl buffer, pH 8.0. Proteins were also precipitated with a similar solution (but without ASB-14) to a final concentration of ASB-14 of 0.15%. Strips of IPG (7 cm, pH 3–10 or 6–11) were rehydrated in these solutions for 12 h at 20 °C with a potential of 20 V. The 2D separations by isoelectric focusing and then SDS-PAGE in a 13% polyacrylamide gel in Tricine buffer (18) were carried out as described previously (14).

Chromatographic Purification of Subunits by Reverse Phase HPLC—Subunits of subcomplexes Iα, Iβ, and Iγ were isolated by reverse phase HPLC on a column of Aquapore RP-300 (PerkinElmer Life Sciences) in 0.1% trifluoroacetic acid with a gradient of acetonitrile (12). Each peak was collected separately.

Mass Spectrometric Analysis of Subunits—Every band or spot on 1D or 2D gels was analyzed by peptide mass fingerprinting of tryptic peptides. In every subunit, at least one tryptic peptide was sequenced and identified, and the new subunit, B16.6, was discovered from CNBr peptides, and the position of ND6 was determined with a polyclonal antibody (data not shown). The hydrophobic subunit B14.7, which was discovered in the 2D analysis of complex I (see below), also stained weakly with Coomassie Blue dye, and it is possible that its staining is suppressed by co-migration with subunit ND3 (14). However, all of these hydrophobic subunits were detected clearly by silver staining (data not shown). The subunits marked on the...
Subcomplexes I and B of the cytochrome VB, VIA, and VIB of cytochrome-impurities were detected sporadically. They include subunits from preparations of the various complexes. A number of minor gels were found consistently at abundant levels in independent analyses of subcomplexes. By analysis of every spot in every train in every gel, this pretreatment improved the resolution of subunit B16.6 (14), diminished the resolution of subunit ASHI, and led to the complete loss of the 10-kDa subunit from subcomplexes. In the gels in Fig. 3, a number of subunits are present as multiply resolved “trains” of spots, which often indicate partial post-translational modifications. Each spot in every train was analyzed by peptide mass fingerprinting, and for each train the MALDI spectra from component spots were very similar. Therefore, the components in each train derive from the same protein, and since the isolated subunits gave unique protein masses by ESI-MS analysis, the trains are artifacts probably arising from partial carbamylation of lysine residues by cyanate derived by disproportionation of urea and/or partial deamidation of asparagines (22, 23).

Subunits Identified by Reverse Phase HPLC Fractionation and Mass Spectrometry

The subunits of subcomplexes Iα, Iβ, and Iλ, but not of complex I, were resolved by reverse-phase HPLC (see Fig. 4), and the subunits in each peak were identified by ESI-MS (see below) or by SDS-PAGE. The new subunit, ESSS, was discovered in the analysis of subcomplex Iβ (14). In the chromatographic separation of subunits of subcomplexes, the hydrophobic subunits ND1–ND6, ND4L, and B14.7 were not recovered from the column, and therefore their masses were not measured by ESI-MS (see below). In numerous independent analyses of subcomplex Iβ, a protein with a mass of 10,566 (±2) Da has been observed to coelute with subunit SGDH (see Fig. 4C). In subsequent experiments with a modified gradient, the two subunits were resolved partially. The mixture of the two subunits was digested in separate digests with trypsin, endoproteinase Asp-N, and CNBr and with trypsin and CNBr sequentially in a double digest. Only peptides from subunit SGDH were observed. On N-terminal analysis of the mixture, only the sequence of SGDH was observed, but by trypsinic mass mapping and of many by tandem mass spectrometry, 34 of the 45 sequenced subunits of the intact complex were identified, and the new subunit, B14.7, was discovered during the analysis of the intact complex (14). The positions of the subunits on the gels are consistent with calculated isopotential points and molecular masses. The seven hydrophobic subunits ND1–ND6 and ND4L were not detected nor were subunits AGGG, ESSS, and SDAP, all components of the hydrophobic subcomplex Iβ. Their absence illustrates the well known unsuitability of 2D gels for analysis of membrane proteins. Because they are insoluble or at best sparingly soluble in the solutions used for the rehydration of the IPG strips they fail to enter the isoelectric focusing gel. Also subunit MLRQ was not detected in the gels shown in Fig. 3, but in other gels (not shown) this subunit has been identified as a rather indistinct series of spots (14).

