Structural and Functional Study of NADH:ubiquinone Oxidoreductase (Complex I) from *Aquifex aeolicus*

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften

vorgelegt beim Fachbereich 14
Biochemie, Chemie und Pharmazie
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von Wenxia Liu
aus Shandong China

Frankfurt am Main (2018)
(D30)

vom Fachbereich Biochemie, Chemie und Pharmazie der Johann Wolfgang Goethe Universitaet als Dissertation angenommen

Dekan: Prof. Dr. Clemens Glaubitz

1. Gutachter: Prof. Dr. Klaas Martinus Pos
2. Gutachter: Prof. Dr. Dr. h.c. Hartmut Michel

Datum der Disputation:

Diese Doktorarbeit wurde vom 20. November 2013 bis zum 17. November 2018 unter Leitung von Prof. Dr. Hartmut Michel in der Abteilung für Molekulare Membranbiologie am Max-Planck-Institut für Biophysik in Frankfurt am Main durchgeführt.

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig angefertigt habe und keine weiteren Hilfsmittel und Quellen als die hier aufgeführten verwendet habe.

_____

(Wenxia Liu)

Frankfurt am Main, den

Publications:

W. Liu, A. Liu, H. Gao, Q. Wang, L. Wang, E. Warkentin, Z. Rao, H. Michel, G. Peng. Structural properties of the peroxiredoxin AhpC2 from the hyperthermophilic eubacterium *Aquifex aeolicus*, Biochimica et Biophysica Acta (BBA) - General Subjects, 1862 (2018) 2797-2805.

G. Peng, B. Meyer, L. Sokolova, **W. Liu**, S. Bornemann, J. Juli, K. Zwicker, M. Karas, B. Brutschy, H. Michel. Identification and characterization two isoforms of NADH:ubiquinone oxidoreductase from the hyperthermophilic eubacterium *Aquifex aeolicus*, Biochimica et Biophysica Acta (BBA) -Bioenergetics, 1859 (2018) 366-373.

Table of contents

List of Tables

List of Figures

I. Zusammenfassung

NADH:Ubichinon-Oxidoreduktase (Komplex I) ist das erste und größte Enzym der Atmungskette. Es katalysiert die Reduktion von Ubichinon mit NADH über eine Vielzahl Enzym-gebundener Redoxzentren – ein Flavinmononukleotid (FMN) und acht Eisen-Schwefel (Fe-S)- Zentren – und verknüpft diese exotherme Reaktion mit dem endergonen Transport von vier Protonen durch die Membran. Bakterien enthalten eine Minimalversion von Komplex I mit einer Molekülmasse von etwa 550 kDa, die sich aus 14 konservierten Kernuntereinheiten aufbaut. Komplex I hat eine L-förmige Struktur bestehend aus 2 Komponenten (Arme). Der hydrophile Arm, der in das Cytosol (oder die Mitochondrienmatrix) hineinragt, enthält die Bindestelle für NADH, für den Zwei- zu Einelektronenschalter FMN und für alle Einelektron-übertragenden Fe-S Zentren and stellt überwiegend den Ort der katalytischen Reaktion dar. Der Membranarm besteht aus den membranständigen Untereinheiten and wirkt als Protonenpumpe. Die Ubichinon-Bindestelle liegt an der Kontaktstelle der beiden Arme.

Im Mittelpunkt dieser Arbeit steht der Komplex Iaus dem hyperthermophilen Bakterium *Aquifex aeolicus*. In Komplex I- Präparationen wurden 20 teilweise homologe Untereinheiten identifiziert, welche zwei Isoformen, NQOR1 und NQOR2, zugeordnet werden konnten. Die Kristallstruktur des hydrophilen Arms von NQOR1 bei 2.9 Å Auflösung erlaubte einen detaillierten Blick auf die räumliche Anordnung der Untereinheiten NuoB, E, F, G, I_1, and D_2 sowie ihrer Redoxzentren. Trotz seiner hohen Ähnlichkeit mit anderen Organismen enthält nur das *A. aquifex* Enzym in $NuoI_1$ eine N-terminale Helix, die sich bis in die Zellmembran hinein erstreckt, und ein zusätzliches tetranukleares Fe-S Zentrum. Dieses N8-Zentrum schafft eine elektronische Verbindung zwischen dem peripheren N7 und dem N4 Zentrum der zentralen Elektronentransferkette and könnte so einen alternativen Elektronentransferweg aufbauen. Das kleine Redoxprotein Alkylhydroperoxid-Reduktase C (AhpC2), welches in der gereinigten Komplex I-Probe mittels Massenspektrometrie detektiert wurde, könnte ein potentielles Substrat oder zumindest ein Bindungspartner für den *A. aelicus* Komplex Idarstellen.

Auf dieser Grundlage hatte ich in meiner Doktorarbeit das Ziel, (1) die beiden Isoformen aus der, aus *A. aeolicus* gereinigten Komplex I - Probe zu trennen, um sie einzeln untersuchen zu können; (2) die Fe-S-Zentren insbesondere das neue N8-Zentrum durch EPR-Messungen zu charakterisieren; (3) vorhandene Hinweise einer Wechselwirkung zwischen Komplex I und AhpC2 zu belegen; und (4) ein heterologes genetisches System für den *A. aeolicus* Komplex I zu etablieren. Dieses Projekt stellt eine große Herausforderung dar, doch ohne ein geeignetes genetisches System ist eine umfassende Untersuchung von Komplex I aus *A. aeolicus* nur beschränkt realisierbar. Genetische Veränderungen in *A. aeolicus* sind noch nicht möglich.

(1) Darstellung getrennter Isoformen von *A. aeolicus* Komplex I

Die Präparation von Komplex I aus *A. aeolicus* führte immer zu einer Probe, die beide Isoformen nebeneinander enthielten. Sie ko-eluierten stets als Mischung in denselben chromatographischen Fraktionen und erschienen als eine Bande auf den IEF and BN-PAGE Gelen. In dieser Arbeit wurden 2 Lösungsansätze verfolgt: (a) Trennung der zusammen gereinigten Komplex I -Isoformen durch Immunopräzipitation und (b) heterologe Produktion einzelner Untereinheiten, Subkomplexe und der vollständigen Isoform NQOR1 in *E. coli*.

(a) Trennung der beiden Komplex I Isoformen mittels Immunopräzipitation

Affinitätschromatographie gereinigte polyklonale Antikörper wurden zur Präzipitation der S1-Probe von *A. aeolicus* Komplex I eingesetzt. Anti-NuoI$_1$ oder Anti-NuoD$_2$ sollten an NQOR1, und Anti-NuoI$_2$ oder Anti-NuoD$_1$ an NQOR2 binden. Anfangs wurde der Pierce Classic IP Kit für die Experimente verwendet. Das molare Verhältnis von Antikörper/Antigen (Ab/Ag) war 1:1. Komplex I wechselwirkte nur schwach mit Anti-NuoI$_1$ und gar nicht mit Anti-NuoI$_2$. Unglücklicherweise überlagerten sich die Banden von NuoD$_1$ (68.7 kDa) und NuoD$_2$ (67.9 kDa) mit dem ko-eluierten Antikörper (75 kDa). Um die Ko-Eluation mit Antikörpern zu vermeiden, wurde der Pierce™ Co-Immunoprecipitation Kit für weitere Experimente verwendet. Für die Anti-NuoI$_1$ und Anti-NuoI$_2$ polyklonalen Antikörper wurden molare Verhältnisse von 2:1, 5:1 und 10:1 für Ab/Ag getestet. Als optimales Ab/Ag Verhältnis hat sich 10:1 erwiesen. Trotzdem

war die Bindung noch immer zu schwach und die größte Menge des Antigens befand sich im Durchfluss. Das Protein, das mit Anti-NuoI$_1$ präzipitiert, wurde, überraschenderweise, mit allen vier Antikörpern gefunden. Das Protein, das mit Anti-NuoI$_2$ präzipitiert, konnte mit Anti-NuoI$_2$ and Anti-NuoD$_1$ detektiert warden. Kein Signal wurde mit Anti-NuoI$_1$ detektiert und ein extrem schwaches mit Anti-NuoD$_2$. Eine Immunopräzipitation mit Anti-NuoD$_1$ und Anti-NuoD$_2$ bei einem molaren Ab/Ag Verhältnis von 10:1 wurde ebenfalls getestet, jedoch war kein Signal im Western Blot sichtbar. Der vielversprechendste Kandidat für eine Immunopräzipitation zur Trennung von NQOR1 und NQOR1 ist also Anti-NuoI$_2$, aber weiter Optimierungsversuche sind notwendig. Die Immunopräzipitationsexperimente wurden nicht weiterverfolgt, weil sich ab einem bestimmten Zeitpunkt der zweite Lösungsansatz erfolgsversprechender erwies.

