Identification of a Novel Subunit of Respiratory Complex I from *Thermus thermophilus*†

Philip Hinchliffe, Joe Carroll, and Leonid A. Sazanov*

Medical Research Council, Dunn Human Nutrition Unit, Wellcome Trust/MRC Building, Hills Road, Cambridge CB2 2XY, U.K.

Received January 17, 2006; Revised Manuscript Received February 23, 2006

ABSTRACT: The hydrophilic domain (peripheral arm) of the proton-translocating NADH:quinone oxidoreductase (complex I) from the thermophilic organism *Thermus thermophilus* HB8 has been purified and characterized. The subcomplex is stable in sodium dodecyl sulfate up to 80 °C. Of nine iron−sulfur clusters, four to five (one or two binuclear and three tetranuclear) could be detected by EPR in the NADH-reduced enzyme. The preparation consists of eight different polypeptides. Seven of them have been positively identified by peptide mass mapping and N-terminal sequencing as known hydrophilic subunits of *T. thermophilus* complex I. The eighth polypeptide copurified with the subcomplex at all stages, is strongly associated with the other subunits, and is present in crystals of the subcomplex, used for X-ray data collection. Therefore, it has been identified as a novel complex I subunit and named Nqo15. It is encoded in a locus separate from the *nqo* operon, containing the 14 other known complex I genes. ORFs encoding Nqo15 homologues are present in the genomes of the closest relatives of *T. thermophilus.* Our data show that, contrary to previous assumptions, bacterial complex I can contain proteins in addition to a "core" complement of 14 subunits.

NADH:ubiquinone oxidoreductase, or complex I, is the first of four enzyme complexes in the respiratory chain (*1*). Complex I catalyzes the transfer of two electrons from NADH to quinone coupled to the translocation of four protons across the inner mitochondrial membrane, thus helping to provide the proton-motive force required for the synthesis of ATP. The complex contains one flavin mononucleotide (FMN)[1] and eight or nine iron−sulfur (FeS) clusters (*2*−*4*). The mitochondrial enzyme contains 46 different subunits (*5*) while the bacterial enzyme is thought to contain just 14 (the "core" subunits). These 14 subunits of bacterial complex I all have their homologues in mitochondrial complex I (*6*). In addition, bacterial complex I presents many of the same properties as the mitochondrial one, such as noncovalently bound FMN, multiple iron−sulfur clusters, and its distinctive L shape. Therefore, the bacterial enzyme represents a useful "minimal" model for complex I studies.

Within the L shape of complex I are two major domains: the membrane arm and the peripheral arm (hydrophilic domain). The latter domain protrudes into the bacterial cytoplasm and has been shown to contain the NADH binding site as well as at least six electron paramagnetic resonance (EPR)-detectable FeS clusters (N1a, N1b, and N2−N5) and the FMN (*1*, *7*). In addition, some bacterial complex I enzymes (e.g., *Escherichia coli* and *Thermus thermophilus*) contain the tetranuclear cluster N7, though this has not been detected by EPR in the intact enzyme (*8*). Furthermore, the tetranuclear clusters N6a and N6b have shown EPR signals in subcomplexes and overexpressed subunits but not in the intact enzyme (*9*, *10*). Recently, we reported the organization of the iron−sulfur clusters in the hydrophilic domain of complex I from *T. thermophilus* (*4*). There are nine clusters, of which seven [six tetranuclear (N2−N5, N6a, and N6b) and one binuclear (N1b)] are involved in the electron transfer chain connecting NADH and quinone binding sites. Cluster N7 is not likely to be a part of the main pathway due to its distal location, while binuclear cluster N1a may act as an antioxidant (*4*).

The only available structural information for intact complex I is from electron microscope analysis and of low resolution. Most reports have agreed on the L-shaped structure, though it has been proposed that, in *E. coli*, this constitutes an inactive form of the enzyme, which assumes an alternative, horseshoe-like structure upon activation (*11*). However, we observed an L shape of *E. coli* complex I under conditions optimal for activity (*12*), and no other "horseshoe-like" structures have been reported in the literature.

The thermophilic organism *T. thermophilus* is potentially a source of a stable enzyme for structural studies. This rod-shaped bacterium was originally isolated from a Japanese thermal spa (*13*). The optimal growth temperature for the organism is between 65 and 72 °C. Previous efforts to obtain complex I from this bacterium have not resulted in pure preparations, and no subunits were conclusively identified (*14*). Here we report a characterization of the hydrophilic

† This research was funded by the Medical Research Council.

* To whom correspondence should be addressed. Telephone: +44 1223 252910. Fax: +44 1223 252915. E-mail: sazanov@mrc-dunn.cam.ac.uk.

[1] Abbreviations: FMN, flavin mononucleotide; FeS, iron−sulfur; EPR, electron paramagnetic resonance; DDM, dodecyl maltoside; BTP7.3, bis-tris propane (pH 7.3); PMSF, phenylmethanesulfonyl fluoride; PIC-E, Roche's Complete Protease Inhibitor Cocktail-EDTA free; FeCy, ferricyanide; OG, octyl glucoside; BCA, bicinchoninic acid; PVDF, polyvinylidene fluoride; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; PEG, polyethylene glycol.

10.1021/bi0600998 CCC: $33.50
Published on Web 03/17/2006

domain of complex I purified from *T. thermophilus* HB8 cells. The membrane arm dissociates early during purification. On its own, the hydrophilic domain was found to be extremely stable and all subunits were positively identified by peptide mass mapping and N-terminal sequencing. A previously unknown polypeptide has been found to associate strongly and crystallize with the subcomplex, indicating that it is a novel subunit of the complex, which we named Nqo15.

EXPERIMENTAL PROCEDURES

Materials. All detergents were purchased from Glycon (Luckenwalde, Germany). Chromatography columns and instrumentation were from Amersham Biosciences (Uppsala, Sweden) except the BioScale DEAE and UNO Q-12 columns were from Bio-Rad (Hemel Hampstead, Herts, U.K.). Complete protease inhibitor tablets were from Roche Diagnostics (Lewes, U.K.), and all other chemicals were from Sigma (Poole, Dorset, U.K.).

