Supplement:

Methods

Cultivation of *T. elongatus* WT strain and preparation / solubilization of thylakoid membranes

T. elongatus WT strain was grown in BG-11 medium (Rippka et al. 1979) at 45°C under illumination of increasing intensity of 50-200 µmol of photons m^{-2} s^{-1} in a 20L foil fermenter at 5 % CO_2 and then transferred to air level CO_2. The culture was harvested and the thylakoid membranes were prepared as described in (Kuhl et al. 2000). The thylakoid membranes were suspended in solubilisation buffer (20 mM MES pH 6.0; 25 mM $MgCl_2$; 1 % (w/v) β-DM; 1 mM TLCK; 1 mM Pefabloc; 5 mM Na-ascorbate) at a final chlorophyll concentration of 1 mg/ml and incubated for 45 min at 20°C and gentle agitation. An equal volume of dilution buffer (20 mM MES pH 6,0; 1 mM TLCK; 1 mM Pefabloc; 5 mM Na-ascorbate) was added, to get a final concentration of 0.5 % β-DM. Insoluble material was removed by centrifugation at 45.000 g.

Purification of NDH-1L by liquid chromatography

The NDH-1L-complex was purified by Ni^{2+} affinity chromatography, followed by size exclusion chromatography. The solubilized thylakoids were filtered through a 0.45 µm membrane and then applied to a chelating sepharose fast flow column (GE healthcare, 7.854 ml) with a flow rate of 1 ml/min, previously saturated with 0.1 M $NiSO_4$ and equilibrated with equilibration buffer (20 mM MES pH 6.0; 0.5 M Mannitol; 0.03 % (w/v) β-DM; 150 mM NaCl; 5 mM Na-ascorbate). The column was washed with equilibration buffer containing 5 mM histidine and the proteins were eluted with a 10-100 mM histidine gradient in equilibration buffer. After concentration of the fractions containing the complex (Amicon concentrator, 15 ml, CutOff 100 kDa), the sample was applied to a Superose 6 size exclusion chromatography column (GE healthcare), previously equilibrated with SEC-buffer (20 mM MES pH 6.0; 0.5 M Mannitol; 0.03 % (w/v) β-DM; 150 mM NaCl; 5 mM Na-ascorbate). After isocratic elution with a flow rate of 0.15 ml/min and a following concentration step, the purified samples were stored at -80 °C.

Electrophoresis

5 µg of purified proteins were loaded on a polyacrylamide gradient BN-Gel (3.5 - 16 %) in a Mini Protean 3 electrophoresis unit (BioRad). The electrophoresis was run as described in (Battchikova et al. 2005). Afterwards, the BN-gel lane was cut out, incubated and loaded on to a 14 % SDS-PAGE Gel with 6 M Urea. The proteins were visualized by silver staining (Blum et al. 1987).

Mass spectrometry analysis of intact proteins

The masses of intact small proteins of the NDH-1 complex were determined by MALDI-ToF mass spectrometry according to (El-Mohsnawy et al. 2010).

The masses of the polypeptides of NDH-1 complex from the Cramer laboratory were measured by liquid chromatography with electrospray-ionization mass spectrometry according to (Whitelegge 2002).

Sample preparation for protein identification via 1D-nLC-ESI-MS/MS

Analysis of 2D-gel spots: Excised protein spots from 2D-Gels were destained and the corresponding proteins were digested *in-gel* with trypsin and/or chymotrypsin (Shevchenko et al. 1996).

Analysis of isolated NDH-1 complexes in solution: The concentrated sample was diluted with 60% methanol, 40% 25 mm ammonium bicarbonate buffer and trypsin and/or chymotrypsin (each 1:100, w/w) were added to the sample. The proteolysis was performed overnight at 37 °C.