(b) Heterologe Produktion von Komplex I in *E. coli*

Die 13 Gene, welche für den Komplex I kodieren, sind über mehrere Operone in der genomischen DNA von *A. aeolicus* verstreut. Aufgrund seiner enormen Größe (> 500 kDa) und des Einbaus von bis zu 10 Fe-S-Zentren ist es bisher nicht gelungen den vollständigen *A. aeolicus* Komplex I heterolog zu produzieren. Bisher wurde nur von der Produktion eines NuoE - NuoF Subkomplexes berichtet. Als Expressionswirt wurde der bestens etablierte Modellorganismus *E. coli* gewählt, der ein komplettes ISC und SUF System zur Reifung der Fe-S –Zentren besitzt.

Die Produktion von *A. aeolicus* Komplex I in *E. coli* wurde als Projekt mit mehreren Teilschritten geplant. Im ersten Schritt wurden Gene, die für NuoB, D$_1$, D$_2$, G, I$_1$, and I$_2$ kodieren, separat produziert, wodurch vergleichende Untersuchungen zwischen den Isoformen (NuoD$_1$ vs NuoD$_2$, and NuoI$_1$ vs NuoI$_2$) möglich werden. Jedes *nuo* Gen wurde aus dem *A. aeolicus* Genom amplifiziert und in eine multiple Klonierungstelle des pET32b Vektors eingesetzt. Alle sechs Untereinheiten konnten erfolgreich exprimiert werden. NuoB und NuoI$_1$ konnten rein und homogen dargestellt werden. NuoG, NuoD$_1$, and NuoD$_2$ aggregierten jedoch trotz mehrerer Optimierungsversuche. Zusätzliche Expressionsvektoren für die NuoI$_2$ Produktion wurden von, in der Abteilung modifizierten, pTTQ18 / pBAD / pQE Vektoren abgeleitet. pBAD-C3-*nuo*I$_2$ / TOP10

erwies sich also am geeignetsten zur Produktion von $NuoI_2$, welches darauffolgend in homogener Form aufgereinigt werden konnte. Im zweiten Schritt wurde das künstliche Operon *nuo*BD_2EFGI_1 in den Vektor pBAD33 mit dem Ziel eingesetzt, ein vollständig assembliertes NADH-Dehydrogenase Fragment (NuoEFG) zu exprimieren. Mit diesem Ansatz gelang es jedoch nur, den NuoE - NuoF Subkomplex darzustellen. Der dritte Schritt umfasste die Produktion des gesamten NQOR1 Komplexes. Aufgrund der Verteilung der Gene von NQOR1 auf drei Orten im Genom und ihrer immensen Gesamtlänge (~ 14 kb), wurden zwei Vektoren für eine Ko-Expression verwendet. 13 *nuo* Gene wurden nach Reorganisation in die Operone *nuo*EFG und *nuo*$A_2BD_2H_1I_1J_1K_1L_1M_1N_1$, eingesetzt, welche anschließend in die Vektoren pBAD-CM1 bzw. pBAD33 kloniert wurden. Die nativen Translationsinitiationsregionen des *A. aeolicus* Genoms wurden ebenfalls in beiden Operone eingesetzt. Nach Verifizierung durch Sequenzierung wurden die Expressionsvektoren in den *nuo*-Deletion *E. coli* Stamm BA14 zwecks induktiver Expression ko-transformiert. Danach wurde die heterolog produzierte strep-getaggte NQOR1 mit Hilfe einer Anionenaustaucher- und Strep-Tactin Affinitätschromatographie gereinigt. Auf einem anschließend angefertigten BN-PAGE-Gel waren sechs Banden sichtbar. All zeigten eine NADH-Dehydrogenase Aktivität, die direkt auf dem Gel detektiert wurde.

Die Banden, nach zunehmender Molekularmasse von 1 bis 6 nummeriert, wurden ausgeschnitten und die Zusammensetzung jeder Bande durch MS bestimmt. Kein Protein aus *E.coli* wurde in den Banden 1-5 detektiert. Bande 5 enthält alle 13 Untereinheiten und damit einen kompletten Komplex I. Banden 1-4 repräsentieren Sub-Komplexe von NQOR1, denen eine oder mehrere Membranuntereinheiten fehlen. Es ist denkbar, dass Membranuntereinheiten abgebaut wurden trotz des zugesetzten Cocktails aus Proteaseinhibitoren-, Metallchelatoren- und einer reduzierenden Substanz. Von großem Interesse ist Bande 6, die bei einer hohen Molekülmasse (> 720 kDa) eluiert und den hydrophilen Arm (NuoB, D_2, E, F, G, and I_1) und die Untereinheit $NuoH_1$ des Membranarmes, sowie Transkriptionsterminationfaktoren, ribosomale Proteine, Polypeptidelongationsfaktoren für die Proteinbiosynthese, Chaperone, anabole und katabole Proteine und einges mehr enthält. Es scheint so, dass ein aktives Ribosom, assoziiert mit dem hydrophilen Arm von Komplex I, während des Translations- und

Assemblierungsprozesses abgefangen wurde. Daher könnte die unvollständige Komplex I
- Komponente ein Zwischenprodukt des Assemblierungsprozesses darstellen.

Die heterologe Produktion des intakten Komplexes Iwurde durch
elektronenmikroskopische Aufnahmen bestätigt. Die charakteristische L-förmige Struktur
des negativ gefärbten Komplex I ist darin klar erkennbar. 2-D Klassifizierungsversuche
ergaben jedoch einen verkürzten Membranarm im Vergleich zu dem aus *A. aeolicus*
präpariertem Komplex I. Zu wenige Teilchen entsprachen dem vollständigen NQOR1. In
vollständiger Übereinstimmung mit den biochemischen und strukturellen Daten
katalysiert die rekombinant produzierte NQOR1 die Reduktion von Quinon mit NADH.

Wenig experimentelle Daten liegen über den Assemblierungsmechanismus von
bakteriellem Komplex I vor, obwohl bereits ein mögliches Szenario aus Studien am
E.coli Enzym vorgeschlagen wurde. Danach wird NuoA als erstes produziert, wodurch
sich der Ribosom- *nuo*-mRNA -Komplex an die bakterielle Membran anlagert und die
Translation von *nuo*B zu *nuo*I in Membrannähe ermöglicht. Als Nächstes wird NuoH, das
von einer polyzistronischen mRNA translatiert wird, gefolgt von NuoJ and NuoK an den
hydrophilen Arm gebunden und der vollständige Komplex I durch Anlagerung des
unabhängig gebildeten Assemblierungszwischenproduktes NuoL, NuoM and NuoN
komplettiert. Ich bin ebenfalls der Meinung, dass die Synthese und Assemblierung der
Nuo Untereinheiten in einer konzertierten Weise erfolgt, dass jedoch die
Zusammensetzung der Subkomplexe in den BN-PAGE-Banden, insbesondere in der
Bande 6, eine anderen Abfolge der Assemblierung nahelegt. Aufgrund der Lokalisierung
der Gene, die für die Q (NuoB, D_2, and I_1) und N (NuoE, F and G)-Module kodieren, in
zwei Operonen, sollten ihre Transkription, Translation, und Assemblierung unabhängig
verlaufen. Anschließend schließen sie sich dann zum hydrophilen Arm, welcher in Bande
6 gefunden wurde, zusammen. Die N-terminale hydrophobe Helix von $NuoI_1$ könnte nun
dazu dienen, den hydrophilen Arm in der Membran zu verankern. Danach könnte
zunächst $NuoH_1$, gefolgt von $NuoA_2$, $NuoJ_1$, $NuoK_1$, and $NuoN_1$ und schließlich $NuoM_1$
and $NuoL_1$ angelagert werden.

(2) Identifizierung der Fe-S Zentren durch EPR Experimente

Der *A. aeolicus* Komplex I enthält FMN und zehn Fe-S Zentren, die für den Elektronentransfer von NADH zu Ubiquinon erforderlich sind. N1a, N1b, N2, N3, N4, N5, N6a, und N6b Zentren sind strikt konserviert. Das N7 Zentrum wurde nur in Komplex I von einigen Bakterien wie *E. coli* und *Thermus thermophilus* gefunden. Das N8 Zentrum ist ein neues vierkerniges Fe-S Zentrum, welches bisher nur in *A. aeolicus* Komplex Igefunden wurde. Für EPR Messungen wurde der aus *A. aeolicus* direkt präparierte Komplex I, das heterolog produzierte NQOR1 und die heterolog erhaltenen einzelnen Untereinheiten herangezogen. Zwei Fe-S Zentren (N1b and N2) konnten eindeutig dem nativen *A. aeolicus* Komplex Izugeordnet werden. Kein EPR Signal für Fe-S Zentren wurde in den einzelnen Untereinheiten NuoB, NuoI$_1$ and NuoI$_2$, weder im oxidierten noch im reduzierten Zustand detektiert. Oxidiertes NuoG zeigt ein EPR Signal für ein 3Fe-4S Zentrum, welches bisher nicht in Komplex I detektiert wurde, während reduziertes NuoG Ähnlichkeiten zu einem 2Fe-2S Zentrum hat. Im Vergleich zum N1b [2Fe-2S] Zentrum des vollständigen Komplexes I ist es schwächer, breiter und inhomogener. Vermutlich wurden die Fe-S Zentren während der Präparation der separaten Untereinheiten verloren oder beschädigt. Die EPR Spektren des rekombinanten NQOR1 und des direkt isolierten *A. aeolicus* Komplexes I ähneln sich sehr, was dafür spricht, dass alle Fe-S Zentren und alle Unterheiten mit hoher Besetzung in *E.coli* produziert wurden.