Purification of the Hydrophilic Domain. The subcomplex was purified from *T. thermophilus* HB8 cells which were either purchased from the University of Georgia Bioexpression and Fermentation Facility or grown in house in a 60 L fermentor. Cells were grown at 70 °C under reduced oxygen (~5% dissolved) to induce complex I expression, until late log phase. All subsequent steps were performed at 4 °C. Cells (175 g) were resuspended at 0.3 g/mL in RBP buffer [50 mM bis-tris propane (pH 7.3) (BTP7.3), 0.002% phenylmethanesulfonyl fluoride (PMSF), and Roche's Complete Protease Inhibitor Cocktail-EDTA free (PIC-E) tablet] using a glass−Teflon homogenizer and passed through a Z-plus 2.2 kW disruptor twice at 15 000 psi and three times at 30 000 psi. Cell debris was removed by centrifugation at 9600*g* (15 min) followed by centrifugation at 18800*g* for 35 min. The membrane fraction was collected by centrifugation at 150000*g* for 4 h, the supernatant discarded, and the pellet resuspended in 150 mL of RBP in a glass−Teflon homogenizer. Membranes were used immediately or stored at −80 °C. Dodecyl maltoside (DDM, 10% solution) was added dropwise to stirred membranes to a final level of 3%. PMSF (0.002%), a PIC-E tablet, and NaCl (to a final concentration of 80 mM) were also added. The solution was stirred on ice for 2 h; nonsolubilized material was removed by centrifugation at 150000*g* for 1 h, and the supernatant was passed through a 0.45 μm filter. The sample was passed through a series of three ion-exchange columns and a final gel filtration step using an ÄKTA Explorer 100 instrument. Eluate absorbance was monitored at 280 and 420 nm. All ion-exchange columns were equilibrated before use with buffer A (20 mM BTP7.3, 0.1% DDM, 10% glycerol, and 0.002% PMSF). The sample was applied to a HiLoad 26/10 Q-Sepharose column and eluted with a 60 mL gradient from 8 to 12% followed by a 700 mL gradient from 12 to 22% of buffer B (A with 1 M NaCl). Fractions (14 mL) were collected and those with NADH:ferricyanide (FeCy) activity pooled and diluted with an equal volume of buffer A. Pooled, diluted material from the Q-Sepharose column was applied to a Bio-Rad Bio-Scale DEAE 20 column. This was eluted with a 10 mL gradient from 0 to 18% followed by a 700 mL gradient from 18 to 32% of buffer B. Fractions (9 mL) were collected and those with NADH:FeCy activity pooled and diluted with 4 volumes of buffer A. Pooled, diluted material from the DEAE column was applied to a Bio-Rad UNO Q-12 column. This was eluted with a 10 mL gradient from 0 to 8% followed by an 800 mL gradient from 8 to 18% of buffer B. Fractions (9 mL) were collected, assayed for NADH:FeCy activity, and analyzed by SDS−PAGE. Selected fractions were pooled and concentrated (Vivacell 70; 50 000 molecular weight cutoff) to 3 mL. The concentrate was diluted to 20 mL in buffer 1 [20 mM BTP7.3 and 1% octyl glucoside (OG)] and exchanged into 3 mL of buffer 2 (20 mM BTP7.3, 1% OG, and 0.5 M NaCl) using a Vivapure Maxi Q Spin Column (Vivascience). Protein was concentrated to ~0.5 mL (Vivaspin 20; 100 000 molecular weight cutoff) and applied to a Superdex 200 10/30 gel filtration column equilibrated in buffer SFG20 (20 mM BTP7.3, 1% OG, 0.002% PMSF, and 150 mM NaCl). Eluted fractions (0.5 mL) were assayed for NADH:FeCy activity and analyzed by SDS−PAGE. Selected fractions were pooled, concentrated to 1 mL (Vivaspin 20; 50 000 molecular weight cutoff) and diluted with 3 volumes of buffer 1. Material was concentrated further to give 0.5 mL of protein at 12−15 mg/mL in 20 mM BTP7.3, ~40 mM NaCl, and 1% OG. Protein was used immediately or stored under liquid nitrogen in 10−100 μL aliquots. A short summary of the purification procedure was given in our recent publication (*4*).

Analytical Methods. Protein concentrations were measured using the Pierce bicinchoninic acid (BCA) protein assay according to the manufacturer's instructions. SDS−PAGE was performed with pre-prepared Novex Tris-glycine polyacrylamide gels containing a 10 to 20% acrylamide gradient (Invitrogen), according to the manufacturer's instructions. Unless otherwise stated, any samples loaded onto the gels were boiled at 105 °C for 3 min immediately before being used. Protein bands were visualized with Coomassie Blue R250 or silver staining according to a standard protocol. The FMN content in the preparation was assayed fluorometrically (*15*), and the iron contents were assayed by spectrophotometry (*16*). Oxidation of NADH or deamino-NADH by complex I was assessed at 340 nm and 50 °C, the maximal temperature at which assays were possible without evaporation problems. Quinone analogue assays were performed by addition of the quinone analogue at 0.1 mM (decylubiquinone or ubiquinone-5) and ~8 μg of protein to assay buffer (10 mM BTP7.3, 25 mM NaCl, and 0.03% DDM). The reaction was started by addition of 0.1 mM NADH. With $K_3Fe(CN)_6$ (ferricyanide, FeCy) as the acceptor, 1−5 μg of protein and 1 mM potassium ferricyanide were added to buffer containing 50 mM potassium phosphate (pH 7.0) and 0.1% DDM, and the reaction was started by adding 0.1 mM NADH.

EPR Spectroscopy. Samples (~4 mg of protein/mL in 20 mM BTP7.3, 20 mM NaCl, and 1% OG) were reduced with 2 mM NADH and frozen immediately in liquid nitrogen. EPR spectra were recorded on a Bruker EMX X-band spectrometer using an ER 4119HS high-sensitivity cavity maintained at low temperature by an ESR900 continuous flow liquid helium cryostat (Oxford Instruments, Abingdon, U.K.); the sample temperature was measured with a calibrated Cernox resistor (Lake Shore Cryotronics Inc.). The spectra were recorded under nonsaturating conditions. The following conditions were applied: microwave frequency, 9.385 GHz; modulation amplitude, 10 G; modulation frequency, 100 kHz; and time constant, 82 ms.

Identification of Subunits. Purified protein was subjected to SDS−PAGE. For N-terminal sequencing, protein bands were electroblotted onto polyvinylidene fluoride (PVDF) membrane, located by Coomassie staining, and excised. N-Terminal sequencing was performed at the MRC Laboratory of Molecular Biology (Cambridge, U.K.). N-Terminal sequences were matched against the known amino acid sequences of the subunits of *T. thermophilus* complex I using the Swissprot database on the NCBI website (http://www.ncbi.nih.gov/). Reference sequences were submitted by Yano et al. (*17*).

Proteins resolved by SDS−PAGE and stained with Coomassie R250 were identified by peptide mass fingerprinting and tandem MS peptide sequence data. Excised gel bands were digested "in gel" (*18*) with trypsin [20 mM Tris-HCl (pH 8) and 5 mM $CaCl_2$ at 37 °C], without prior reduction and alkylation. Portions of the digest were examined in positive ion mode by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry with a TofSpec 2E instrument (Micromass, Altrincham, U.K.) in the presence of α-cyano-4-hydroxycinnamic acid as the matrix. Spectra were calibrated with trypsin autolysis peptides (*m*/*z* 2163.057 and 2273.160) and a matrix-related ion (*m*/*z* 1060.048). Peptide mass data were searched against an NCBInr protein database using Mascot (http://www.matrixscience.com) with trypsin constraints, variable modifications (methionine oxidation and cysteine propionamide), a maximum of one missed cleavage, and a mass tolerance of 70 ppm.