Protein identification via 1D-nLC-ESI-MS/MS

After desalting by ZipTips (Millipore) the samples were resuspended in buffer A (0.1 % formic acid in water) and subjected to 1D-nLC-ESI-MS/MS using an autosampler. An UPLC BEH C_{18} column (1.7 µm, 75 µm x 150 mm, Waters, Milford, MA, USA) and an UPLC Symmetry C_{18} trapping column (5 µm, 180 µm x 20 mm, Waters, Milford, MA, USA) for LC as well as a PicoTip Emitter (SilicaTip, 30 µm, New Objective, Woburn, MA, USA) were used in combination with the nanoACQUITY gradient UPLC pump system (Waters, Milford, MA, USA) coupled to an LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). The analytical column oven was set to 45 °C. For elution of the peptides a multiple step gradient of buffer A to buffer B (0.1 % formic acid in acetonitrile) was applied (0-5 min: 1 %

buffer B; 5-10 min: 5 % buffer B; 10-175 min: 40 % buffer B; 175-200 min: 99 % buffer B; 200-210 min: 1 % buffer B) at a flow rate of 0.4 µL/min and a spray voltage of 1.5-1.8 kV. The LTQ Orbitrap was operated by instrument method files of Xcalibur (Rev. 2.0.7). The linear ion trap and orbitrap were operated in parallel, i.e. during a full MS scan on the orbitrap in the range of 300-2000 m/z at a resolution of 60,000. MS/MS spectra of the four most intense precursors were detected in the ion trap. The heated desolvation capillary was set to 200 °C. The relative collision energy for collision-induced dissociation was set to 35 %. Dynamic exclusion was enabled with a repeat count of one and a one-minute exclusion duration window. Singly charged and more than triply charged ions were rejected from MS/MS.

SEQUEST Analysis

The SEQUEST algorithm was used for MS/MS data interpretation. To obtain reliable protein identification, only peptides with a ΔCn score above 0.1 were used. In addition the cross-correlation scores of double and triple charged peptides had to be greater than 2.5 and 3.5, respectively. As modifications the oxidation of methionine was permitted.

Table S1: NDH-1 subunit analysis after *in-gel* digestion with trypsin

NDH-1 SU	ORF[a]	kDa[b]	TMH[c]	XC[d]	Coverage (%)[e]
NdhF1	tll0720	71.972	16	30.1	1.37
NdhH	tlr1288	45.216	--	120.3	33.76
NdhD1	tll0719	56.078	12	68.3	9.92
NdhB	tll0045	55.144	14	30.2	6.6
NdhA	tlr0667	41.347	13	60.2	16.09
NdhK	tlr0705	25.742	--	70.2	24.47
NdhI	tlr0668	22.415	--	40.2	20.41
NdhG	tlr0669	21.569	5	30.3	13.00
NdhJ	tlr1430	19.343	--	10.2	8.33
NdhN	tlr1130	16.636	--	10.1	7.33
NdhM	tll0447	12.567	--	40.3	45.05
NdhO	tsl0017	7.867	--	10.2	18.57

[a] Cyanobase ORF ID.
[b] Calculated molecular weight.
[c] Number of predicted transmembrane helices.
[d] XC score: $((\Delta Cn^2)+Sp) \cdot Xcorr$
[e] Protein coverage given by percentage of identified amino acids.

Table S2: NDH-1 subunit analysis after digestion with trypsin in solution

NDH-1 SU	ORF[a]	kDa[b]	TMH[c]	XC[d]	Coverage (%)[e]
NdhA	tlr0667	41.347	13	50.3	15.04
NdhB	tll0045	55.144	14	40.2	12.43
NdhC	tlr1429	15.003	3	10.3	15.91
NdhD1	tll0719	56.078	12	50.2	17.32
NdhE*	tlr0670	11.133	3	10.3	28.71
NdhF1	tll0720	71.972	16	40.3	9.76
NdhG	tlr0669	21.569	5	40.3	21.50
NdhH	tlr1288	45.216	--	110.3	34.01
NdhI	tlr0668	22.415	--	90.3	45.41
NdhJ	tlr1430	19.343	--	40.2	32.14
NdhK	tlr0705	25.742	--	50.2	21.52
NdhL	tsr0706	8.571	2	10.1	11.84
NdhM	tll0447	12.567	--	60.2	40.54
NdhN	tlr1130	16.636	--	50.3	46.00
NdhO	tsl0017	7.867	--	20.2	35.71

[a] Cyanobase ORF ID.
[b] Calculated molecular weight.
[c] Number of predicted transmembrane helices.
[d] XC score: $((\Delta Cn^2)+Sp) \cdot Xcorr$
[e] Protein coverage given by percentage of identified amino acids.
* Identified after cleavage with trypsin and chymotrypsin.