To a minor extent, the 2D gel patterns were influenced by pretreatment of samples with chloroform/methanol (see “Experimental Procedures”). On the pH 6–11 gel, this pretreatment improved the resolution of subunit B16.6 (14), diminished the resolution of subunit ASHI, and led to the complete loss of the 10-kDa subunit from subcomplexes. In the gels in Fig. 3, a number of subunits are present as multiply resolved “trains” of spots, which often indicate partial post-translational modifications.
the mass of the unknown subunit is not compatible with it being an N-terminal fragment of SGDH, and therefore its N terminus is modified. The blocking group was not removed by treatment with methanolic HCl. Therefore, the blocking group is not formyl. The absence of an N-formyl group and the molecular mass of the unknown subunit show that the unknown protein is neither an intact ND subunit (mitochondrial DNA gene product) nor a fragment of any of them. For a number of reasons, it cannot be a fragment of other nuclear encoded blocked subunits (B-subunits) of complex I. For example, all of the B-subunits yielded peptides in various digests, and the mass of the unknown subunit is not compatible with it being a fragment of any of them. Therefore, this recalcitrant 46th subunit requires further exploration.

The accurate measurement of intact protein masses together with knowledge of the N-terminal sequences of subunits helped to detect and identify many post-translational modifications. All of the proteins were identified by peptide mass fingerprinting of tryptic peptides. Many of the modifications have been described before, and most of them have been verified subsequently by tandem MS experiments. These data will be presented elsewhere. A summary of many of these post-translational modifications is given in Table I and in the following sections.

Mitochondrial Import Sequences—By comparison of the N-terminal sequences of mature subunits assembled into bovine complex I with protein sequences deduced from bovine cDNA sequences, it is evident that at least 16 subunits have import sequences (24) that are removed during entry into the mitochondrial (see Table I). Uncertainty remains about the 49-kDa subunit and subunit SDAP (see Footnotes f and h in Table I). Another 20 subunits, all of them with molecular masses below 22 kDa, appear to lack N-terminal extensions that could act as import sequences, and therefore the information for import into the organelle resides in the mature proteins (see Table I). The translational initiator methionine residues of these subunits have a variety of fates. In subunit

![Fig. 3. Resolution of the subunits of bovine complex I and its subcomplexes on 2D gels. A, B, C, and D show complex I and subcomplexes Iα, Iβ, and Iβ, respectively. The first dimensions were conducted with either IPG 3–10 (left-hand panels) or IPG 6–11 (right-hand panels). Before gel analysis, samples were dialyzed and concentrated. For other details, see “Experimental Procedures.” Except for subunit KFYI, where the identification was made solely by peptide mass fingerprinting, proteins were identified by peptide mass fingerprinting and tandem MS analyses. Each analysis has been conducted several times with independent samples of the complexes. The subunits are numbered as follows: 1, 75 kDa; 2, 51 kDa; 3, 49 kDa; 4, 30 kDa; 5, 24 kDa; 6, PSST; 7, TYKY; 8, 42 kDa; 9, 39 kDa; 10, 18 kDa; 11, 15 kDa; 12, 13 kDa; 13, 10 kDa; 15, ASHL; 17, KFYI; 19, MNLL; 20, MWFE; 21, PDSW; 22, PGIV; 24, SGDH; 25, B22; 26, B18; 27, B17.2; 28, B17; 29, B16; 30, B15; 31, B14.7; 32, B14.5a; 33, B14.5b; 34, B14; 35, B13; 36, B12; 37, B9; 38, B8. The numbers correspond to the descending order of subunits in Table I. Subunits 14 (AGGG), 16 (ESSS), 18 (MLRQ), and 23 (SDAP) were not detected.](image)
The Acyl Carrier Protein, Subunit SDAP—The sequence of subunit SDAP identified it as an acyl carrier protein. The mature protein mass calculated from the cDNA sequence is 10,109.6 Da, and its mass was measured by ESI-MS to be 10,751.6 Da (25). With subunit isolated from recent preparations of bovine complex I, a mass of 10,674.2 Da has been determined. The difference between the two measured values (77.4 Da) corresponds to the formation in earlier enzyme preparations of a mixed disulfide between the subunit (probably via cysteine 72) and β-mercaptoethanol (calculated value 76.1 Da), a component of buffers used formerly in the isolation of bovine complex I. In recent preparations, the β-mercaptoethanol has been replaced by dithiothreitol, and thereby formation of mixed disulfide has been avoided. In the earlier experiments, incubation of the protein under alkaline and reducing conditions decreased the mass from 10,751.6 to 10,449.4 Da, consistent with the hydrolysis of a thioester linkage and loss of an acyl group with a mass of 302.2 Da (25). The mass of the deacyl protein (10,449.4 Da) corresponds to modification of the subunit with the pantetheine-4'-phosphate moiety (10,449.8 Da) as anticipated for an acyl carrier protein. In view of the recent measurements, it is now clear that the loss of 302.2 Da corresponded to the removal of both a β-mercaptoethyl and an acyl group. Thus, the revised mass of the acyl group is 224.4 Da, leading to a free acid with a mass of 241.4 Da. This value is close to the calculated mass of 3-hydroxytetradecanoic acid (244.4 Da), the proposed identity of the acyl group on the acyl carrier protein found in complex I from *Neurospora crassa* (26, 27). Biotination of the pantetheine-4'-phosphate group would also produce a mass consistent with the observed value of 10,676.1 Da.