Peptide sequence analysis was performed on a Q-TOF mass spectrometer equipped with ESI (Micromass) and coupled on-line to a capillary high-performance liquid chromatography system (CapLC; Micromass). Peptide mixtures were separated using a PepMap C18 column (180 μm × 100 mm, LC Packings, Amsterdam, The Netherlands) with an acetonitrile gradient in 0.1% formic acid. Acquired tandem MS spectra were interpreted manually, assembled into Peptide Sequence Tags (*19*), and compared with protein sequence databases.

Crystallization. Crystallization screening was conducted using sitting drop vapor diffusion at 296 K with Corning 96-well microplates. Reservoir solutions (100 μL) were dispensed using a Genesis 150 Workstation (Tecan, U.K.) and crystallization drops (0.2 μL of reservoir solution and 0.2 μL of protein solution) established using a PixSys SQ series robot (Cartesian Technologies). An initial set of 1344 conditions were set up from commercially available kits: Crystal Screen, Grid Screen, MembFac, Natrix, Quick Screen (Hampton Research), Cryo I+II, Wizard I+II (Emerald Biostructures), JBScreen 1−10, JBScreen Membrane 1−3 (Jena Bioscience), Clear Strategy Screen I+II, MemStart, MemSys, Structure Screen I+II, and Stura Footprint Screens (Molecular Dimensions). Crystals were observed under a few of these conditions which contained similar precipitants [polyethylene glycol (PEG) 4000 or PEG 8000], salts (magnesium chloride, calcium chloride, or sodium chloride) and buffers [Tris-HCl (pH 8.5) and HEPES (pH 7.5)].

Refinement of crystal conditions was performed in CrysChem 24-well plates (Hampton Research) with drops consisting of 1 μL of protein and 1 μL of reservoir reagent equilibrated against a 700 μL reservoir. The best crystals were grown in reagents comprising 0.1 M HEPES (pH 7.5), 0.4−0.5 M NaCl, 0.1 M $MgCl_2$/$CaCl_2$, and 8.5−9.5% PEG 4000.

RESULTS

Purification Procedure. Like our observations with *E. coli* complex I (*12*), the content of complex I in *T. thermophilus* membranes is increased when cells are grown under a limiting oxygen supply, as judged from NADH:FeCy activity found in membranes. However, the dissolved oxygen concentration needed to be higher (∼5% vs 2% for *E. coli*) for optimal yields with *Thermus*, probably due to the limited oxygen solubility at high temperatures. Despite numerous trials using different detergents, the membrane arm of *T. thermophilus* complex I appears to be dissociated from the complex upon solubilization. This was indicated by gel filtration experiments with the solubilized sample: all NADH:FeCy activity eluted with an apparent molecular mass of ∼400 kDa, lower than the value of ∼600 kDa expected for the intact complex with detergent bound. Membranes were solubilized using 3% DDM and passed through a series of three ion-exchange columns and a final gel filtration step to obtain a pure preparation of the hydrophilic domain (peripheral arm) of complex I (Figure 1). Although DEAE and UNO-Q steps gave relatively small increases in specific activity, both of these steps were essential for obtaining a pure preparation. In the final gel filtration step, activity and nearly all protein coeluted, with an apparent molecular mass of ∼400 kDa, consistent with the combined molecular mass of the peripheral arm (280 kDa) and the bound detergent. Omitting the detergent at this step resulted in the aggregation of protein, indicating that some hydrophobic areas are present in the subcomplex, presumably at the interface with the membrane arm. SDS−PAGE analysis of fractions from the gel filtration step indicated that all putative subunits coeluted together and no significant impurities were present in the main peak (first five fractions at the left, Figure 2).

A summary of the purification procedure is given in Table 1. The final yield is sufficiently substantial for structural studies. The complex has a ferricyanide reductase activity of 106 ± 7.2 units (micromoles of NADH oxidized per minute per milligram of protein). This compares favorably with intact *E. coli* complex I, where a pure preparation gives a specific activity of 57.9 units (*12*). Activities with the quinone (and menaquinone) analogues decylubiquinone (1.66 ± 0.06 units) and ubiquinone-5 (1.79 ± 0.17 units) were significant but much lower than those found with the intact *E. coli* enzyme (*12*). This is likely due to the absence of the membrane arm in the *T. thermophilus* preparation. Complex I may have one or several quinone-binding sites, all of which are predicted to be located mostly within the membrane domain, due to the hydrophobic nature of this substrate (*1*). Assays of the FMN and iron contents of the preparation gave values of 33.8 ± 4.0 Fe atoms and 0.80 ± 0.08 FMN molecule per complex I. This result is consistent with the presence of 32 Fe atoms per monomer, from two binuclear and seven tetranuclear clusters, identified by X-ray crystallography (*4*), and indicates that one FMN molecule is bound per complex. This is in contrast to the suggestions that complex I may contain two bound FMNs (*20*).

Subunit Identification. SDS−PAGE analysis of the preparation showed only one band at the very top of the gel when