Table S3: Identification of NdhP and NdhQ by specific peptides

	MH+[a]	ΔM (ppm)[b]	P[c]	z[d]	Xcorr[e]	ΔCn[f]	Coverage[g] (%)
NdhP[h]	1782.710	4.8	$2.3e^{-11}$	2	5.145	-	36.36
NdhQ[i]	1299.651	0.9	$5.3e^{-5}$	2	2.744	0.575	24.44

[a] Measured masses of the precursor ion.
[b] Mass difference of calculated and measured masses in ppm.
[c] Peptide probability calculated by the Bioworks software.
[d] Charge of the peptide.
[e] Cross-correlation score calculated by the sequest algorithm.
[f] ΔCn value calculated by the sequest algorithm.
[g] Protein coverage given by percentage of identified amino acids.
[h] After digestion with trypsin.
[i] After digestion with chymotrypsin.

Table S4: Masses of intact NdhP and NdhQ

	Calculated average mass[a]	Measured.average mass[b]	ΔM (Da)[c]	TMH[d]	Modification[e]
NdhP	4902.67	4902.65	0.02	1	N-Formyl
NdhQ	4710.53	4710.52	0.01	1	Minus Met-1

[a] Calculated average masses (MH+).
[b] Measured average masses (MH+).
[c] Mass difference of calculated and measured masses.
[d] Number of predicted transmembrane helices.
[e] Predicted common modifications based on the measured masses.

Gene and translated amino acid sequences of the novel NDH-1 subunits

ndhP

Sequence:

ATGGATGCTGTGATTAGCGTAAAGCCCATTTTGCTGGCTATGACGCCTGTATTTA
TTCTGTTGTGTTTGTTTTTTGGCACCCGCAATGGCTTCTACGACACGGATCAATA
CCACGGTAACGGTTCTGCCCAC

Genomic region: 1189596-1189465

Translated amino acid sequence:

MDAVISVKPILLAMTPVFILLCLFFGTRNGFYDTDQYHGNGSAH

ndhQ

Sequence:

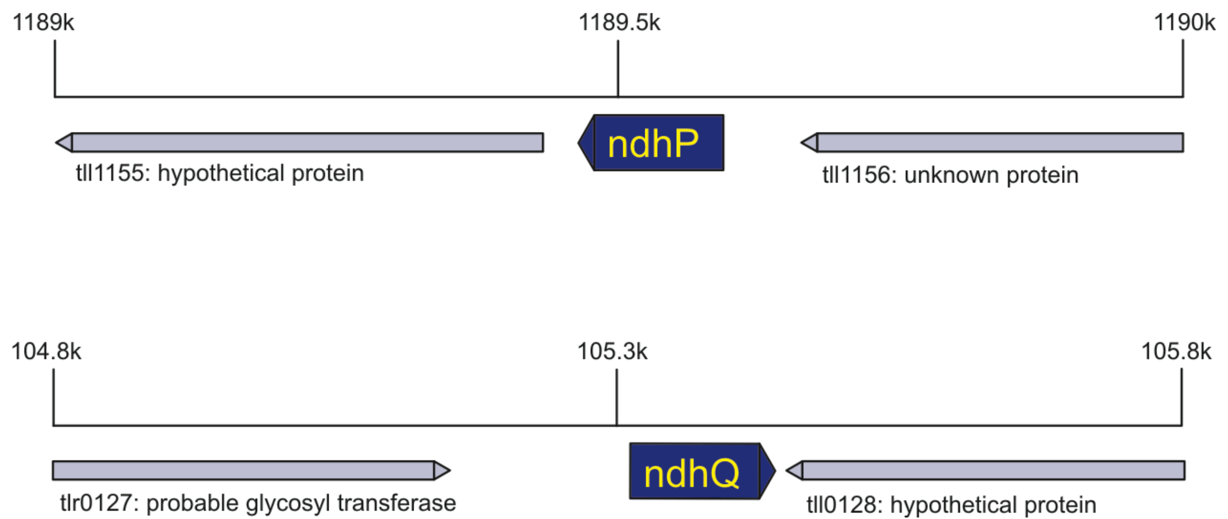
ATGGCCACGGATTTTAATCGCGGCATTATGAAGTTTGATGGTGCCGACAGCCCG
GCGATGATTGCGATTTCTGCGGTCTTGATTCTTGGCTTTATTGCAGGACTGATTT
GGTGGGCACTCCACACCGCTTACGCC