Polymorphisms—For both the 51-kDa and the 42-kDa subunits, two sequences have been determined, each differing in a single amino acid. In the 51-kDa subunit, residue 393 has been determined as cysteine (28) and tryptophan (29). Both ESI-MS and tandem MS experiments (see Supplemental Data Fig. S2-2) show that residue 393 is a tryptophan, but the Cambridge cDNA sequence, where the existing data have been reexamined and the existing clone has been reisolated from the same library and resequenced, shows cysteine at this position. By sequencing the cDNA of the 42-kDa subunit, residue 255 has been found to be asparagine (30) and alternatively lysine.2 The ESI-MS experiments support lysine, but the cDNA sequences provide clear evidence for both amino acids. In both cases, the available evidence is consistent with the presence of polymorphisms in the bovine population.

Residual Problems—At the end of this extensive recharacterization of the subunit composition of bovine complex I, a number of apparent anomalies and discrepancies remain between calculated and determined masses of nuclear encoded subunits. They require further experimental exploration.

The measured mass of subunit B12 exceeds the calculated...
The Subunit Composition of Bovine Complex I

### TABLE I

| Subunit* | Mass measurements by ESI-MS and post-translational modifications of nuclear encoded subunits of bovine complex I |
|----------|----------------------------------------------------------------------------------------------------------------|
|          | Observed | Calculated | Mass difference | Post-translational modifications* |
|          | Da       | Da         |                |                                |
| 75 kDa   | ND       | 76,960.5   |                 | ∆ import, 4Fe-4S, 2Fe-2S        |
| 51 kDa   | 48,502.5 | 48,499.4   |                 | ∆ import, 4Fe-4S               |
| 48 kDa   | 49,198.8 | 49,174.6   | +24.2           | ∆ import, Fe-S                 |
| 30 kDa   | 26,434.2 | 26,431.9   |                 | ∆ import                      |
| 24 kDa   | 23,814.8 | 23,814.5   |                 | ∆ import, 2Fe-2S               |
| PSST     | 20,093.6 | 20,077.6   | +16.0           | ∆ import, 4Fe-4S?              |
| TYKY     | 20,194.1 | 20,196.0   |                 | ∆ import, 2 × 4Fe-4S           |
| 42 kDa   | 36,705.0 | 36,707.0   |                 | ∆ import                      |
| 39 kDa   | 39,122.7 | 39,115.1   | −7.6            | ∆ import                      |
| 18 kDa   | 15,337.5 | 15,337.3   |                 |                               |
| 15 kDa   | 12,354.4 | 12,667.6   | −133.2          | −Met                          |
| 13 kDa   | 10,534.4 | 10,535.7   |                 | ∆ import                      |
| 10 kDa   | 8,438.3  | 8,437.4    |                 | ∆ import                      |
| AGGG     | 8,493.4  | 8,493.4    |                 | ∆ import                      |
| ASHI     | 18,738.3 | 18,737.0   |                 | ∆ import                      |
| ESSS     | 14,451.7 | 14,453.1   |                 | ∆ import                      |
| KFYI     | 5,829.0  | 5,828.7    |                 | ∆ import                      |
| MLRQ     | 9,323.3  | 9,324.7    |                 | None                          |
| MNLL     | 6,961.0  | 7,097.4    | −131.3          | −Met                          |
| MWFE     | 8,106.0  | 8,105.4    |                 | None                          |
| PDSW     | 20,832.7 | 20,964.9   | −132.2          | −Met                          |
| PGIV     | 19,959.1 | 20,092.2   | −132.1          | −Met                          |
| SDAP     | 10,674.2 | 10,109.6   | +564.6          | ∆ import, ACP                 |
| SGDH     | 16,727.9 | 16,726.4   |                 | ∆ import                      |
| B22      | 21,698.9 | 21,788.9   | −90.0           | −Met + acetyl                 |
| B18      | 16,477.9 | 16,397.8   | +80.1           | −Met + myristyl               |
| B17.2    | 17,131.2 | 17,089.5   | +41.7           | +Acetyl                       |
| B17      | 15,436.7 | 15,524.1   | −87.4           | −Met + acetyl                 |
| B16.6    | 16,583.9 | 16,673.5   | −89.6           | −Met + acetyl                 |
| B15      | 15,039.8 | 15,184.3   | −89.1           | −Met + acetyl                 |
| B14.7    | ND       | 14,758.1   | −90.4           | −Met + acetyl                 |
| B14.5a   | 12,586.2 | 12,675.1   | −90.4           | −Met + acetyl                 |
| B14.5b   | 14,095.8 | 14,096.3   |                 | None                          |
| B14      | 14,136.3 | 14,095.8   | +40.0           | +Acetyl                       |
| B14      | 14,965.4 | 15,035.5   | −88.1           | −Met + acetyl                 |
| B13      | 13,226.2 | 13,315.7   | −89.5           | −Met + acetyl                 |
| B12      | 11,037.5 | 11,140.7   | −103.0          | −Met + X*                     |
| B9       | 9,261.0  | 9,348.9    | −87.9           | −Met + acetyl + X*            |
| B8       | 10,990.3 | 11,079.8   | −89.5           | −Met + acetyl                 |

* Subunit names prefixed by B gave no N-terminal sequence by Edman degradation. They have modified (“blocked”) N termini. All other subunits gave N-terminal sequences by Edman degradation that have been reported before (1).