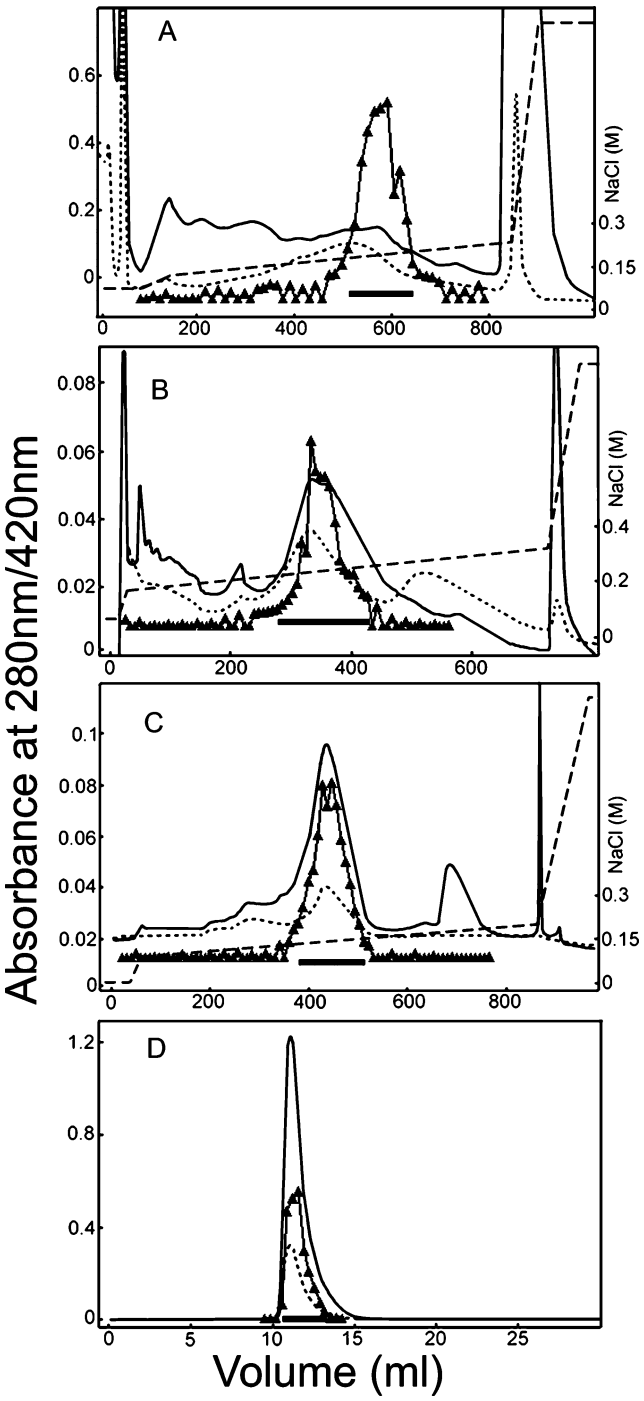

FIGURE 1: Purification of the hydrophilic domain of complex I from *T. thermophilus*: (A) Q-Sepharose column chromatography, (B) DEAE Bio-Scale column chromatography, (C) UNO-Q12 column chromatography, and (D) Superdex 200 10/30 column chromatography. The solid line represents the 280 nm absorbance, the dotted line the 420 nm absorbance, the dashed line the NaCl concentration, the black bar the fractions used in subsequent steps, and the triangles represent complex I activity in individual fractions.

samples were prepared at room temperature (Figure 3A). Only after boiling the sample in SDS sample buffer before loading onto the gel was it possible to resolve this high-molecular mass species into individual subunits (Figure 3D). This indicated that the hydrophilic domain is remarkably stable, with a large proportion of protein still forming a complex even after incubation at 80 °C in SDS (Figure 3C). SDS−PAGE analysis of boiled samples revealed eight

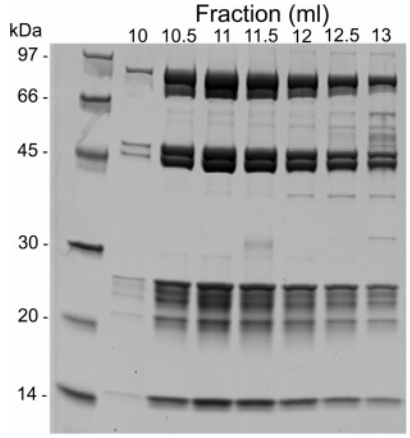

FIGURE 2: SDS−PAGE analysis of Superdex 200 fractions. Protein was stained with Coomassie Blue. Molecular markers are indicated at the left. The top ~80 kDa band is diffused due to relatively large amounts of loaded protein.

Table 1: Purification of the Hydrophilic Domain of Complex I from *T. thermophilus*[a]

purification step	volume (mL)	protein (mg)	total activity[b] (μmol/min)	specific activity (μmol^{-1} mg^{-1})
membrane fraction	160	2480	5300	2.14
solubilized membranes	250	1940	4200	2.16
Q-Sepharose	112	120	2300	19.2
DEAE	135	60	1200	20
UNO-Q12	155	30	1000	25
Superdex 200	5	5	530	106

[a] Starting material was 175 g of *T. thermophilus* cells. [b] Activity was measured by following the oxidation of NADH while FeCy was reduced, at 50 °C.

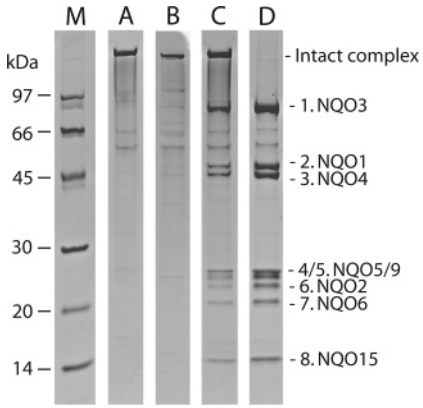

FIGURE 3: Subunit identification and SDS−PAGE analysis of the purified hydrophilic domain of complex I after incubation for 3 min in SDS at different temperatures: (A) room temperature, (B) 50 °C, (C) 80 °C, and (D) 100 °C. Subunits identified by peptide mass fingerprinting and N-terminal sequencing are shown at the right. Lane M contained molecular mass markers with sizes shown at the left. Protein was stained with Coomassie Blue.

polypeptide bands, each of which was analyzed by peptide mass mapping and N-terminal sequencing (Table 2). From the resulting data, seven of these bands were identified as components of the *T. thermophilus* complex I hydrophilic domain. Six of the bands contained single polypeptides, and one band comprised both Nqo5 and Nqo9, which run very close to each other on most gels. Subunit Nqo1 appears to be missing six N-terminal amino acids. Initially, the eighth band defied identification, although we obtained a protein

Table 2: Subunit Identification by Peptide Mass Mapping and N-Terminal Sequencing

band no.	apparent molecular mass (kDa)	subunit (NCBI GI identifier)	peptide mass fingerprint data: no. of peptides matched	peptide mass fingerprint data: sequence coverage (%)	N-terminal sequence: observed	N-terminal sequence: expected
1	85	NQO3 (62297831)	29	46	MVRVKVND	MVRVKVNDRI
2	53	NQO1 (2499313)	20	55	SG- - -(L)- -[b]	MTGPILSGLD
3	49	NQO4 (2499314)	23	69	MRE(E)FLE(E)	MREEFLEEIP
4/5	25	NQO5 (2499323)	13	67	M(R/T)LE(A/R) (V/L)(A/L)(Q/E)	MRLERVLEEA
		NQO9 (2499326)	7	48		MTLKALAQSL
6	23	NQO2 (2499324)	9	61	GF(F)D-KQD	MGFFDDKQDF
7	21	NQO6 (2499331)	10	41	ALKDLFE(R)	MALKDLFERD
8	14	NQO15 (55771878)	8	65[a]	SAS-EREL	MSASSERELY