Genomic region: 105314-105448

Translated amino acid sequence:

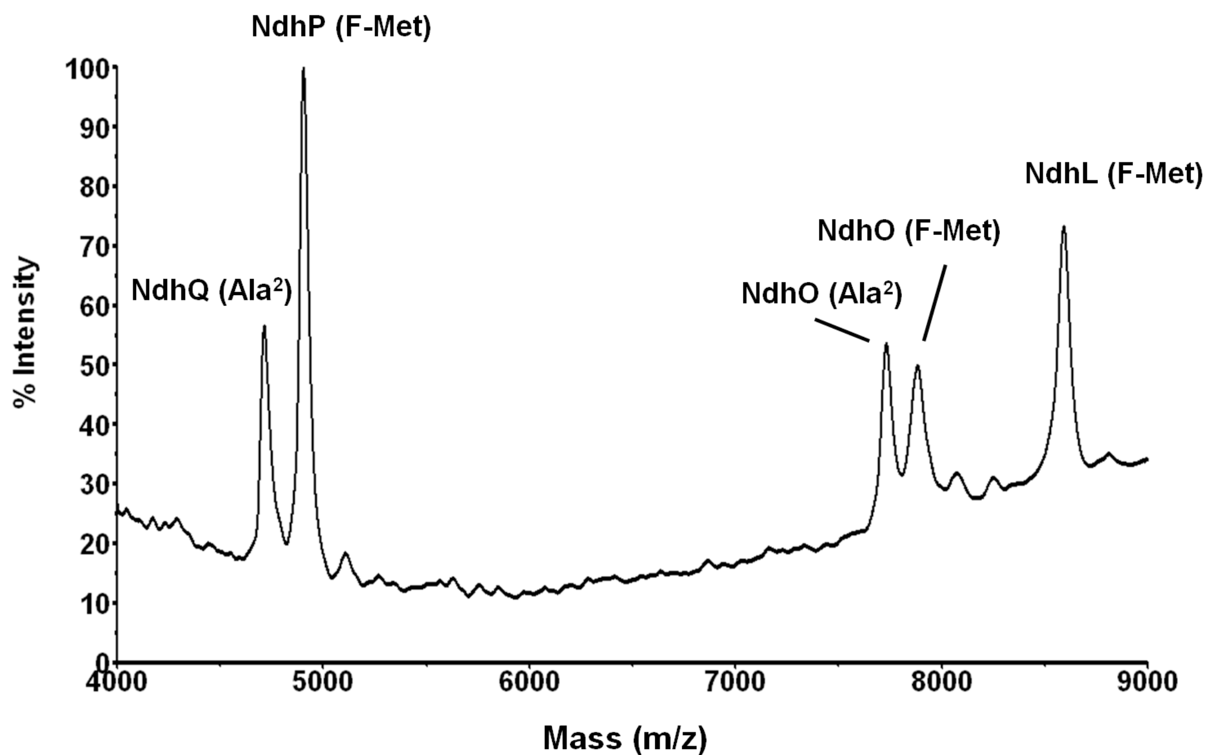
MATDFNRGIMKFDGADSPAMIAISAVLILGFIAGLIWWALHTAYA

Figure S1



Position of the *ndhP* and *ndhQ* genes within the genome of *T. elongatus*. Open reading frames were predicted with Glimmer (http://bioinformatics.biol.rug.nl/websoftware/orf/orf_start.php) and included into the genome of *T. elongatus* as annotated in Cyanobase (http://genome.kazusa.or.jp/cyanobase).