* A ∆ import indicates that the DNA sequence encodes an N-terminal extension that acts as a mitochondrial import sequence. This import sequence is not in the mature protein.

* The 24-kDa subunit is known to contain a [2Fe-2S] cluster (50), and the TYKY subunit contains canonical ligation motifs for two [4Fe-4S] clusters (51). The 51-kDa subunit is thought to contain one [4Fe-4S] cluster (1, 52, 53). The 75-kDa subunit contains 11 conserved cysteines that are likely to ligate one [4Fe-4S] and one [2Fe-2S] cluster (1), although a second [4Fe-4S] cluster has also been suggested (53). The location of the [4Fe-4S] cluster “N2” remains uncertain. It is possible that it is coordinated by three cysteines from the PSST subunit and either a non-cysteine ligand or a fourth cysteine from the 49-kDa subunit (38). In the acidic conditions used for HPLC and electrospray ionization, the Fe-S clusters are lost from the protein, and so they do not influence protein molecular mass measurements.

* ND, not determined.

* Calculated with residue 393 of the 51-kDa subunit and residue 255 of the 42-kDa subunit as tryptophan and lysine.

* The bovine cDNA codes for residue 3 onward, and residues 1 and 2 were determined by direct protein sequencing. The human cDNA sequence encodes a plausible import sequence.

* The cDNA encodes the sequence MMNLl. In the mature protein, methionine 1 is mostly removed. The observed and calculated masses refer to the sequences MMNLL... and MNLL..., respectively. See Supplemental Data Fig. S2-1.

* The bovine cDNA for subunit SDAP was extended in a 5’ direction ~500 bp beyond the codon for residue 1. This sequence did not contain either a translational initiator or a stop codon in-phase. Also the encoded protein sequence did not have the characteristic features of a mitochondrial import sequence. Therefore, it was concluded that either the 5’ sequence had been added artefactually or that it represented an unspliced intron (M. J. Runswick and J. E. Walker, unpublished results). The corresponding human cDNA appears to encode an import sequence.

122 Molecular & Cellular Proteomics 2.2
mass of the mature protein sequence by 28 Da. Apart from residues 1–12, peptide mass mapping of fragments produced by endoproteinase Lys-C, trypsin, and CNBr cleavage have excluded the presence of post-translational modifications in the protein. Histidine residues are found at residues 2, 4, 6, and 8, and the methylation of two of these residues would account for the mass difference, but at present there are no mass measurements on peptides containing residues 1–12.

The observed and calculated mass values for the 49-kDa, 39-kDa, and PSST subunits do not agree (see Table I). Tryptic peptide mass fingerprinting, tandem MS sequencing, and database searches, including expressed sequence tags, have not produced any evidence for either polymorphisms or post-translational modifications. Therefore, these discrepancies remain unexplained at present.

DISCUSSION

Practical Considerations—This reanalysis of the subunit composition of bovine complex I has benefited from improvements in the quality of the complex and its subcomplexes and from advances in protein analytical methods to separate the subunits and to analyze their sequences. Three different methods have been applied to the fractionation of the subunits of complex I and its subcomplexes. Each of them exploits different properties of the subunits, and since the three methods lead to protein compositions that differ only in minor details, it is increasingly unlikely that any more subunits of complex I from bovine heart mitochondria remain undetected.

In the present analysis, a significant improvement in 2D gels derives from replacing isoelectric focusing tube gels with IPG strips (31). Many of the small subunits of the complex are rather basic and in earlier analyses were lost from isoelectric focusing tube gels even before reaching equilibrium conditions (for example, see Refs. 32 and 33), whereas in the present work, the use of IPG strips allowed the gels to be run to equilibrium without losing basic subunits. Another improvement came from the introduction of the detergent ASB-14 (in place of CHAPS) to improve the solubilization of hydrophobic proteins (34). The reanalysis has also been aided by recent developments in mass spectrometric analysis of proteins. For example, tandem mass spectrometry has been useful for obtaining amino acid sequences from complex mixtures of peptides, especially from proteins with modified N termini.