[a] For this subunit also, two matching peptides, EADFPDFIYR and FFALADR, were identified by tandem MS. [b] For N-terminal sequences, a dash indicates that no data were obtained, brackets indicate that identification of that residue was not definite, and two residues in brackets indicate that either is a possibility.

```
T.ther. ----------------------MSASSERELYEAWVELLSWMREYAQAKGVRFEKEADFP 38
D.geo.  ----------------------MAHAQDERLYTQWVELLGWLEAEASTRGLGFEKVADFP 38
D.rad.  MCAGCPFAAEGTLWLRDVKLPQMANADKAALYRQWVELLGWLGDEARSRGLSIDKVADFP 60
                           *: :..  **  *****.*:   * ::*: ::* ****

T.ther. DFIYRMERPYDLPTTIMTASLSDGLGEPFLLADVSPRHAKLKRIGLRLPRAHIHLHAHYE 98
D.geo.  DYIYRMERPYDLPTTVMSVALTVG-GQPLLLAAVSPRHVDLKGISLRLMGGSKHWHLHAG 97
D.rad.  DYIYRMERPYDLPTTVMSVNVSQD-GQALLVAAVSPRHVDLGGISLRLMGGSKHWHLHAG 119
        *:**************:*:. :: .:*:.:*:* *****..*  *.***  .  * * *
                                                      ↑ ↑ ↑

T.ther. PGKGLVTGKIPLTKERFFALADRAREALAFA-- 129
D.geo.  ER-GLLEGQRPFTRERLAVLLNGALRG--VAAV 127
D.rad.  AG-GLLEGKRPFTRERLSVLLDGVMRGNGTVAV 151
           **: *: *:*:**: .* : . ..   .
```

FIGURE 4: Comparison of the deduced primary structure of subunit Nqo15 from *T. thermophilus* with its homologues. The alignment was performed with the ClustalW server at the European Bioinformatics Institute (*32*). Abbreviations: T.ther., *T. thermophilus* HB8 (YP_143762/55771878); D.geo., *D. geothermalis* (ZP_00396143); D.rad., *D. radiodurans* (NP_294450). Arrows indicate the conserved histidines which contribute to a possible iron binding site (*29*).

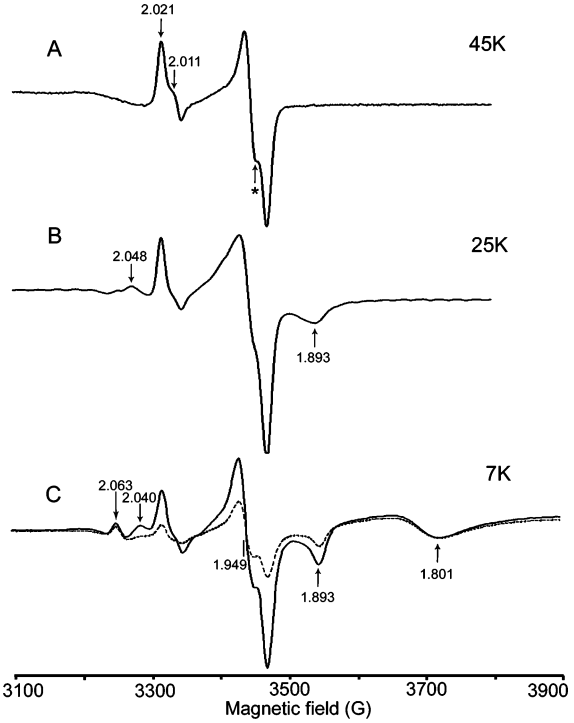

FIGURE 5: EPR spectra of the *T. thermophilus* complex I hydrophilic domain reduced with 2 mM NADH. Spectra were recorded at (A) 45 K with a 0.2 mW microwave power, (B) 25 K with a 20 mW microwave power, and (C) 7 K with a 0.2 mW (—) or 2 mW (- - -) microwave power. Arrows indicate *g* values.

N-terminal sequence and both peptide mass and sequence data. However, following the recent availability of the genome sequences for *T. thermophilus* HB27 (*21*) and HB8 (GenBank entry AP008226; Masui et al., 2004, unpublished), a matching protein was located in the NCBI database. This polypeptide, consisting of 129 residues and with a molecular mass of 14.7 kDa (accession numbers YP_143762/55771878 for HB8 and YP_004097 for HB27), is a hypothetical cytosolic protein with no assigned function. The only homologous proteins present in databases are ORFs encoding polypeptides of unknown function from *Deinococcus radiodurans* (NP_294450, 49% identical) and *Deinococcus geothermalis* (ZP_00396143, 50% identical) (Figure 4). These species are closely related to *T. thermophilus*.

This previously unknown polypeptide copurified with the complex I hydrophilic domain through membrane preparation and solubilization, four column chromatography steps, and various dilution, concentration, and detergent-exchange processes. In SDS−PAGE analysis, it was only located as a discrete band after boiling purified protein. It is clearly very strongly associated with the hydrophilic domain and can thus be identified as a 15th subunit of *T. thermophilus* complex I, which we propose to name Nqo15.

EPR Spectroscopy. EPR spectroscopic analysis of the complex I hydrophilic domain, reduced by NADH, was performed to analyze iron−sulfur clusters. At 45 K, binuclear type iron−sulfur clusters are predominant and we could detect one major axial or near-axial EPR signal for which $g_z = 2.021$ and $g_y = 1.940$ (Figure 5A). In addition, a possible contribution from a second binuclear cluster could be detected as a main peak shoulder at $g_z = 2.011$. At 25 K, additional signals, most likely due to tetranuclear clusters, could be detected at $g_z = 2.048$ and $g_x = 1.893$, while the apparent g_y value shifted to ∼1.943 (Figure 5B). The magnitude of the $g_z = 2.048$ signal was diminished at lower temperatures and at 7 K it was replaced by a $g_z = 2.040$ signal (Figure 5C). Also, another signal for which $g_z = 2.063$ and $g_x = 1.801$ appeared at low temperatures, presumably from a rapidly relaxing tetranuclear cluster (Figure 5C). These signals are similar to the EPR spectra observed previously with *T. thermophilus* membrane particles and partially purified enzyme (*22*). This suggests that the iron−sulfur clusters in our preparation were not damaged during the purification procedure.

The clusters can be tentatively assigned by comparison with previous EPR data obtained with intact enzymes and overexpressed subunits. For binuclear cluster N1b, located in subunit Nqo3, an axial or near-axial spectrum was

observed in different species, in which $g_z = 2.03/2.02$ and $g_y = 1.94$ (*8*, *23*−*25*). Therefore, cluster N1b is likely to be a main contributor to the high-temperature spectrum shown in Figure 5A, where $g_z = 2.021$ and $g_y = 1.940$. Overexpressed and purified *T. thermophilus* Nqo2 subunit, containing the single binuclear cluster N1a, has been shown to exhibit a rhombic spectrum where $g_z = 2.002$, $g_y = 1.946$, and $g_x = 1.915$ (*17*). These values could be somewhat different in situ, and it is possible that the observed shoulder at $g_z = 2.011$ (Figure 5A) is due to cluster N1a. Another shoulder in the g_y region, indicated with an asterisk in Figure 5A, may also be due to cluster N1a, although it is possible that cluster N1b has a partly rhombic spectrum. Thus, we can assign the main observed binuclear cluster as N1b, with a possible additional contribution from cluster N1a. Under our experimental conditions (in the presence of NADH), cluster N1a may be only partially reduced, as it has a low midpoint potential, approximately −370 mV (*25*, *26*). However, the reduction of the subcomplex with 10 mM sodium dithionite, in addition to NADH, did not reveal any additional or increased EPR signals at 45 K (data not shown).