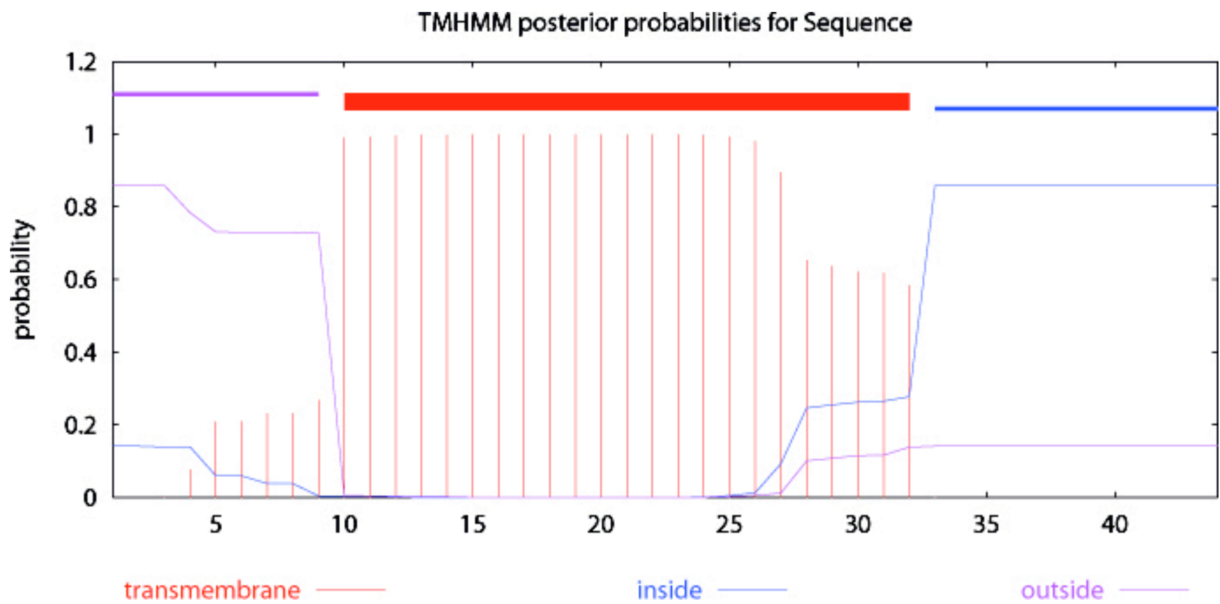
Figure S2



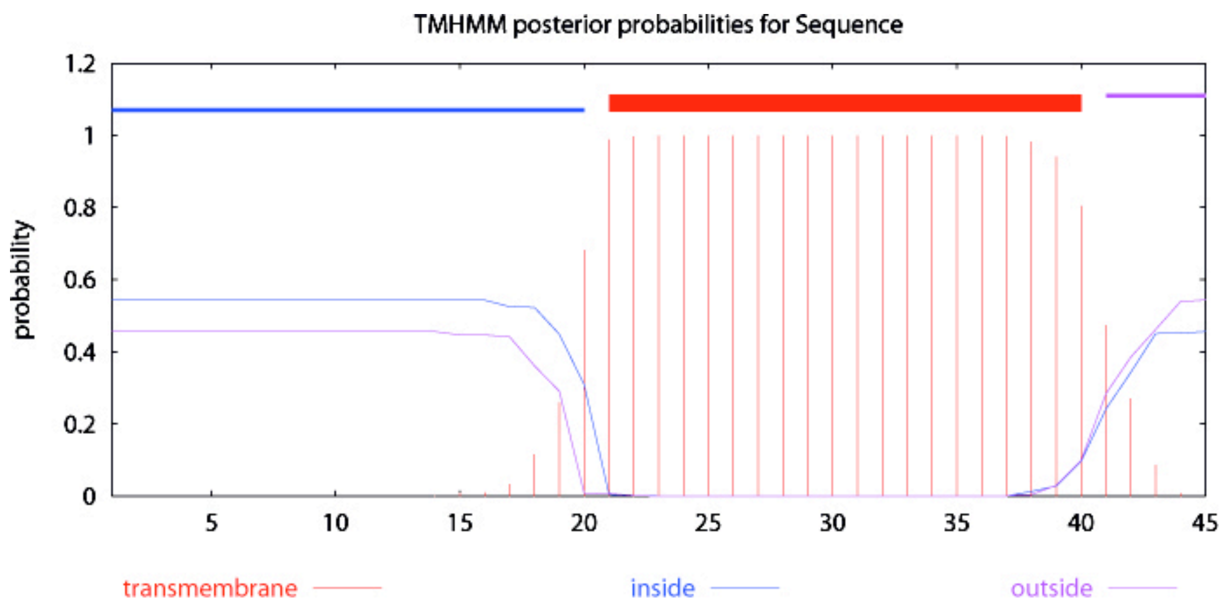
Result of MALDI-ToF mass spectrometry of intact NDH-1 complexes. Samples ware analyzed in the low molecular mass range according to (El-Mohsnawy et al. 2010) and the measured masses were assigned to small NDH-1 subunits with respect to common posttranslational modifications.