Nonetheless the task of deciding what is an authentic subunit of such a complicated mammalian enzyme remains a difficult task because methods of authentication that can be used with simpler complexes are not readily applicable for practical reasons. For example, reconstitution of the enzyme from individual subunits in vitro would be a formidable task for such a complex membrane-associated enzyme, and mutagenesis of individual subunits, which has been applied successfully to complex I in euobacteria and lower eukaryotes, is much more difficult to apply to the bovine enzyme. In addition to the complexity of subunits that constitute the mammalian assembly, various other properties have hampered the determination of its subunit composition. They include the hydrophobicity of many subunits and the significant number of subunits with modified N termini, which hinder or prevent sequence analysis by Edman degradation.

These difficulties notwithstanding, various criteria have been helpful in deciding whether a protein is an authentic subunit of the complex. They include the presence of the subunit in the intact complex and in its subcomplexes and also in many independently isolated samples of both, the absence of potential complex I subunits from other well-characterized protein complexes from the inner membranes of bovine mitochondria, and the presence of homologous subunits in complex I isolated from a range of species. Examples of the utility of the last criterion are the conservation of the "core" subunits of bovine complex I as the minimal complex I in euobacteria (1) and the presence of these and other nuclear encoded subunits in complex I in fungal mitochondria. However, the analyses of the complexes from N. crassa (35, 36) and Yarrowia lipolytica (37) are incomplete, but the current data in N. crassa confirm the presence of homologues of a range of bovine subunits. They suggest also that there may be subunits that are confined to fungi and others that are found only in mammalian enzymes.3

The Subunit Composition of Bovine Complex I—From the analyses described above and in Figs. 1–4, it appears that bovine complex I is an assembly of 46 different proteins, including the seven proteins that are encoded in mitochondrial DNA and the unsequenced (nuclear encoded) protein with a mass of 10,566 Da. The 38 identified nuclear encoded subunits are listed in Table I. The 42-kDa subunit is loosely bound to the complex. A total of 36 subunits have been detected in at least one of the three subcomplexes (see Table II), the exceptions being the KFYI and MLRQ subunits. It should be noted that subunit MLRQ has been found in a minority of the preparations that have now been analyzed, although it has not been found as a component of any other

3 J. Hirst, J. Carroll, I. M. Fearnley, R. J. Shannon, and J. E. Walker, submitted.

1 ACP, acyl carrier protein with serine 44 modified by pantetheine-4′-phosphate with 3-hydroxytetradecanoic acid probably attached via a thioester linkage (see “Results”).

2 This subunit contains a canonical myristylation signal near the N terminal (24). The presence of the N-terminal myristyl group has been verified by tandem MS experiments (J. Carroll and I. M. Fearnley, unpublished results).

3 X indicates that the subunit has an additional modification that has not been characterized fully, probably two methyl groups on histidine residues at positions 2, 4, 6, and 8.
The Subunit Composition of Bovine Complex I

Table II
Subunit compositions of subcomplexes of bovine complex I

| Subcomplex | Subunit Composition |
|------------|---------------------|
| Iα         | 75 kDa              |
| Iβ         | 75 kDa              |
| Iγ         | 51 kDa              |
| Iδ         | 51 kDa              |
| Iε         | 49 kDa              |
| Iθ         | 49 kDa              |
| Iφ         | 30 kDa              |
| Iψ         | 30 kDa              |
| Iξ         | 24 kDa              |
| Iη         | 24 kDa              |
| PSST       | PSST               |
| TYKY       | PSST               |
| TYKY       | SDAPα,β            |
| 18 kDa     | 18 kDa              |
| 13 kDa     | 13 kDa              |
| 10 kDa     | 10 kDa              |
| B17.2      | B17.2              |
| B16.6      | B16.6              |
| B14.7      | B14.7, α, d        |
| B14.5a     | B14.5a             |
| B13        | B13                |
| B8         | B8                 |
| (42 kDa)   |                   |
| 39 kDa     |                   |
| 15 kDa     |                   |
| MWFE       |                   |
| PGIV       |                   |
| SDAPα,β,d  |                   |
| (B15)      |                   |
| B14        |                   |
| B9         |                   |
| ND6        |                   |

* Not detected in 2D gels.
* Present in substantial amounts in both subcomplexes Iα and Iβ.
* Not detected in 1D gels.
* Not detected by HPLC.
* Detected by immunoblotting of 1D gels with a polyclonal antibody.

Subcomplex Iα contains 15 subunits at apparent stoichiometric levels (12). They include all seven of the nuclear encoded core subunits, which are conserved across a wide range of species and contain all of the bound redox cofactors of the complex. The subcomplex is predominantly hydrophilic, and it constitutes the globular arm of the complex, which protrudes into the matrix. However, in the absence of detergent it tends to aggregate, probably because subunit B16.6 has one predicted transmembrane helix (12). The presence of low levels of subunit B14.7, which also has predicted transmembrane helices (14), may also contribute to this process. Two related subcomplexes, IαS and IS, prepared by varying the composition of the chaotropic detergent, have also been reported, but their subunit compositions have not been characterized fully (13).