Of the tetranuclear clusters of complex I, neither N6a/b nor N7 has been detected in situ previously (*1*), and N5 signals have been very weak (*24*). Thus, these clusters are not likely to be detectable by EPR in our preparation either. Cluster N2 has not been observed previously in *T. thermophilus*, but its spectra are axial in other species, where $g_z = 2.05$ and $g_y = 1.92/1.91$ (*23*−*25*). Other tetranuclear clusters have rhombic spectra (*25*). Overall, spin relaxation rates of complex I clusters increase (and optimal EPR sample temperature decreases) with the cluster identification number, from N1 to N5 (*25*). Therefore, it is likely that the $g_z = 2.048$ signal, which appears at 25 K (Figure 5B) and disappears at 7 K (Figure 5C), is due to cluster N2. Its g_y value may be close to 1.93/1.94 and so is not discernible in the spectra. The $g_x = 1.893$ signal, visible at 7 and 25 K, may be due to the cluster which $g_z = 2.040$ signal dominates at 7 K and may also be present at 25 K, overlapping in the g_z area with the signal from cluster N2. From comparison of *g* values and spin relaxation rates (*25*), cluster N3 is the most likely candidate responsible for the observed rhombic signal at $g_z = 2.040$ and $g_x = 1.893$. Most of the g_y values in our spectra appear to coincide in approximately the 1.94 region. However, a comparison of the 7 K spectra obtained at 2 mW versus 0.2 mW (Figure 5C) indicates that the relative decrease in the magnitudes of the $g_z = 2.040$ and $g_x = 1.893$ signals as compared to $g_z = 2.063$ and $g_x = 1.801$ signals is associated with an increased contribution of the signal for which $g_y = 1.949$. Thus, the rapidly relaxing tetranuclear cluster N4 is likely to be responsible for these 7 K signals with $g_z = 2.063$, $g_y = 1.949$, and $g_x = 1.801$.

Crystallization. Purified protein was monodisperse and highly stable, two favorable characteristics for crystal growth. Tests using gel filtration chromatography have shown that the subcomplex is stable in a range of detergents, including OG. Therefore, after purification in DDM, protein was exchanged at the last step into OG for crystal trials. OG was chosen for its small micelle size, increasing the chance of crystal contacts forming between molecules, compared to large micelle, low-CMC detergents such as DDM. The final concentration step proved to be vital, because an overconcentrated protein solution produced mainly detergent crystals or precipitate, while an underconcentrated protein solution gave only phase separation in the crystal trials. Thus, for crystal formation, an exact concentration of both protein and detergent was important. The best crystals appear to be produced when detergent is close to, but not reaching, the phase separation point, as noted previously (*27*).

Crystals obtained in extensive initial screens, using commercial reagents and a robotic setup (1344 conditions), were thin, small, and highly aggregated (0.2 mm × 0.1 mm × 0.02 mm). Subsequent optimization based around initial conditions produced larger single crystals, though still of plate morphology, only 0.02−0.05 mm thick (Figure 6). Crystals were a pale brown color due to iron from the many iron−sulfur clusters present in the enzyme. The presence of $MgCl_2$ or $CaCl_2$ (∼0.1 M) was vital for the improvement of crystals, in conjunction with a high NaCl concentration (0.3−0.6 M). The use of additives (e.g., heptanetriol, benzamidine, or ethanol) did not alter the crystal morphology. Thus, at the in-house X-ray source, these thin plate crystals diffracted to only 6−8 Å. At the synchrotron, however, some crystals diffracted to a maximum of 3.2 Å (*4*). SDS−PAGE analysis of purified crystals indicated that all the subunits were present in the crystals, including, importantly, the newly identified Nqo15 (Figure 7).

DISCUSSION

Our purification method yields relatively large amounts of the monodisperse hydrophilic domain of complex I, with no significant impurities, suitable for structural studies. The preparation shows expected activity with different electron acceptors, and EPR signals (Figure 5) are similar to those observed previously with intact *T. thermophilus* membranes (*22*), indicating that the overall structure of the domain is not perturbed upon separation from the membrane arm and during purification. We have tentatively identified most of the clusters which can be observed by EPR in an intact complex I: tetranuclear N2−N4, as well as binuclear N1b and possibly N1a.

The membrane domain is lost upon solubilization, even in mild detergents, possibly due to the specific lipid requirements of the complex. As we observed previously for *E. coli* complex I, native lipids can be important for maintaining the stability of the enzyme (*12*). Alternatively, the relatively low temperature during purification (4 °C vs 70 °C, the temperature at which cells are grown) would have weakened hydrophobic interactions, and that might have destabilized the intact complex. The hydrophilic domain on its own is stable in SDS up to 80 °C (Figure 3). This is similar to observations with a smaller subcomplex of complex I (NuoEFG) from the hyperthermophile *Aquifex aeolicus*, which is also stable at 80 °C (*28*). While the purification procedure described for *A. aeolicus* results in an intact complex, the yield is ∼100 times lower than that from the procedure described here.

N-Terminal sequencing and peptide mass mapping confirmed that the subunit composition of the preparation was as expected for the hydrophilic domain of complex I, but with the addition of a previously unidentified subunit, Nqo15, clearly associated with the complex. SDS−PAGE analysis indicated that crystals, used for X-ray data collection,