Figure S3

NdhP



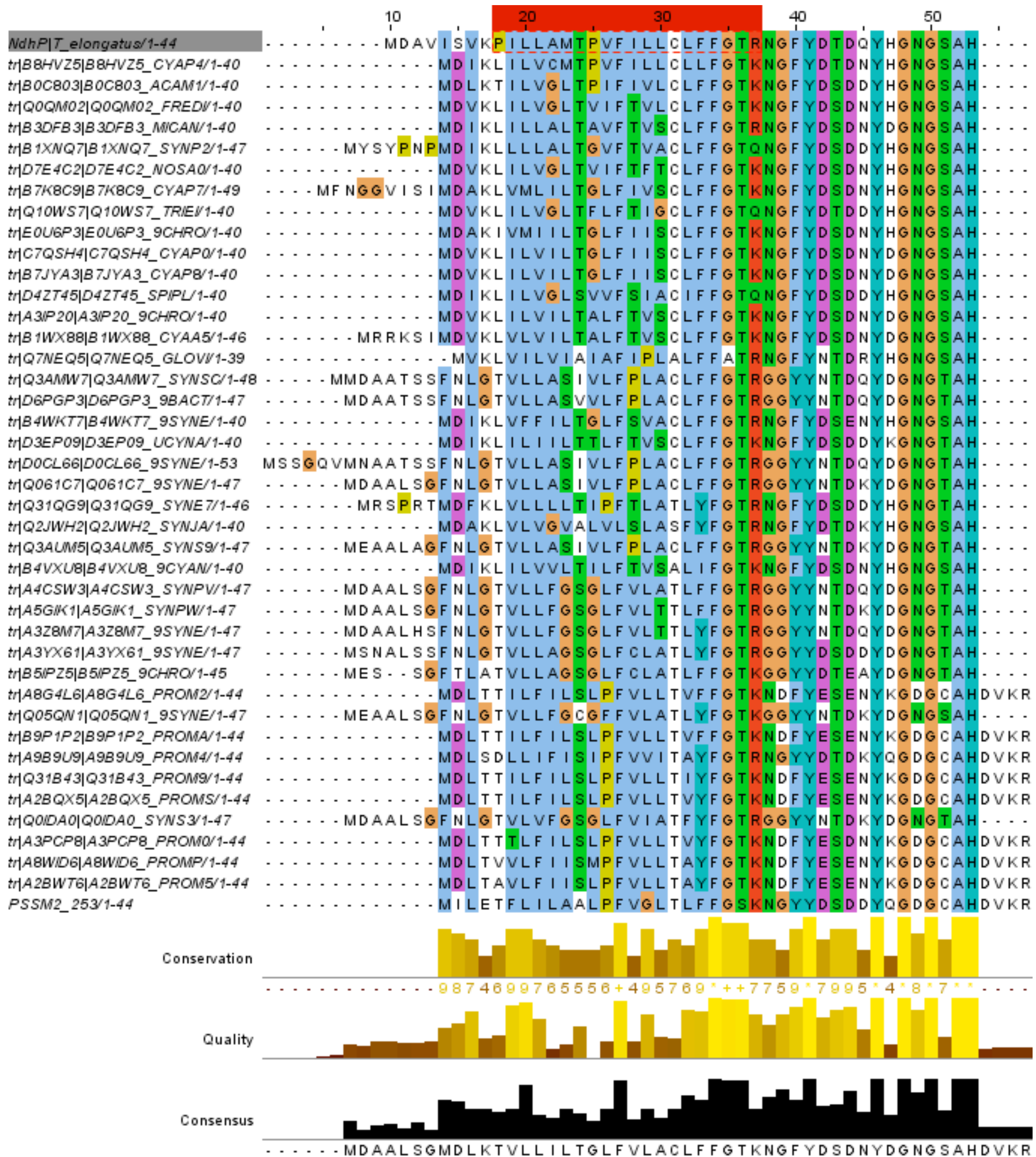
NdhQ



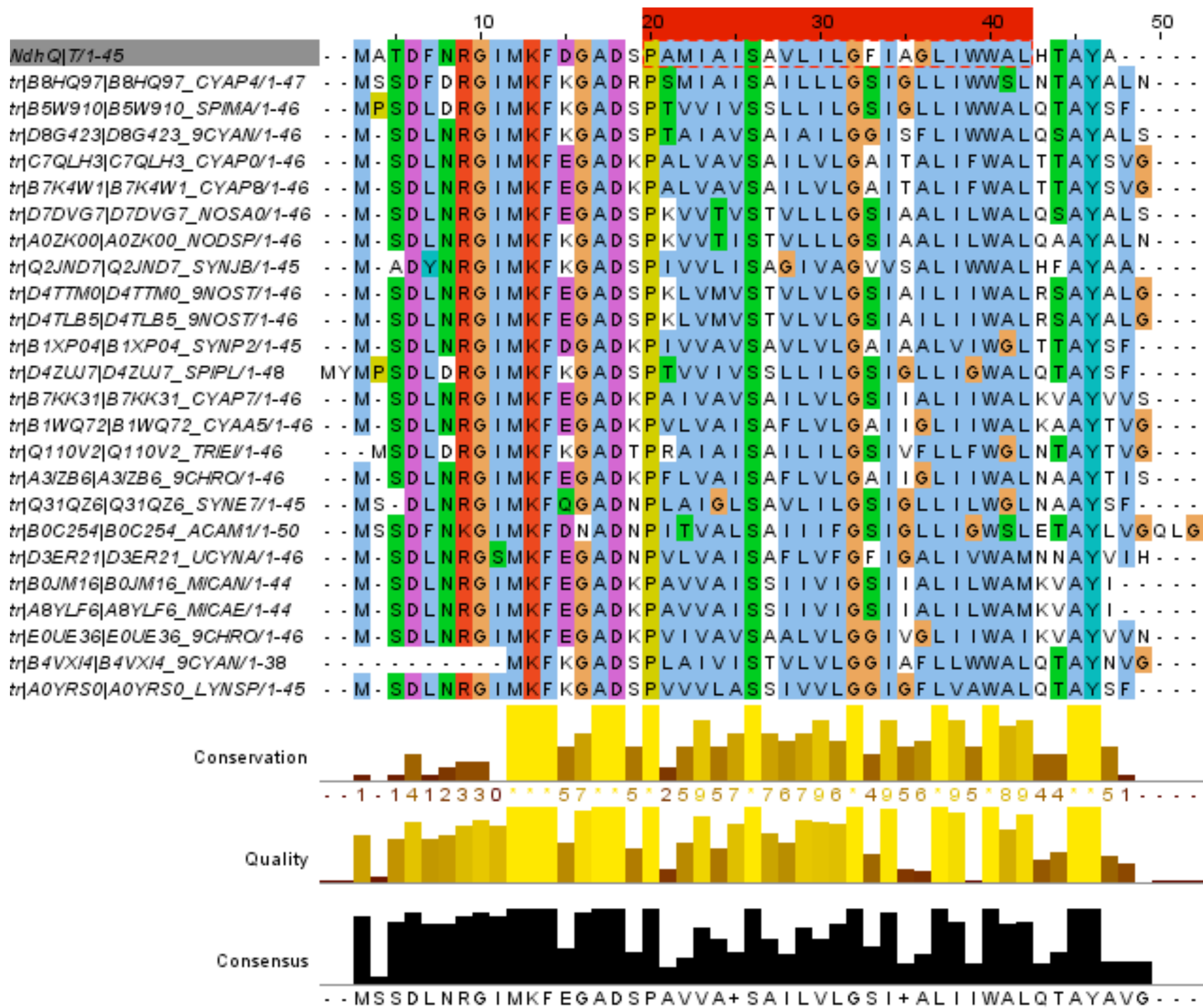
Prediction of transmembrane helices for NdhP and NdhQ. The translated amino acid sequences of *ndhP* and *ndhQ* were used to predict transmembrane helices with TMHMM (http://www.cbs.dtu.dk/services/TMHMM/). The red bar indicates the position of the TMH within the sequence.