(3) Untersuchungen der Wechselwirkung zwischen AhpC2 und Komplex I

Frühere massenspektrometrische Daten zeigten, dass AhpC2 ein potentielles Substrat oder Bindingspartner für den *A. aeolicus* Komplex I darstellt, und eine Rolle als terminaler Elektronendonor/akzeptor in einem alternative Elektrontransferweg, vermittelt durch das neue Fe-S Zentrum N8, spielen könnte. Aminosäuresequenzanalysen ordnen AhpC2 der 1-Cys Peroxiredoxinfamilie zu, deren Mitglieder eine breite Palette von Hydroperoxidsubstraten einschließlich Phospholipidhydroperoxiden (PLOOH) reduzieren können. Für molekulare Untersuchungen etablierte ich erfolgreich ein heterologes Expressionssystem für AhpC2 in *E. coli* und präparierte das Protein in hoher Reinheit und Homogenität. Die Struktur von AhpC2 wurde mittels Röntgenkristallographie und Elektronenmikroskopie als ringförmiger dodekamerer

Proteinkomplex aufgeklärt. Das Redoxpotential von AhpC2 von etwa -310 mV ähnelt damit jenem von NADH (-320 mV) und ist niedriger als dasjenige der Fe-S Zentren von Komplex I mit Ausnahme von N1a (-340 mV). Ein Elektronentransfer zwischen AhpC2 and Komplex I wäre also möglich. Ferner untersuchte ich die Wechselwirkung zwischen AhpC2 und dem aus *A. aeolicus* gereinigten Komplex I mit der Oberflächenplasmonresonanz-Methode. Es stellte sich heraus, dass AhpC2 eine sehr hohe Affinität zu Komplex I hat. Der K_D Wert ist 0.478 nM and 4.84 nM für eine AhpC2 -Probe, die mit bzw. ohne NaCl behandelt wurde. Elektronenmikroskopische Aufnahmen an negativ gefärbten AhpC2 Proben zeigten nur einen dodekameren AhpC2-Komplex in Abwesenheit von NaCl, bei Behandlung mit 1 M NaCl zerfallen sie jedoch in Monomere. Offensichtlich beeinflusst der oligomere Zustand von AhpC2 seine Affinität zu Komplex I und vielleicht auch seine Funktion. Obwohl Ko-Kristallisationsexperimente zwischen Komplex I und AhpC2 zu Kristallen führten, streuten diese zu schwach um für eine Strukturanalyse tauglich zu sein.

Zusammenfassend liefert die erfolgreiche Produktion der vollständig assemblierten und aktiven Komplex I Isoform NQOR1 aus dem hyperthermophilen Bakterium *A. aeolicus* in *E. coli* eine vielversprechende Plattform für zukünftige strukturelle und funktionelle Untersuchungen von *A. aeolicus* Komplex I. Die hohe Qualität des rekombinanten NQOR1 wurde durch EPR, Enzymaktivität und EM –Daten und deren Vergleich mit den analogen Daten von aus *A. aeolicus* gereinigten Komplex Iuntermauert. Ferner stellt das in dieser Arbeit vorgestellte heterologe Expressionssystem eine solide Vorlage für die Produktion der Isoform NQOR2 dar. Das heterologe genetische System erlaubt uns auch Mutagenese und Quervernetzungsexperimente durchzuführen, wodurch viele Türen für zukünftige Untersuchungen geöffnet werden. Abschließend deutet die hohe Affinität zwischen Komplex I und AhpC2 auf eine wesentliche Rolle von AhpC2 in Verbindung mit der Funktion von Komplex I *in vivo* hin. Deren Aufklärung stellt ein interessantes Zunftsprojekt dar.

II. Summary

NADH:ubiquinone oxidoreductase (Complex I) is the first and largest enzyme in the respiratory chain. It catalyzes the transfer of two electrons from NADH to ubiquinone via a series of enzyme-bound redox centers - Flavin mononucleotide (FMN) and iron-sulfur (Fe-S) clusters – and couples the exergonic reaction with the endergonic translocation of four protons across the membranes. Bacteria contain the minimal form of complex I, which is composed of 14 conserved core subunits with a molecular mass of around 550 kDa. Complex I has an L-shaped structure which can be subdivided into two major parts (arms). The hydrophilic arm protruding into the bacterial cytosol (or mitochondrial matrix) harbors the binding site for the substrate NADH, the two- to one-electron switch FMN and all one-electron transferring Fe-S clusters and therefore considered as the catalytic unit. The membrane arm consists of the membrane-spanning subunits and conducts the proton pumping process. The Quinone binding site is located at the interface of both arms.

The previous work of my research group focused on complex I from the hyperthermophilic bacterium *Aquifex aeolicus*. In purified complex I samples 20 partially homologous subunits have been identified, which could be assigned to two different isoforms, named NQOR1 and NQOR2. The crystal structure of the hydrophilic arm of NQOR1, determined at 2.9 Å resolution, reveals the structural arrangement of NuoB, E, F, G, I_1, and D_2 as well as their redox centers which largely corresponds to the architecture of the hydrophilic arms of other organisms. However, it is interesting to discover that the N-terminal helix of $NuoI_1$ is unexpectedly inserted into the membrane and a novel tetranuclear Fe-S cluster termed N8. It electronically connects the peripheral N7 with the N4 of the main electron transfer chain, which may constitute an alternative electron transfer pathway. A small redox protein alkyl hydroperoxide reductase C (AhpC2), identified by mass spectrometry in the purified complex I sample, might be a potential substrate or binding partner of *A. aeolicus* complex I.

On this basis my doctoral work pursues the following objectives: (1) to separately prepare NQOR1 and NQOR2 for individual structural and functional studies; (2) to characterize the Fe-S clusters by EPR measurements with the special focus on the spectral properties of the novel N8 cluster; (3) to investigate the interactions between AhpC2 and complex I, and (4) to establish a heterologous genetic system to produce the *A. aeolicus* complex in *E. coli*. This highly challenging project is attempted because comprehensive studies on complex I from *A. aeolicus* are severely hampered due to the lack of a genetic system. Genetic manipulations in *A. aeolicus* are not feasible yet. Therefore, the construction of a heterologous genetic system for the *A. aeolicus* complex I will supply a platform for these studies.

(1) Isolation and purification of the individual isoforms of *A. aeolicus* complex I

In complex I samples prepared from *A. aeolicus* cells, two isoforms were always co-eluted as a mixture in the same chromatographic fractions and appeared as one band on IEF and BN-PAGE gels. In this work, two solution strategies for this severe obstacle were pursued: (a) the separation of the two isoforms after purification from *A. aeolicus* using immunoprecipitation and (b) the heterologous production of single subunits, sub-complexes and the entire NQOR1 from both isoforms in *E. coli*.

(a) Separation of two isoforms of complex I using immunoprecipitation

Polyclonal antibodies after affinity purification were used to precipitate complex I two isoforms from the preparations of *A. aeolicus* S1 sample. Anti-NuoI$_1$ or anti-NuoD$_2$ potentially target NQOR1, and anti-NuoI$_2$ or anti-NuoD$_1$ target NQOR2. The precipitated protein was further validated by western blot using these four antibodies. The initial experiment was performed using the Pierce Classic IP Kit. The molar ratio of Ab/Ag was 1:1. Unfortunately, the bands of NuoD$_1$ (68.7 kDa) and NuoD$_2$ (67.9 kDa) were masked by the co-eluted antibody (75 kDa). Complex I only revealed a weak binding to anti-NuoI$_1$ and no binding to anti-NuoI$_2$. To eliminate the co-elution of antibodies, the Pierce™ Co-Immunoprecipitation Kit was used for further experiments. Different molar ratios of Ab/Ag (2:1, 5:1 and 10:1) were tested using anti-NuoI$_1$ and anti-NuoI$_2$. The optimal Ab/Ag ratio was proved to be 10:1. Nevertheless, the binding was still too weak

and most of the antigen appeared in the flow through. The protein precipitated by anti-$NuoI_1$ could be, unexpectedly, detected by all of four antibodies, while the protein precipitated by anti-$NuoI_2$ could be detected by anti-$NuoI_2$ and anti-$NuoD_1$ as expected, no signal was detected by anti-$NuoI_1$ and extremely weak signals were detected by anti-$NuoD_2$. Immunoprecipitation using anti-$NuoD_1$ and anti-$NuoD_2$ with the Ab/Ag molar ratio of 10:1 does not exhibit a protein signal in the Western blot. The most promising candidate to utilize immunoprecipitation for separating NQOR1 and NQOR2 appears to be anti-$NuoI_2$, but further optimization is required. This approach was abandoned because the second strategy was more successful at a certain stage.