The 15 subunits of subcomplex Iα are a subset of the 23 present in approximately stoichiometric amounts in subcomplex (Table II), which is consistent with the recovery of the subcomplex Iα from subcomplex Iα treated appropriately with N,N-lauryledimethylamine oxide (17). The additional eight subunits are predicted to contain nine to 10 transmembrane helices, five of them in ND6, which supports the view that subcomplex Iα consists of the extrinsic arm and part of the membrane arm of complex I.

The compositions of the subcomplexes reported here are similar, but not identical, to compositions reported previously. For example, in one analysis of subcomplex Iα, the additional subunits MLRQ and ND2 (<10%) were reported to be present, but subunits B17.2, B16.6, B14.7, B14.5a, B15, and ND6 were not detected (15). In another independent preparation of subcomplex Iα, additional subunits B15 and B14 were found, but subunits B17.2, B16.6, and B14.7 were not detected (17). These differences may also reflect minor variations either in splitting the complex or in purification of subcomplexes or both. The most important point is that the preparation procedure now used yields a subcomplex Iα of consistent subunit composition.

In the purification of subcomplexes Iα and Iβ, several subunits were found in the breakthrough fractions, and almost pure 42-kDa subunit was recovered in fractions before subcomplex Iα. Similar breakthrough fractions have been referred to as subcomplex Iγ (17). However, this fragment splits readily into two smaller fragments, and so it is not clear that subcomplex Iγ represents a true fragment of complex I. The reported protein composition of subcomplex Iγ is similar to that of the breakthrough fractions plus the 42-kDa subunit, although in the present work, several additional subunits (PGIV, B15, B14.5b, B14.7, B9, and MWFE) were detected. These differences may reflect variation in the initial fragment-
tation and separation procedures. However, most of these subunits are found also in one or more of subcomplexes Iα, Iβ, and Iβ (see Table II). Only subunits KFYI, ND1, ND2, ND3, and ND4L are found uniquely in the fractions known as subcomplex Iγ (17). Both subunits ND1 and ND4L are in the breakdown through fractions and also in the fraction containing the 42-kDa subunit. It may be that rather than being an integral subcomplex these subunits represent a mixture of several small fragments that “splitter” away during the preparation of subcomplexes Iα and Iβ.

Concluding Remarks—Complex I from bovine heart mitochondria is one of the most complicated enzymes known. It is a membrane-bound assembly made of 46 different subunits that contains an FMN and up to eight Fe-S clusters.3 Assuming that the complex contains one copy of each of the 46 subunits, its molecular mass is 980 kDa.

It is important to establish its subunit composition for two reasons. First, knowledge of the subunit composition is an essential component of the process of determining the structure of complex I and of understanding its enzymic mechanism. Inevitably the analysis of the subunit composition has raised new issues that require biochemical explanation. Foremost among them are the roles of the numerous supernumerary subunits, especially those such as the acyl carrier protein, with biochemical properties predicted from their sequences that are not apparently related to the electron transfer and associated proton pumping activities of complex I (1). Second, it provides a basis for the characterization of the human enzyme. Human complex I has an increasing medical importance as the catalogue of pathologies grows. The list is dominated by neuromuscular diseases that are associated with mutations in both mitochondrially and nuclear encoded subunits. Other diseases may be related to mutations in undefined factors that influence the assembly of the complex (41). The human enzyme is much less readily available in the quantities required for an exhaustive chemical analysis of its subunits. So far it has been isolated only in small quantities by native gel electrophoresis. Before the advent of the human genome, the cDNAs for many nuclear encoded subunits of bovine complex I were used to clone human homologues (42–44). These experiments demonstrated a very high degree of sequence conservation between the two species, and other proteins identified later in the human genome sequence as subunits of complex I by homology with bovine sequences support this conclusion. Antibodies raised against some recombinant bovine subunits have been used to confirm directly the presence of homologues in the human enzyme (45, 46), but for most nuclear encoded subunits direct confirmation of the presence of homologues of bovine subunits in the human complex is lacking. Another aspect of mammalian complex I that remains unexplored is whether other tissues contain isoforms of some of the supernumerary heart subunits similar to the tissue-specific isoform subunits of cytochrome-c oxidase (47–49). Thus, the exact definition of human complex I in all tissues will require more experimentation with both bovine and human material.