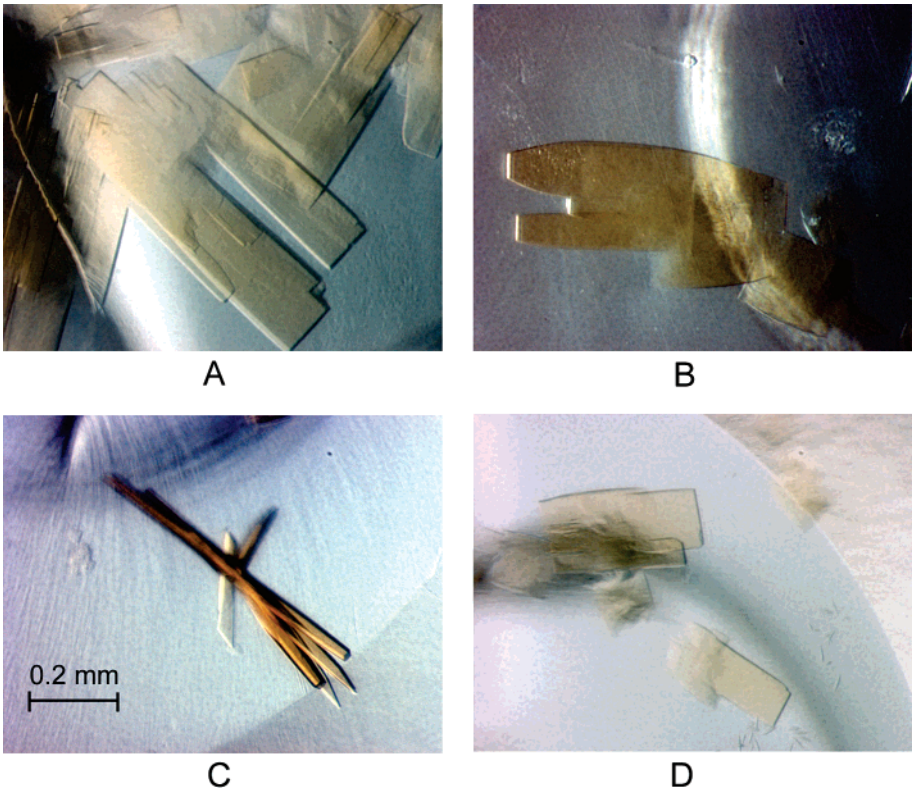

FIGURE 6: Crystals of the hydrophilic domain of *T. thermophilus* complex I. Complex I (1 μL, 15 mg/mL) was mixed with 1 μL of crystallization reagent: (A) 0.1 M HEPES, 0.45 M NaCl, 10% PEG 4000 (w/v), and 0.1 M $MgCl_2$, (B) 0.1 M HEPES, 0.5 M NaCl, 9% PEG 4000 (w/v), and 0.1 M $CaCl_2$, (C) 0.1 M HEPES, 0.5 M NaCl, 10% PEG 4000 (w/v), and 0.1 M $MgCl_2$, and (D) 0.1 M HEPES, 0.5 M NaCl, 10% PEG 4000 (w/v), and 0.15 M $MgCl_2$. All trials were performed in Cryschem 24-well sitting drop plates at room temperature. The pale brown color of the crystals, due to the iron−sulfur clusters, is clearly visible. Photographs were taken 5−7 days after setup and at the same magnification.

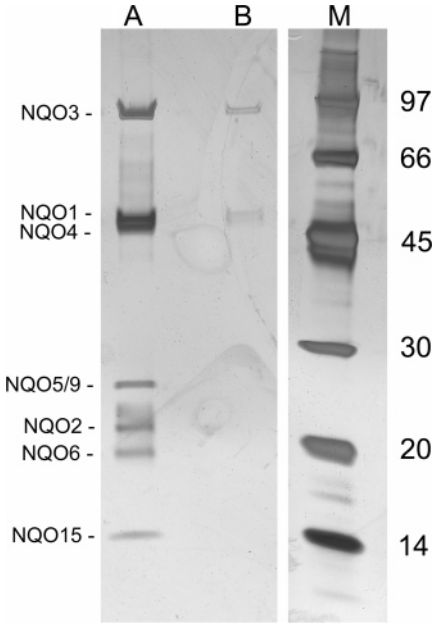

FIGURE 7: SDS−PAGE analysis of crystals of the hydrophilic domain of *T. thermophilus* complex I: (A) crystals washed three times in crystallization reagent and (B) crystal loop that has been immersed into a crystal well containing crystals with crystallization reagent but no crystal collected, and then washed in a manner similar to that used for panel A. Subunits are indicated at the left. Lane M contained molecular mass markers with sizes shown at the right. Protein was silver stained.

contain all subunits identified in the preparation, including the novel subunit Nqo15 (Figure 7). This provided further evidence that it is a bona fide subunit of *T. thermophilus* complex I. Finally, the atomic structure of the domain, which we recently determined, shows that Nqo15 is an integral part of the complex (*29*).

Generally, the 14 genes encoding bacterial complex I subunits are organized in one operon or gene cluster. The order of these *nuoA*−*N* or *nqo1*−*14* genes within the cluster is usually conserved (*6*). From bacterial sources, so far only *E. coli* complex I has been purified in an intact form (*12*, *30*) with all protein components positively identified (*12*). Only the 14 subunits (with NuoC and NuoD fused) encoded within the *nuo* operon were detected in the preparation (*12*). Complex I from *A. aeolicus* has also been purified in an intact form, and seven known *nuo* subunits have been identified so far in the purified sample (*28*). The *nuo*/*nqo* gene cluster in some bacteria contains additional ORFs, not encoding known complex I subunits. However, disruption of several of these ORFs in *Rhodobacter capsulatus* did not affect complex I activity (*31*). Therefore, it is generally assumed that bacterial complex I contains only 14 core subunits, which are conserved from bacteria to mammals (*1*, *6*). That seems to be the case for the *E. coli* enzyme (*12*). However, our data clearly show that at least in *T. thermophilus* and its relatives, complex I contains an additional subunit, Nqo15. It is encoded in a locus separated from the *nqo* operon by ~ 360 kb. Our finding opens up a possibility that, in some other bacteria, complex I may also contain more than 14 subunits.

Since a protein homologous to Nqo15 could be identified so far only in close relatives of *T. thermophilus* (Figure 4),

it is possible that one role of this subunit is to stabilize the complex, due to the harsh conditions in which these organisms grow. The atomic structure is consistent with such a proposal (*29*). In addition, the fold of Nqo15 was found to be similar to the unique fold of frataxins (iron chaperones), with several conserved Nqo15 histidines (Figure 4) contributing to the possible iron binding site (*29*). Therefore, another role of Nqo15 may be in the storage of iron for Fe−S cluster regeneration. Thermophiles which are less closely related to *Thermus*, or mesophilic organisms, may not have a similar subunit. Alternatively, it is possible that its structural and functional analogue is present in other species but that the degree of sequence similarity is low. Also, many bacteria have extended versions of Nqo3 or Nqo2 subunits in comparison to *T. thermophilus*, and these extra domains might play a role similar to that of Nqo15.

ACKNOWLEDGMENT

We thank Dr. Sew Peak-Chew and Dr. Farida Begum (MRC LMB) for performing N-terminal sequencing. In addition, we thank Dr. Judy Hirst (MRC Dunn) for help with EPR spectroscopy.