(b) Heterologous production of *A. aeolicus* complex I in *E. coli*

13 genes encoding complex I are dispersed in different operons in the genomic DNA of *A. aeolicus*. Due to its large size (> 500 kDa) and the incorporation of up to ten iron-sulfur (Fe-S) clusters, no complete *A. aeolicus* complex I has been heterologously produced so far. Solely, the production of the NuoE - NuoF subcomplex was reported. Here, *E. coli* was chosen as the expression host which possesses a complete ISC and SUF system for Fe-S cluster assembly.

Production of the *A. aeolicus* complex I in *E. coli* was planned in several steps. In the first step, single genes encoding for NuoB, D_1, D_2, G, I_1, and I_2 were separately expressed for further characterizations and comparisons with the respective isoforms ($NuoD_1$ vs $NuoD_2$, and $NuoI_1$ vs $NuoI_2$). Each *nuo* gene was amplified from the *A. aeolicus* genome and first inserted into multiple cloning sites of the pET32b vector. All six subunits could be successfully expressed. NuoB and $NuoI_1$ could be purified to homogeneity, but NuoG, $NuoD_1$, and $NuoD_2$ aggregated even after diverse optimization attempts. For the successful production of $NuoI_2$, the in-house modified pTTQ18 / pBAD / pQE vectors termed pBAD-A2/C3, pTTQ-A2/C3 and pQE-A2/C3 were applied. pBAD-C3-*nuo*I_2 / TOP10 was most suitable to produce $NuoI_2$ which could be homogeneously purified. In the second step, an artificial operon *nuo*BD_2EFGI_1 was cloned into pBAD33 with the aim to express the genes in an essentially natural fashion and thus to obtain a fully assembled NADH dehydrogenase fragment (NuoEFG). However, only a subcomplex composed by NuoE and NuoF could be isolated. In the last step, production of the entire NQOR1 was

attempted. Due to the dispersion of the genes encoding NQOR1 in three loci and the immense size (~ 14 kb) of the *nuo* genes, two vectors were used for co-expression. 13 *nuo* genes reorganized into two operons as *nuo*EFG and *nuo*$A_2BD_2H_1I_1J_1K_1L_1M_1N_1$ were cloned to vectors pBAD-CM1 and pBAD33, respectively. The native translation initiation regions from *A. aeolicus* were also incorporated into the two operons. After verification by sequencing, the expression vectors were co-transformed into the *nuo* deletion *E. coli* strain BA14 for inductive expression. Afterwards, the heterologously produced NQOR1 was purified by anion exchange and Strep-Tactin affinity chromatography. The obtained strep-tagged NQOR1 was evaluated by BN-PAGE and by an in-gel activity assay. Unexpectedly, six bands were present in the BN-PAGE gel. All of them exhibited NADH dehydrogenase activity.

The bands, numbered 1-6 in the order of increased molecular masses, were sliced and the compositions of each band were determined by mass spectrometry. No proteins from *E. coli* were detected in bands 1-5. Band 5 contained all 13 subunits and thus the fully assembled complex of *A. aeolicus*. Bands 1-4 revealed subcomplexes of NQOR1 lacking one or a few of the membrane subunits. One possibility might be that the membrane subunits are degraded, although protease inhibitors, a reducing agent and metal chelators were added to avoid proteolysis during protein production and purification. Most interestingly, band 6 found at a higher molecular mass (> 720 kDa) contains the complete hydrophilic arm (NuoB, D_2, E, F, G, and I_1) and subunit $NuoH_1$ of the membrane arm, as well as transcription termination factors, ribosomal proteins, polypeptide elongation factors for protein biosynthesis, chaperonins and proteins involved in anabolism and catabolism, etc.. It seems that an active ribosome with an attached hydrophilic arm of complex I had been captured during the translation and protein assembly process. Therefore, the incomplete complex I could be an intermediate of the assembly process. Little hard data are known about the assembly of bacterial complex I, although a possible assembly pathway was proposed based on studies on the in *E. coli* complex I. Accordingly, NuoA was produced first, anchors the ribosome – *nuo*-mRNA complex to the bacterial membrane and conducts the translation of *nuo*B to *nuo*I close to the membrane. Then, NuoH translated from the polycistronic mRNA was attached to the

hydrophilic arm followed by NuoJ and NuoK. Finally, an assembly intermediate consisting of NuoL, NuoM, and NuoN was associated. I also think that the synthesis and assembly of Nuo subunits proceed in a concerted fashion, but the compositions of the subcomplexes in the BN-PAGE bands, especially of that of band 6, suggest a different scenario concerning the sequence of events. Due to the localization of the genes encoding the subunits of the Q (NuoB, D_2, and I_1) and the N (NuoE, F and G) modules of NQOR1 inside two operons, their transcription, translation, and assembly should proceed independently. The Q and N modules subsequently associate to the hydrophilic arm found in the protein complex of band 6. The N-terminal helix of $NuoI_1$ might serve as an anchor to connect in the next step the hydrophilic arm with the membrane surface. Afterwards, $NuoH_1$ might be attached to the hydrophilic arm, then $NuoA_2$, $NuoJ_1$, $NuoK_1$, and $NuoN_1$ and finally $NuoM_1$ and $NuoL_1$.

The successful assembly of the intact complex I was confirmed using single-particle electron microscopy. The L-shaped structure could be clearly observed in negative staining images. The 2-D classification was attempted, but the number of particles of the fully assembled complex I was insufficient yet to reconstitute an entire model. Only a low-quality model could be, so far, obtained with a truncated membrane arm. In complete agreement with the biochemical and structural data, the recombinant NQOR1 catalyzes the electron transfer from NADH to ubiquinone.

(2) Identification of the Fe-S clusters by EPR measurement

A. aeolicus complex I contains FMN and ten Fe-S clusters which are required for the transfer of electrons from NADH to ubiquinone. N1a, N1b, N2, N3, N4, N5, N6a, and N6b clusters are conserved. N7 cluster is discovered in some bacteria, such as *E. coli* and *T. thermophilus*. N8 cluster is a novel tetranuclear Fe-S cluster, which only occurs in the *A. aeolicus* complex I. Complex I purified from *A. aeolicus*, as well as recombinant single subunits and NQOR1, were analyzed by EPR measurements. Two Fe-S clusters (N1b and N2) could be assigned unambiguously in the native *A. aeolicus* complex I. No EPR signals of Fe-S clusters were detected in the single subunits NuoB, $NuoI_1$, and $NuoI_2$, neither in the oxidized nor in the reduced state. Oxidized NuoG revealed an EPR signal for a 3Fe-4S cluster which is normally absent in complex I, while reduced NuoG

exhibited a signal which might originate from a 2Fe-2S cluster. This signal has similarities to that of N1b of entire complex I but is much weaker, broader and more inhomogeneous. It was assumed that the Fe-S clusters were lost or damaged during the preparation of the subunits. The EPR spectra obtained from NQOR1 were similar to those of the native *A. aeolicus* complex suggesting that all Fe-S cluster and the subunits of the hydrophilic arm are bound with a high occupancy.

(3) Investigation of the interactions between AhpC2 and complex I

Previous mass spectrometric data indicated that AhpC2 is a potential substrate or binding partner of the *A. aeolicus* complex I which might play a role as terminal electron donor/acceptor in an alternative electron transfer pathway via the novel Fe-S cluster N8. To test this hypothesis, we first performed a heterologous production and characterization of AhpC2. Amino acid sequence analysis classified AhpC2 as a member of the 1-Cys peroxiredoxin family, which has the capability to reduce a broad range of hydroperoxide substrates including phospholipid hydroperoxides (PLOOH). This protein could be produced in *E. coli* and purified to homogeneity. The structures of AhpC2 were determined by X-ray crystallography and EM, which both reveal a ring-shaped dodecamer. The redox potential of AhpC2 was determined to around -310 mV, which is comparable to that of NADH (-320 mV) and lower than that of the Fe-S clusters in complex I except for the one of N1a (-340 mV). Therefore, an electron transfer between AhpC2 and complex I is possible. Furthermore, the interaction of AhpC2 with the purified complex I isolated from *A. aeolicus* was investigated by surface plasmon resonance (SPR). Interestingly, it came out that AhpC2 has a very high affinity to complex I. The K_D value is 0.478 nM and 4.84 nM for an AhpC2 sample treated with and without NaCl, respectively. Negative stain electron microscopic data revealed a dodecameric structure of AhpC2 that dissociates into monomers after the treatment with 1 M NaCl. Obviously, the oligomeric state of AhpC2 influences its affinity to complex I and perhaps also the function AhpC2. Co-crystallization experiments between AhpC2 and complex I were also attempted, but the diffraction power of the resulting crystals was too low for further structural analysis.