Note Added in Proof—Two studies (54, 55) have confirmed the presence in human mitochondria of 44 homologues of the 45 sequenced bovine complex I subunits. Subunit ND4L was not detected.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Appendix 1—The on-line version of this article (available at http://www.mcponline.org) contains a supplemental section. Part 1 summarizes the MS data used to identify components of subcomplexes Iα and Iβ and the protein compositions of other fractions reported in Figs. 1, B and D, and 2, A and B (Tables S1-1 to S1-4 and Figs. S1-1 to S1-54). Part 2 contains tandem MS data supporting some of the sequence observations reported in this paper (Figs. S2-1 and S2-2). Part 3 contains electrospray ionization MS data supporting the new mass data reported in Table I (Figs. S3-1 to S3-7).

‡ To whom correspondence should be addressed: Medical Research Council Dunn Human Nutrition Unit, The Welcombe Trust/MRC Bldg., Hills Rd., Cambridge CB2 2XY, UK. Tel.: 44-1223-252701; Fax: 44-1223-252705; E-mail: walker@mrc-dunn.cam.ac.uk.

REFERENCES

1. Walker, J. E. (1992) The NADH-ubiquinone oxidoreductase (complex I) of respiratory chains. Q. Rev. Biophys. 25, 253–324
2. Weiss, H., Friedrich, T., Hofhaus, G., and Preis, D. (1991) The respiratory-chain NADH dehydrogenase (complex I) of mitochondria. Eur. J. Biochem. 197, 563–576
3. Saraste, M. (1999) Oxidative phosphorylation at the mitochondrial NADH-ubiquinone oxidoreductase (complex I). FEBS Lett. 444, 216–220
4. Schulz, B. E., and Chan, S. I. (2001) Structures and proton-pumping strategies of mitochondrial respiratory enzymes. Annu. Rev. Biophys. Biomol. Struct. 30, 23–65
5. Wikström, M. (1984) Two protons are pumped from the mitochondrial matrix per electron transferred between NADH and ubiquinone. FEBS Lett. 169, 300–304
6. Smelting, J., Sengers, R., Trijbels, F., and van den Heuvel, L. (2001) Human NADH:ubiquinone oxidoreductase. Bioenerg. Biomembr. 33, 259–266
7. Smelting, J., van den Heuvel, L., and Di Mauro, S. (2001) The genetics and pathology of oxidative phosphorylation. Nat. Rev. Genet. 2, 342–352
8. Chomyn, A., Mariotti, P., Cleeter, M. W. J., Ragan, C. I., Matsuno-Yagi, A., Hafet, Y., Doolittle, R. F., and Attardi, G. (1985) Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory-chain NADH dehydrogenase. Nature 314, 592–597
9. Chomyn, A., Cleeter, M. W. J., Ragan, C. I., Riley, M., Doolittle, R. F., and Attardi, G. (1986) URF6, last unidentified reading frame of human mitochondrial DNA. Science 234, 614–618
10. Grigorieff, N. (1999) Structure of the respiratory NADH:ubiquinone oxidoreductase complex I, Curr. Opin. Struct. Biol. 9, 476–483
11. Guénebaut, V., Schiltt, A., Weiss, H., Leonard, K., and Friedrich, T. (1998) Consistent structure between bacterial and mitochondrial NADH:ubiquinone oxidoreductase (complex I), J. Mol. Biol. 276, 105–112
12. Fearnley, I. M., Carroll, J., Shannon, R. J., Runswick, M. J., Walker, J. E., and Hirst, J. (2001) GRIM-19, a cell death regulatory gene product, is a subunit of bovine mitochondrial NADH:ubiquinone oxidoreductase (complex I), J. Biol. Chem. 276, 38345–38348
13. Finel, M., Majander, A. S., Tyynelä, J., Dejong, A. M. P., Albracht, S. P. J., and Wikström, M. (1994) Isolation and characterisation of subcomplexes of the mitochondrial NADH-ubiquinone oxidoreductase (complex I), Eur. J. Biochem. 226, 237–242
14. Carroll, J., Shannon, R. J., Fearnley, I. M., Walker, J. E., and Hirst, J. (2002) Definition of the nuclear encoded protein composition of bovine heart mitochondrial complex I. Identification of two new subunits. J. Biol. Chem. 277, 50311–50317
15. Finel, M., Skehel, J. M., Albracht, S. P. J., Fearney, I. M., and Walker, J. E. (1992) Resolution of NADH-ubiquinone oxidoreductase from bovine heart mitochondria into two subcomplexes, one of which contains the nuclear-encoded import precursor of mitochondrial DNA- and nuclear-encoded subunits of respiratory NADH dehydrogenase (complex I) from the strictly aerobic yeast Yarrowia lipolytica. Biochim. Biophys. Acta 1555, 187–191

16. Walker, J. E., Skehel, J. M., and Buchan, S. K. (1995) Structural analysis of NADH:ubiquinone oxidoreductase from bovine heart mitochondria. Methods Enzymol. 260, 14–34

17. Sazanov, L. A., Peak-Chew, S. Y., Fearney, I. M., and Walker, J. E. (2000) Resolution of the membrane domain of bovine complex I into subcomplexes: implications for the structural organization of the enzyme. Biochemistry 39, 7229–7235