REFERENCES

1. Yagi, T., and Matsuno-Yagi, A. (2003) The proton-translocating NADH-quinone oxidoreductase in the respiratory chain: The secret unlocked, *Biochemistry 42*, 2266−2274.
2. Walker, J. E. (1992) The NADH:ubiquinone oxidoreductase (complex I) of respiratory chains, *Q. Rev. Biophys. 25*, 253−324.
3. Friedrich, T. (2001) Complex I: A chimaera of a redox and conformation-driven proton pump? *J. Bioenerg. Biomembr. 33*, 169−177.
4. Hinchliffe, P., and Sazanov, L. A. (2005) Organization of iron−sulfur clusters in respiratory complex I, *Science 309*, 771−774.
5. Carroll, J., Fearnley, I. M., Shannon, R. J., Hirst, J., and Walker, J. E. (2003) Analysis of the subunit composition of complex I from bovine heart mitochondria, *Mol. Cell. Proteomics 2*, 117−126.
6. Friedrich, T., and Scheide, D. (2000) The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases, *FEBS Lett. 479*, 1−5.
7. Ohnishi, T., Sled, V. D., Yano, T., Yagi, T., Burbaev, D. S., and Vinogradov, A. D. (1998) Structure−function studies of iron−sulfur clusters and semiquinones in the NADH-Q oxidoreductase segment of the respiratory chain, *Biochim. Biophys. Acta 1365*, 301−308.
8. Nakamaru-Ogiso, E., Yano, T., Ohnishi, T., and Yagi, T. (2002) Characterization of the iron−sulfur cluster coordinated by a cysteine cluster motif (CXXCXXXCX27C) in the Nqo3 subunit in the proton-translocating NADH-quinone oxidoreductase (NDH-1) of *Thermus thermophilus* HB-8, *J. Biol. Chem. 277*, 1680−1688.
9. Rasmussen, T., Scheide, D., Brors, B., Kintscher, L., Weiss, H., and Friedrich, T. (2001) Identification of two tetranuclear FeS clusters on the ferredoxin-type subunit of NADH:ubiquinone oxidoreductase (complex I), *Biochemistry 40*, 6124−6131.
10. Yano, T., Magnitsky, S., Sled, V. D., Ohnishi, T., and Yagi, T. (1999) Characterization of the putative 2x[4Fe-4S]-binding NQO9 subunit of the proton-translocating NADH-quinone oxidoreductase (NDH-1) of *Paracoccus denitrificans*. Expression, reconstitution, and EPR characterization, *J. Biol. Chem. 274*, 28598−28605.
11. Bottcher, B., Scheide, D., Hesterberg, M., Nagel-Steger, L., and Friedrich, T. (2002) A novel, enzymatically active conformation of the *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I), *J. Biol. Chem. 277*, 17970−17977.
12. Sazanov, L. A., Carroll, J., Holt, P., Toime, L., and Fearnley, I. M. (2003) A role for native lipids in the stabilization and two-dimensional crystallization of the *Escherichia coli* NADH-ubiquinone oxidoreductase (complex I), *J. Biol. Chem. 278*, 19483−19491.
13. Oshima, T., and Imahori, K. (1971) Isolation of an extreme thermophile and thermostability of its transfer ribonucleic acid and ribosomes, *J. Gen. Appl. Microbiol. 17*, 513−517.
14. Yagi, T., Hon-nami, K., and Ohnishi, T. (1988) Purification and characterization of two types of NADH-quinone reductase from *Thermus thermophilus* HB-8, *Biochemistry 27*, 2008−2013.
15. Burch, H. B. (1957) Fluorometric assay of FAD, FMN and riboflavin, *Methods Enzymol. 3*, 960−962.
16. Huberman, A., and Perez, C. (2002) Nonheme iron determination, *Anal. Biochem. 307*, 375−378.
17. Yano, T., Chu, S. S., Sled, V. D., Ohnishi, T., and Yagi, T. (1997) The proton-translocating NADH-quinone oxidoreductase (NDH-1) of thermophilic bacterium *Thermus thermophilus* HB-8. Complete DNA sequence of the gene cluster and thermostable properties of the expressed NQO2 subunit, *J. Biol. Chem. 272*, 4201−4211.
18. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, 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.
19. Mann, M., and Wilm, M. (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags, *Anal. Chem. 66*, 4390−4399.
20. Albracht, S. P., van der Linden, E., and Faber, B. W. (2003) Quantitative amino acid analysis of bovine NADH:ubiquinone oxidoreductase (Complex I) and related enzymes. Consequences for the number of prosthetic groups, *Biochim. Biophys. Acta 1557*, 41−49.
21. Henne, A., Bruggemann, H., Raasch, C., Wiezer, A., Hartsch, T., Liesegang, H., Johann, A., Lienard, T., Gohl, O., Martinez-Arias, R., Jacobi, C., Starkuviene, V., Schlenczeck, S., Dencker, S., Huber, R., Klenk, H. P., Kramer, W., Merkl, R., Gottschalk, G., and Fritz, H. J. (2004) The genome sequence of the extreme thermophile *Thermus thermophilus*, *Nat. Biotechnol. 22*, 547−553.
22. Meinhardt, S. W., Wang, D. C., Hon-nami, K., Yagi, T., Oshima, T., and Ohnishi, T. (1990) Studies on the NADH-menaquinone oxidoreductase segment of the respiratory chain in *Thermus thermophilus* HB-8, *J. Biol. Chem. 265*, 1360−1368.
23. Flemming, D., Schlitt, A., Spehr, V., Bischof, T., and Friedrich, T. (2003) Iron−sulfur cluster N2 of the *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I) is located on subunit NuoB, *J. Biol. Chem. 278*, 47602−47609.
24. Yano, T., Sklar, J., Nakamaru-Ogiso, E., Takahashi, Y., Yagi, T., and Ohnishi, T. (2003) Characterization of cluster N5 as a fast-relaxing [4Fe-4S] cluster in the Nqo3 subunit of the proton-translocating NADH-ubiquinone oxidoreductase from *Paracoccus denitrificans*, *J. Biol. Chem. 278*, 15514−15522.
25. Ohnishi, T. (1998) Iron−sulfur clusters/semiquinones in complex I, *Biochim. Biophys. Acta 1364*, 186−206.
26. Zu, Y., Di Bernardo, S., Yagi, T., and Hirst, J. (2002) Redox properties of the [2Fe-2S] center in the 24 kDa (NQO2) subunit of NADH:ubiquinone oxidoreductase (complex I), *Biochemistry 41*, 10056−10069.
27. Kuhlbrandt, W. (1988) Three-dimensional crystallization of membrane proteins, *Q. Rev. Biophys. 21*, 429−477.
28. Peng, G., Fritzsch, G., Zickermann, V., Schagger, H., Mentele, R., Lottspeich, F., Bostina, M., Radermacher, M., Huber, R., Stetter, K. O., and Michel, H. (2003) Isolation, characterization and electron microscopic single particle analysis of the NADH:ubiquinone oxidoreductase (complex I) from the hyperthermophilic eubacterium *Aquifex aeolicus*, *Biochemistry 42*, 3032−3039.
29. Sazanov, L. A., and Hinchliffe, P. (2006) Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*, *Science* (in press).
30. Leif, H., Sled, V. D., Ohnishi, T., Weiss, H., and Friedrich, T. (1995) Isolation and characterization of the proton-translocating NADH:ubiquinone oxidoreductase from *Escherichia coli*, *Eur. J. Biochem. 230*, 538−548.
31. Dupuis, A., Chevallet, M., Darrouzet, E., Duborjal, H., Lunardi, J., and Issartel, J. P. (1998) The complex I from *Rhodobacter capsulatus*, *Biochim. Biophys. Acta 1364*, 147−165.
32. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) Multiple sequence alignment with the Clustal series of programs, *Nucleic Acids Res. 31*, 3497−3500.

BI0600998