In summary, the successful production of a fully assembled and active complex I isoform NQOR1 from the hyperthermophilic bacterium *A. aeolicus* in *E. coli* provides a promising platform for further structural and functional studies of *A. aeolicus* complex I. The high quality of the recombinant NQOR1 was supported by EPR, enzyme activity and EM data in comparison with the analogous data obtained for the native *A. aeolicus* complex I. The heterologous expression system described in this work provides a solid reference for producing the second isoform NQOR2. Furthermore, this system allows us to perform mutagenesis and cross-linking experiments, which opens manifold new opportunities for more specific studies in the future. Finally, the high affinity between complex I and AhpC2 implicates a crucial role of AhpC2 in connection with the function of complex I *in vivo*. The elucidation of the function of AhpC2 is an interesting future research project.

III. Abbreviation

The abbreviations used in this work are listed in the following table.

Table A. Abbreviations of chemicals

Abbreviations	Name
AAs	amino acids
ADP	adenosine-5′-diphosphate
AP	alkaline phosphatase
ATP	adenosine-5′-triphosphate
BCA	bicinchoninic acid
BCIP	5-bromo-4-chloro-indolyl-phosphatase
BisTris	1,3-bis(tris(hydroxymethyl)methylamino)propane
ddH2O	double-distilled water
dNTP	deoxyribonucleotide triphosphate
DDM	n-dodecyl-β-D-maltoside
DNA	deoxyribonucleic acid
EDTA	ethylendiaminetetracetic acid
EtBr	ethidium bromide
EtOH	ethanol
FADH2	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
HABA	2-(4'-hydroxy-benzeneazo)-benzoic acid
HAc	acetic acid
HEPES	4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid
IPTG	isopropyl β-D-1-thiogalactopyranoside
MES	2-(N-morpholino)-ethanesulfonic acid
MOPS	3-(N-morpholino)-propanesulfonic acid
NADH	Reduced nicotinamide adenine dinucleotide
NBT	nitro blue tetrazolium
Ni-NTA	nickel nitrilotriacetic acid
PVDF	polyvinylidene difluoride
RNA	ribonucleic acid
SDS	sodium dodecyl sulfate
Tris	tris-hydroxymethyl-aminomethane

Table B. Biological, biochemical and biophysical abbreviations

Abbreviations	Name
BN	blue-native
Complex I	NADH:ubiquinone oxidoreductase
DNase	deoxyribonuclease
EM	electron microscopy
EPR	electron paramagnetic resonance
ETC	electron transport chain
IMS	intermembrane space
MS	mass spectrometry
MW	molecular weight
OD	optical density
ORF	open reading frame
OXPHOS	oxidative phosphorylation
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PMF	proton-motive force
ROS	reactive oxygen source
SEC	size-exclusion chromatography
SPR	surface plasmon resonance
UV-VIS	ultraviolet–visible

Table C. Symbols for measures and units

Abbreviations	Name
Å	angstrom
bp	base pair
°C	degrees celsius
Da	dalton
e.g.	exempli gratia (Latin), for example (English)
et al.	et alii (Latin), and others (English)
etc.	et cetera (Latin), and so on (English)
h	hour
kDa	Kilodalton
min	minute
Pa	pascal
psi	pound per square inch
rpm	rotations per minute

RT	room temperature
s	second
3D	three-dimensional
U (enzymatic)	unit
V (as unit)	volt
v/v	volume/volume
w/v	weight/volume
w/w	weight/weight

Table D. Abbreviations of database and software

Abbreviations	Name
BLAST	Basic Local Alignment Search Algorithm
CCP4	Collaborative Computational Project number 4
ESPript	Easy Sequencing in PostScript
FASTA	FAST-All
NCBI	National Centre for Biotechnology Information
PDB	Protein Data Bank

1. Introduction

1.1 Electron Transport and Oxidative Phosphorylation

Most of the free energy produced during glycolysis, fatty acid oxidation and the citric acid cycle is retained in the reduced coenzymes such as nicotinamide adenine dinucleotide (NADH) and Flavin adenine dinucleotide (FADH$_2$). Both NADH and FADH$_2$ have a sufficiently low redox potential to reduce molecular oxygen to water and release a large amount of free energy to produce Adenosine triphosphate (ATP). The process, in which ATP is formed as a result of the transfer of electrons from NADH or FADH$_2$ to O$_2$, is called oxidative phosphorylation (OXPHOS). OXPHOS, executed in mitochondria or cell plasma membranes, is the major source of ATP in the aerobic organism.

Electrons flow from the electron donor (NADH or FADH$_2$) to the electron acceptor (O$_2$) via the electron transport chain (ETC) or respiratory chain, which is composed by at least four well-organized membrane protein complexes localized in the cell plasma or mitochondrial membrane (Figure 1.1). These protein complexes are NADH-ubiquinone oxidoreductase (complex I, CI), succinate-ubiquinone oxidoreductase (complex II, CII) ubiquinol cytochrome c oxidoreductase (complex III, CIII; also known as the cytochrome bc_1 complex), and cytochrome c oxidase (complex IV, CIV) [1]. Complex -I is the entry point for electrons from NADH, which is used to reduce ubiquinone (Q) to ubiquinol (QH$_2$). QH$_2$ is subsequently re-oxidized by complex III and electrons are transferred to reduce cytochrome c in the intermembrane space (IMS), while complex-IV uses cytochrome c to reduce the ultimate electron acceptor molecule oxygen to water. For each NADH molecule oxidized by complex I, III and IV, 10 protons are translocated across the membrane from the inner cell matrix to the IMS which generates a proton-motive force (PMF) across the membrane. Complex V utilizes this PMF to produce ATP from adenosine diphosphate (ADP) and inorganic phosphate. The conversion of redox energy into transmembrane electrical potential and proton gradients and finally into chemical energy stored in ATP is known as the

chemiosmosis, which was originally proposed by Peter Mitchell in 1961 [2]. The reactions catalyzed by the complexes are summarized in Figure 1.2
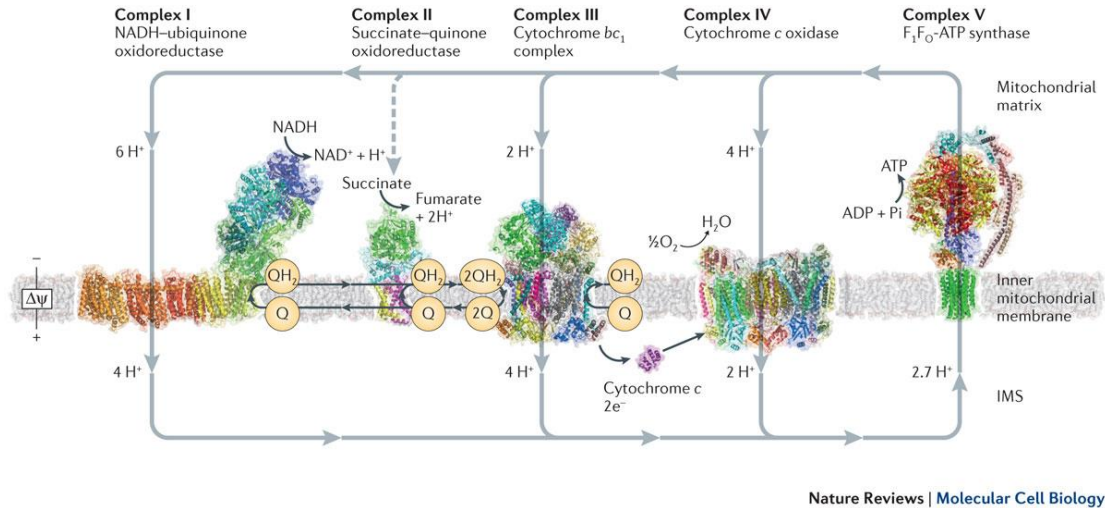
Figure 1.1. Electron transfer coupled to proton translocation in the mitochondrial electron transfer chain.