18. Schlager, H., and von Jagow, G. (1987) Tricine-sodium dodeyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. 166, 388–379

19. Skehel, J. M., Fearney, I. M., and Walker, J. E. (1999) NADH:ubiquinone oxidoreductase from bovine heart mitochondria: sequence of a novel 17.2-kDa subunit. FEBS Lett. 438, 301–305

20. Wilm, M., Shevchenko, A., Huthaevé, T., Brit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379, 466–469

21. van Montfort, B. A., Casas, B., Duurkens, R., Godovac-Zimmermann, J., and Robillard, G. T. (2002) Improved in-gel approaches to generate peptide maps of integral membrane proteins with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J. Mass Spectrom. 37, 322–330

22. Robinson, N. E., and Robinson, A. B. (2001) Molecular clocks. Proc. Natl. Acad. Sci. U. S. A. 98, 844–849

23. Robinson, N. E. (2002) Protein degradation. Proc. Natl. Acad. Sci. U. S. A. 99, 5283–5288

24. Walker, J. E, Arizmendi, J. M., Dupuis, A., Fearney, I. M., Finel, M., Meed, S. M., Pilkington, S. J., Runswick, M. J., and Skehel, J. M. (1992) Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from bovine heart mitochondria. Application of a novel strategy for sequencing proteins using the polymerase chain reaction. J. Mol. Biol. 226, 1051–1072

25. Runswick, M. J., Fearney, I. M., Skehel, J. M., and Walker, J. E. (1991) Presence of an acyl carrier protein in NADH:ubiquinone oxidoreductase from bovine heart mitochondria. FEBS Lett. 286, 121–124

26. Brody, S., and Mikolajczyk, S. (1988) Neurospora mitochondria contain an acyl carrier protein. Eur. J. Biochem. 173, 353–359

27. Sackmann, U., Zensen, R., Houthaeve, T., Brit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379, 466–469

28. van Montfort, B. A., Casas, B., Duurkens, R., Godovac-Zimmermann, J., and Robillard, G. T. (2002) Improved in-gel approaches to generate peptide maps of integral membrane proteins with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J. Mass Spectrom. 37, 322–330

29. Robinson, N. E., and Robinson, A. B. (2001) Molecular clocks. Proc. Natl. Acad. Sci. U. S. A. 98, 844–849

30. Robinson, N. E. (2002) Protein degradation. Proc. Natl. Acad. Sci. U. S. A. 99, 5283–5288

31. Walker, J. E, Arizmendi, J. M., Dupuis, A., Fearney, I. M., Finel, M., Meed, S. M., Pilkington, S. J., Runswick, M. J., and Skehel, J. M. (1992) Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from bovine heart mitochondria. Application of a novel strategy for sequencing proteins using the polymerase chain reaction. J. Mol. Biol. 226, 1051–1072

32. Runswick, M. J., Fearney, I. M., Skehel, J. M., and Walker, J. E. (1991) Presence of an acyl carrier protein in NADH:ubiquinone oxidoreductase from bovine heart mitochondria. FEBS Lett. 286, 121–124

33. Brody, S., and Mikolajczyk, S. (1988) Neurospora mitochondria contain an acyl carrier protein. Eur. J. Biochem. 173, 353–359

34. Sakmann, U., Zensen, R., Houthaeve, T., Brit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379, 466–469

35. van Montfort, B. A., Casas, B., Duurkens, R., Godovac-Zimmermann, J., and Robillard, G. T. (2002) Improved in-gel approaches to generate peptide maps of integral membrane proteins with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J. Mass Spectrom. 37, 322–330

36. Robinson, N. E., and Robinson, A. B. (2001) Molecular clocks. Proc. Natl. Acad. Sci. U. S. A. 98, 844–849

37. Robinson, N. E. (2002) Protein degradation. Proc. Natl. Acad. Sci. U. S. A. 99, 5283–5288

38. Fearney, I. M., Finel, M., Skehel, J. M., and Walker, J. E. (1991) NADH:ubiquinone oxidoreductase from bovine heart mitochondria. cDNA sequence of the import precursors of the nuclear-encoded import precursor of mitochondrial DNA- and nuclear-encoded subunits of respiratory NADH dehydrogenase detected with peptide- and subunit-specific antibodies in mitochondrial myopathies. Biochim. Biophys. Acta 1234, 63–73

39. Fearney, I. M., Finel, M., Skehel, J. M., and Walker, J. E. (1991) Presence of an acyl carrier protein in NADH:ubiquinone oxidoreductase from bovine heart mitochondria. FEBS Lett. 286, 121–124

40. Brody, S., and Mikolajczyk, S. (1988) Neurospora mitochondria contain an acyl carrier protein. Eur. J. Biochem. 173, 353–359