Figure S4

NdhP

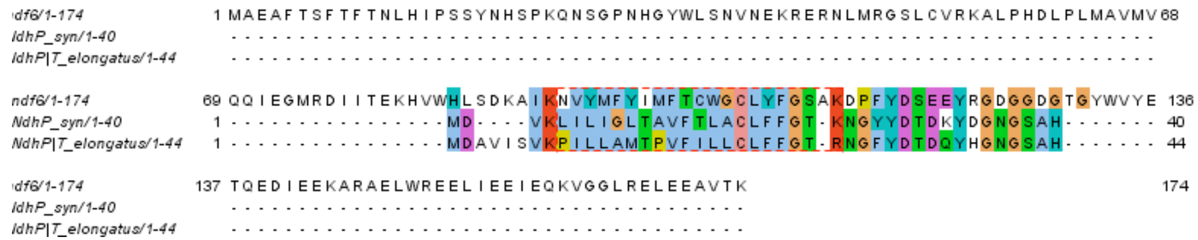


NdhQ



Multiple sequence alignment of proteins similar to NdhP and NdhQ. The multiple sequence alignment was performed with the programmes CLUSTAL-X (Thompson et al. 1997) and Jalview (Waterhouse et al. 2009). The residue colour code is as follows: orange, Gly; yellow, Pro; green, Thr, Ser, Asn and Gln; red, Lys and Arg; blue, Trp, Met, Val, Ile, Ala, Leu and Phe; pink, Cys; cyan, His and Tyr; and magenta, Glu and Asp. The red bar indicates the position of the TMH within the sequence.

Figure S5



Sequence comparison of NdhP and NDF6 (AT1G18730). The sequence alignment was performed with the programmes CLUSTAL-X (Thompson et al. 1997) and Jalview (Waterhouse et al. 2009). The residue colour code is as follows: orange, Gly; yellow, Pro; green, Thr, Ser, Asn and Gln; red, Lys and Arg; blue, Trp, Met, Val, Ile, Ala, Leu and Phe; pink, Cys; cyan, His and Tyr; and magenta, Glu and Asp. The red bar indicates the position of the TMH within the sequence.

References:

Battchikova, N., P. P. Zhang, S. Rudd, T. Ogawa and E. M. Aro (2005). "Identification of NdhL and Ssl1690 (NdhO) in NDH-1L, and NDH-1M complexes of Synechocystis sp PCC 6803." Journal of Biological Chemistry **280**(4): 2587-2595.

Blum, H., H. Beier and H. J. Gross (1987). "Improved Silver Staining of Plant-Proteins, Rna and DNA in Polyacrylamide Gels." Electrophoresis **8**(2): 93-99.

El-Mohsnawy, E., M. J. Kopczak, E. Schlodder, M. Nowaczyk, H. E. Meyer, B. Warscheid, N. V. Karapetyan and M. Rogner (2010). "Structure and Function of Intact Photosystem 1 Monomers from the Cyanobacterium Thermosynechococcus elongatus." Biochemistry **49**(23): 4740-4751.

Kuhl, H., J. Kruip, A. Seidler, A. Krieger-Liszkay, M. Bunker, D. Bald, A. J. Scheidig and M. Rogner (2000). "Towards structural determination of the water-splitting enzyme - Purification, crystallization, and preliminary crystallographic studies of photosystem ii from a thermophilic cyanobacterium." Journal of Biological Chemistry **275**(27): 20652-20659.

Rippka, R., J. Deruelles, J. B. Waterbury, M. Herdman and R. Y. Stanier (1979). "Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria." Journal of General Microbiology **111**(Mar): 1-61.

Shevchenko, A., M. Wilm, O. Vorm and M. Mann (1996). "Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels." Analytical Chemistry **68**(5): 850-858.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins (1997). "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." Nucleic Acids Research **25**(24): 4876-4882.

Waterhouse, A. M., J. B. Procter, D. M. A. Martin, M. Clamp and G. J. Barton (2009). "Jalview Version 2-a multiple sequence alignment editor and analysis workbench." Bioinformatics **25**(9): 1189-1191.

Whitelegge, J. P. (2002). "Plant proteomics: BLASTing out of a MudPIT." Proceedings of the National Academy of Sciences of the United States of America **99**(18): 11564-11566.