The ETC consists of four redox active membrane protein complexes NADH-ubiquinone oxidoreductase (complex I, CI), succinate-ubiquinone oxidoreductase (complex II, CII), ubiquinol cytochrome c oxidoreductase (complex III, CIII; also known as the cytochrome bc1 complex), and cytochrome c oxidase (complex IV, CIV). CI, CIII, and CIV are also proton-pumping enzymes, which generate proton motive force to drive F_1F_0-ATP synthase. Electron transport between complexes is mediated by membrane-embedded ubiquinone (Q) and soluble cytochrome c. Complex II (succinate–quinone oxidoreductase) provides an additional entry point for electrons into the chain. The three-dimensional (3D) structure of each respiratory complex presented here are complex I from *Thermus thermophilus* (protein databank identifier (PDB ID) 4HEA) [3], complex II from *Sus scrofa* (PDB ID 1ZOY) [4], complex III from *Bos taurus* (PDB ID 1BGY) [5] and complex IV from *B. taurus* (PDB ID 1OCC) [6]. The structure of F_1F_0-ATP synthase was generated by merging crystal structures of subcomplexes from the *B. taurus* enzyme within an 18 Å resolution cryo electron microscopy map [7]. $\Delta\Psi$, membrane potential. The figure was adapted from "*A giant molecular proton pump: structure and mechanism of respiratory complex I*" by Leonid A. Sazanov [8]. Copyright © 2018 Copyright Clearance Center, Inc. Adapted with permission.

Complex I: $NADH + H^+ + Q + 4H^+_{in} \rightarrow NAD^+ + QH_2 + 4H^+_{out}$

Complex III: $QH_2 + 2\ cyt\ c^{3+} + 2H^+_{in} \rightarrow Q + 2\ cyt\ c^{2+} + 4H^+_{out}$

Complex IV: $O_2 + 4\ cyt\ c^{2+} + 8H^+_{in} \rightarrow 2H_2O + 4\ cyt\ c^{3+} + 4H^+_{out}$
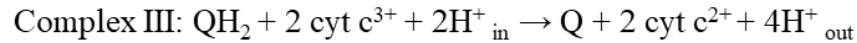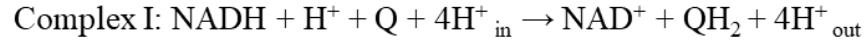
Figure 1.2. Catalytic reactions by respiratory complexes.

Q denotes ubiquinone and QH2 ubiquinol, cyt c denotes cytochrome c, and 'in' denotes the mitochondrial matrix and 'out' the IMS

1.2 NADH:ubiquinone oxidoreductase (complex I)

Complex I is the first enzyme of the ETC in bacteria and mitochondria. It is also one of the largest protein assemblies known in membrane. The enzyme couples the transfer of two electrons from NADH to ubiquinone via a series of enzyme-bound redox centers - Flavin mononucleotide (FMN) and iron-sulfur (Fe-S) clusters - with the translocation of four protons across the membrane. This process provides about 40% of the proton flux required for ATP synthesis and plays an essential role in cellular energy production. In addition, complex I is also the major contributor of reactive oxygen source (ROS) in mitochondria. Deficiencies of mitochondrial complex I can result in energy conversion decrease and ROS production increase, which affect many organs and systems of the body, particularly the nervous system, the heart, and skeletal muscles. Many diseases such as Leber's hereditary optic neuropathy, Parkinson's disease, dystonia, severe lactic acidosis, various encephalomyopathies, and possibly Huntington's disease are related to the dysfunction of complex I.

1.2.1 Subunit composition

Bacteria contain the minimal form of complex I, which is composed of 14 conserved core subunits with molecular mass around 550 kDa. In some bacteria, genes *nuo*C and *nuo*D are fused, which leads to the assembly of a complex consisting of 13 subunits [9]. In addition, extra subunits are discovered in bacteria such as *T. thermophilus* which are not conserved within the complex I family [10]. The subunits of bacterial

complex I are encoded by 14 corresponding genes (*nuo*A to *nuo*N) [11] located in one operon (Figure 1.3), an gene organization that is conserved in several other bacteria, including *Salmonella typhimurium* [12], *Paracoccus denitrificans* [13], *Rhodobacter capsulatus* [14], and *T. thermophilus* [15]. The core subunits can be classified into three distinct modules based on evolutionary relationships: the NADH-dehydrogenase N-module, the connecting Q-module and the pumping P-module [16]. The N-module is built up of the globular subunits NuoE, F, and G which forms the electron input part of the complex. NuoB, NuoC, NuoD, and NuoI constitute the connecting fragment. NuoA, NuoH, NuoJ, NuoK, NuoL, NuoM, and NuoN constitute the P-module. The mitochondrial complex I is even more complicated with a total mass about 1 MDa. In addition to the core subunits that are sufficient to catalyze energy transduction, there are also ~ 30 supernumerary (or accessory) subunits surrounding the core complex. Of the 44 subunits, 7 are encoded in mitochondrial DNA and 37 are encoded in nuclear DNA. Some supernumerary subunits are known to have specific roles, and as a cohort they have been proposed to have roles in functional regulation, protection, complex assembly and stability [17-22]. The composition and nomenclature of the core subunits is shown in Table 1.1. Hereinafter, *E. coli* nomenclature and bovine mitochondrial nomenclature will be used to represent the bacteria and mitochondrial complex I, respectively.
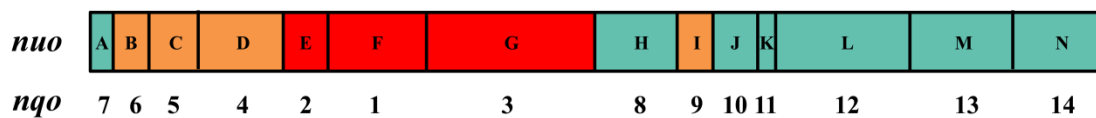
Figure 1.3. Map of *nuo* locus.

The genes encoding three modules of bacterial complex I were colored differently, N-module as red, Q-module as orange, and P-module as cyan.

Table 1.1. The nomenclature*of the core subunits

	Prokaryotes		Eukaryotes		
	Escherichia coli	*Thermus thermophilus*	*Bos Taurus*	*Homo sapiens*	*Yarrowia lipolytica*
Peripheral arm					
N-module (Dehydrogenase)	NuoF	Nqo1	51 kDa	NDUFV1	NUBM
	NuoE	Nqo2	24 kDa	NDUFV2	NUHM
	NuoG	Nqo3	75 kDa	NDUFS1	NUAM
Q-module (Connecting)	NuoD*	Nqo4	49 kDa	NDUFS2	NUCM
	NuoC*	Nqo5	30 kDa	NDUFS3	NUGM
	NuoI	Nqo9	TYKY	NDUFS8	NUIM
	NuoB	Nqo6	PSST	NDUFS7	NUKM
Membrane arm					
P-Module (Pumping)	NuoH	Nqo8	ND1	ND1	NU1M
	NuoA	Nqo7	ND3	ND3	NU3M
	NuoJ	Nqo10	ND6	ND6	NU6M
	NuoK	Nqo11	ND4L	ND4L	NULM
	NuoN	Nqo14	ND2	ND2	NU2M
	NuoM	Nqo13	ND4	ND4	NU4M
	NuoL	Nqo12	ND5	ND5	NU5M

* In some bacteria including *E. coli* NuoC and NuoD are fused.

Note: Nuo nomenclature originates from NADH-ubiquinone oxidoreductase, Nqo nomenclature originates from NADH-quinone oxidoreductase, NDU nomenclature is from NADH dehydrogenase-ubiquinone, and mtDNA-encoded subunits are given the prefix "ND".

1.2.2 Overall structure of complex I

Complex I was first isolated in 1961 [23] and in the following decades characterized concerning the molecular and cofactor composition and the catalytic reaction. Before the X-ray crystal structure of complex I was available, the structural information of complex I was mainly obtained via sequence analysis, electron paramagnetic resonance (EPR) measurement and electron microscopy (EM). The location of certain

Fe-S cluster binding motif was predicted by the primary sequence of the gene cluster of complex I and six Fe-S clusters including the binuclear clusters N1a and N1b, the tetranuclear clusters N2, N3, N4, N5 were identified by EPR analysis [24-31]. Low resolution structures acquired by using EM for the *Neurospora crassa* [32-34], bovine heart [35], *Y. lipolytica* [36] and *E. coli* [34] reveal an L-shaped architecture consisting of a hydrophilic and hydrophobic domain. The former protrudes into the mitochondrial matrix or the bacterial cytoplasm and the latter is embedded in the membrane. However, detailed structural information of complex I was absent before 2005. In 2006, the first crystal structure of the hydrophilic domain of complex I from *T. -thermophiles* was solved at 3.3 angstrom [10], The FMN moiety and nine Fe-S clusters were clearly shown in the structure. In 2010–2011, the structure of the membrane domain of *E. coli* complex I at 3.0 Å resolution [37] and the entire *T. -thermophilus* enzyme at 4.5 Å were determined [38]. And the first X-ray analysis of mitochondrial complex I on obligatory aerobic yeast (*Y. lipolytica*) at 6.3 Å resolution was also reported in 2010 [39]. Up to date, there are around 20 complex I structures reported (Table 1.2). The latest one was the cryo-EM structures of complex I from mouse heart [40]. The representative structures of bacterial and mitochondrial complex I are shown in Figure 1.4.

Due to the fact that the core of complex I is highly conserved among bacteria and mitochondria and the target of this project is bacterial source, the introduction hereafter will be mainly focused on the bacterial complex I, otherwise indicated unless.
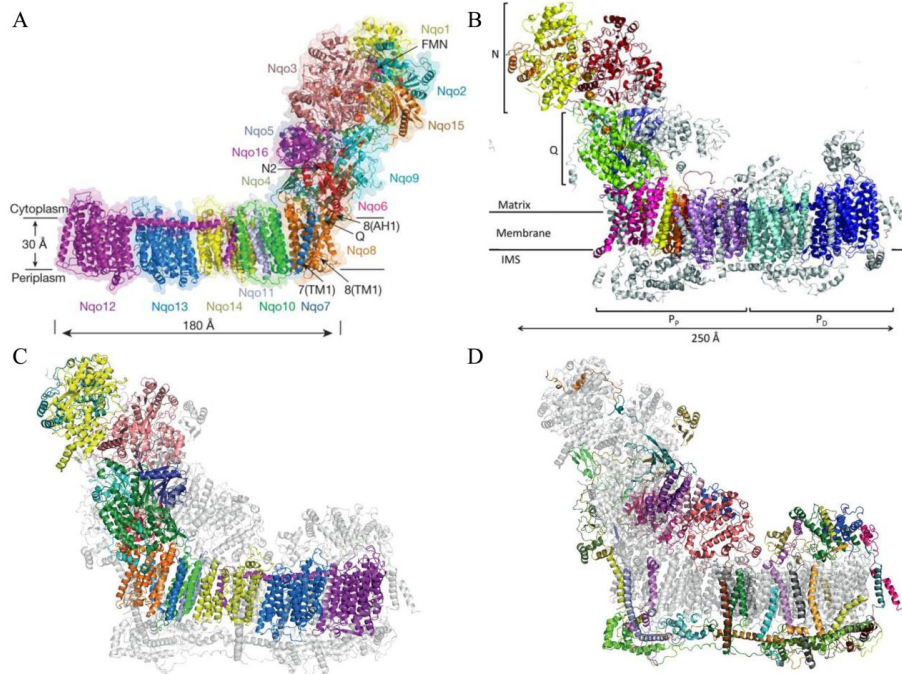
Figure 1.4. Overall structure of complex I.

(A) Crystal structure of entire complex I from *T. thermophiles* at 3.3 Å. The redox centers are shown as magenta and red-orange spheres, respectively, the FMN and N2 cluster were labelled. All the subunits depicted in different colors were labelled. (B) Crystal structure of mitochondrial complex I from *Y. lipolytica* at 3.6 to 3.9 Å. The subunits were coloured differently: Red, 75-kDa; yellow, 51-kDa; orange, 24-kDa; green, 49-kDa; violet, 30-kDa; blue, PSST; cyan, TYKY; dark blue, ND5; cyan, ND4; lilac, ND2; red, ND4L; orange, ND6; yellow, ND3; and pink, ND1. (C) Cryo-EM structure of ovine complex I. The core subunits were colored and supernumerary subunits in grey and transparent. (D) Cryo-EM structure of ovine complex I with core subunits in grey and supernumerary subunits in color. Figure 1.4A was adapted from "*Crystal structure of the entire respiratory complex I*" [3], Figure 1.4B from "*Mechanistic insight from the crystal structure of mitochondrial complex I*" [41], and Figure 1.4C, D from "*Atomic structure of the entire mammalian mitochondrial complex I*" [42]. Copyright © 2018 Copyright Clearance Center, Inc. Adapted with permission.

Table 1.2. List of complex I structure

	Structure	Method	Resolution (Å)	PDB	Reference
T. thermophilus	Peripheral arm	Crystallization	3.3	2FUG	[10]
T. thermophilus	Peripheral arm (oxidized)		3.1	3I9V, 3IAS	[43]
T. thermophilus	Peripheral arm (reduced)		3.1	3IAM	[43]
T. thermophilus	Entire		4.5	3M9S	[38]
T. thermophilus	Membrane arm		3.3	4HE8	[3]
T. thermophilus	Entire		3.3	4HEA	[3]
E. coli	Membrane arm		3.9	3M9C	[38]
E. coli	Membrane arm		3.0	3RKO	[37]
Y. lipolytica	Entire	Crystallization	6.3*	_*	[39]
Y. lipolytica	Entire		3.6	4WZ7	[41]
B. tuarus	Entire	Cryo-EM	4.95	4UQ8	[44]
B. tuarus	Entire		4.2	5LDX, 5LC5, 5LDW	[45]
B. tuarus	Entire		4.1	5O31	[46]
Ovine	Entire		3.9	5LNK	[42]
Homo sapiens	Metrix arm		3.4	5XTB	[47]
H. sapiens	Membrane arm		3.7	5XTC	[47]
Mouse	Entire		3.9	6G2J, 6G72	[40]

* X-ray analysis of the *Y. lipolytica* mitochondrial complex I, no subunit was identified and no models were deposited owing to limited resolution.

1.2.3 Hydrophilic (Peripheral) arm of complex I

The core of complex I peripheral arm is Y-shaped (Figure 1.5) and composed of two functional modules: the N-module and the Q-module. It contains the NADH binding site, the non-covalently bound FMN, and eight conserved redox-active Fe-S clusters (N1a, N1b, N2, N3, N4, N5, N6a, N6b) [10, 48, 49]. N1a, N1b are binuclear ([2Fe–2S]) Fe-S clusters, whilst N2, N3, N4, N5, N6a, N6b, and N7 are tetranuclear ([4Fe–4S]) Fe-S clusters. N7 is an additional, non-conserved Fe-S cluster found in some bacteria such as *E. coli* and *T. thermophiles* [31]. EPR spectroscopy has been the most informative

technique for the study of Fe-S clusters in complex I, which provides useful information on the spin-spin interactions between neighboring redox centers. However, not all of the known Fe-S clusters can be identified by EPR. To date, only five Fe-S clusters (N1b, N2, N3, N4, and N5) of isolated complex I from bovine heart mitochondria [49-51] and six of nine Fe-S clusters N1a, N1b, N2, N3, N4, and N7 of the *E.coli* complex I are unambiguously identified by EPR measurement [52-54]. Signals of N6a and N6b cluster were also reported recently [55-57]. The information of the redox centres in complex I was summarized in Table 1.3. The properties of the EPR spectra of from complex I Fe-S clusters were shown in Figure 1.6 and Table 1.4.
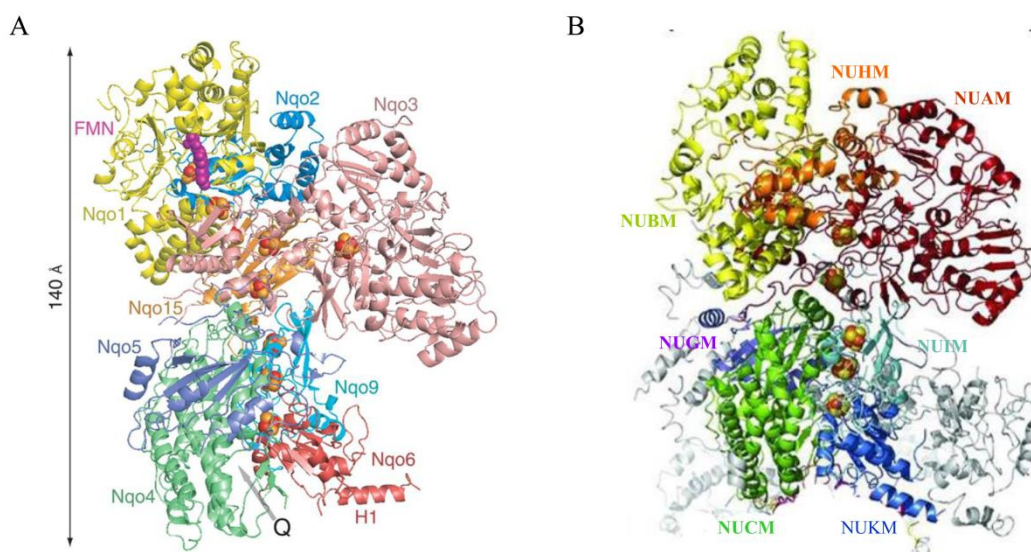
Figure 1.5. Architecture of the hydrophilic domain of complex I.

(A) The hydrophilic arm of *T. thermophilus* complex I. Eight hydrophilic subunits (Nqo1-6, Nqo9, Nqo15) are labelled and colored differently; FMN is shown as magenta spheres, Fe-S clusters as red spheres for Fe atoms and yellow spheres for S atoms. (B) The hydrophilic arm of mitochondrial complex I from *Y. lipolytica*. Figure 1.5A and 1.5B were adapted from "*Structure of the Hydrophilic Domain of Respiratory Complex I from Thermus thermophilus*" [10] and "*Mechanistic insight from the crystal structure of mitochondrial complex I*" [41], respectively. Copyright © 2018 Copyright Clearance Center, Inc. Adapted with permission.

Table 1.3. Substrates and redox centres of complex I

Redox centre	Location (*T. thermphilus*/*E. coli*/ *Bos Taurus*/ *Homo sapiens* / *Y. lipolytica*)	Redox Potential (mV)	Type
NADH	Nqo1/NuoF/75 kDa/NDUFS1/NUAM	-320	Substrate
FMN	Nqo1/NuoF/75 kDa/NDUFS1/NUAM	-340	
N3	Nqo1/NuoF/75 kDa/NDUFS1/NUAM	-250	[4Fe–4S]
N1b	Nqo3/NuoG/75 kDa/NDUFS1/NUAM		[2Fe–2S]
N4	Nqo3/NuoG/75 kDa/NDUFS1/NUAM	-250	[4Fe–4S]
N5	Nqo3/NuoG/75 kDa/NDUFS1/NUAM	-250	[4Fe–4S]
N6a	Nqo9/NuoI/TYKY/NDUFS8/NUIM	-250	[4Fe–4S]
N6b	Nqo9/NuoI/TYKY/NDUFS8/NUIM	-250	[4Fe–4S]
N2	Nqo6/NuoB/PSST/NDUFS7/NUKM	-100	[4Fe–4S]
N1a	Nqo2/NuoE/24 kDa/NDUFV2/NUHM	-370	[2Fe–2S]
N7	Nqo3/NuoG/75 kDa/NDUFS1/NUAM	-250	[4Fe–4S]
Ubiquinone		110	Substrate
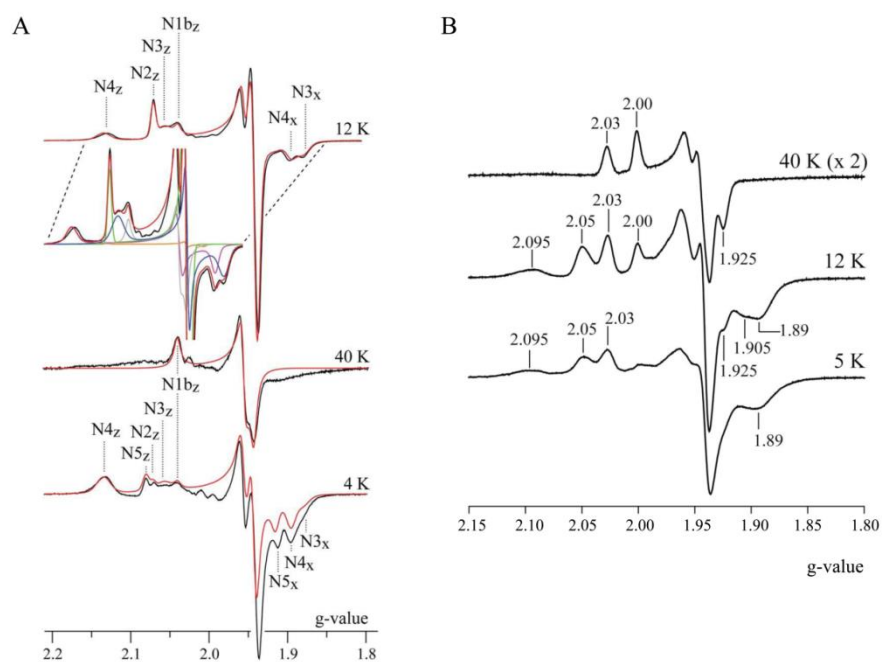
Figure 1.6. EPR spectra Fe-S clusters in complex I.

(A) EPR spectra of complex I from *B. taurus* mitochondria reduced by NADH. The spectra recorded at 12, 40 and 4 K (black) are compared with their simulated spectra (red). The inset for the 12 K spectrum (not on the x-axis scale) shows how the modelled 12 K spectrum comprises N1b (grey), N2 (green), N3 (blue), N4 (magenta) and N5 (orange). The g-values of the individual Fe-S clusters are $g_{z,y,x}$=2.024, 1.941, 1.926 for N1b, $g_{z,y,x}$=2.055, 1.926, 1.925 for N2, $g_{z,y,x}$=2.041, 1.927, 1.865 for N3, $g_{z,y,x}$=2.114, 1.930, 1.882 for N4 and $g_{z,y,x}$=2.064, 1.928, 1.898 for N5. (B) EPR spectra of isolated *E. coli* complex I. Complex I was dialyzed anaerobically against 20 mM Tris-HCl pH 7.5, 0.1 mM NADH for 1 h at 0 °C, then further reduced by 1 mM dithionite and frozen immediately. (Top) Spectrum at 40 K comprises two [2Fe-2S] clusters, N1b ($g_{z,y,x}$= 2.03, 1.94, 1.94) and N1a ($g_{z,y,x}$= 2.00, 1.95, 1.92). (Middle) Spectrum at 12 K comprises at least two [4Fe-4S] clusters, N4 ($g_{z,y,x}$=2.095 1.93 1.89) and N7 ($g_{z,y,x}$=2.048, ≈1.94, 1.916). (Bottom) Spectrum at 5 K comprises at least two [4Fe-4S] clusters (N4, N7) and one [2Fe-2S] cluster (N1b). Conditions were as follows: microwave power, 0.1 mW; conversion time, 81.92 ms; time constant, 20.48 ms; modulation amplitude 10 G; microwave frequency, 9.38 MHz. Figure adapted with permission from [50, 53]. Copyright © 2018 Copyright Clearance Center, Inc.

Table 1.4. The g values of signals of Fe-S clusters reported in the literature

Fe-S clusters	Species	g-values (g_z, g_x, g_y)	Reference
N3			
	E.coli	2.05, 1.94, 1.88	[52]
	B. taurus	2.041, 1.927, 1.865	[50, 51]
N1b			
	E.coli	2.03, 1.94, 1.94	[52]
	B. taurus	2.024, 1.941, 1.926	[50, 51]
N4			
	E.coli	2.09, 1.93, 1.89	[52]
	B. taurus	2.114, 1.930, 1.882	[50, 51]
N5			
	B. taurus	2.064, 1.928, 1.898	[50, 51]
N2			
	E.coli	2.05, 1.91, 1.91	[52]
	B. taurus	2.055, 1.926, 1.925	[50, 51]
N1a			
	E.coli	2.00, 1.95, 1.92	[52]

N7			
	E.coli	2.05, 1.94, 1.91	[53, 58]
N6a			
	E.coli	2.09, ?, 1.88	[55, 57]
N6b			
	E.coli	2.087, 1.905, 1.889	[56]

The hydrophilic domain contains all the redox centres and is responsible for the oxidation-reduction reactions. The electron donor NADH is embedded into its binding pocket in NuoF in an extended conformation with the nicotinamide ring positioned in front of the isoalloxazine ring of FMN for an effective hydride transfer [43]. The two electrons are delivered via seven conserved Fe-S clusters to reduce ubiquinone to ubiquinol. The main route for electron transfer within complex I presumably extends along chain NADH-FMN-N3-N1b-N4-N5-N6a-N6b-N2-quinone (Figure 1.7) [10]. N1a and N7 cluster don't belong to the main electron transfer chain. The distance between N7 and other Fe-S clusters is more than 20 Å, therefore, it is too far away from the main chain to participate in the electron transfer [10]. It was further demonstrated by mutational and EPR studies that N7 is essential for the stability of complex I, but not involved in the electron transfer [54]. Binuclear Fe–S cluster N1a is located on the opposite side of the flavin, distant from the other clusters, but close enough to the flavin for rapidly exchanging electrons. The function of cluster N1a is still unclear. However, it was proposed by Sazanov *et al* that N1a may play an important role in reducing ROS production by complex I [10, 59]. In addition, it was also pointed out according to mutational analysis that N1a might play a role for stability, for the assembly of the complex, or for the oxidation reaction of NADH due to its close proximity to the substrate binding site [60, 61].
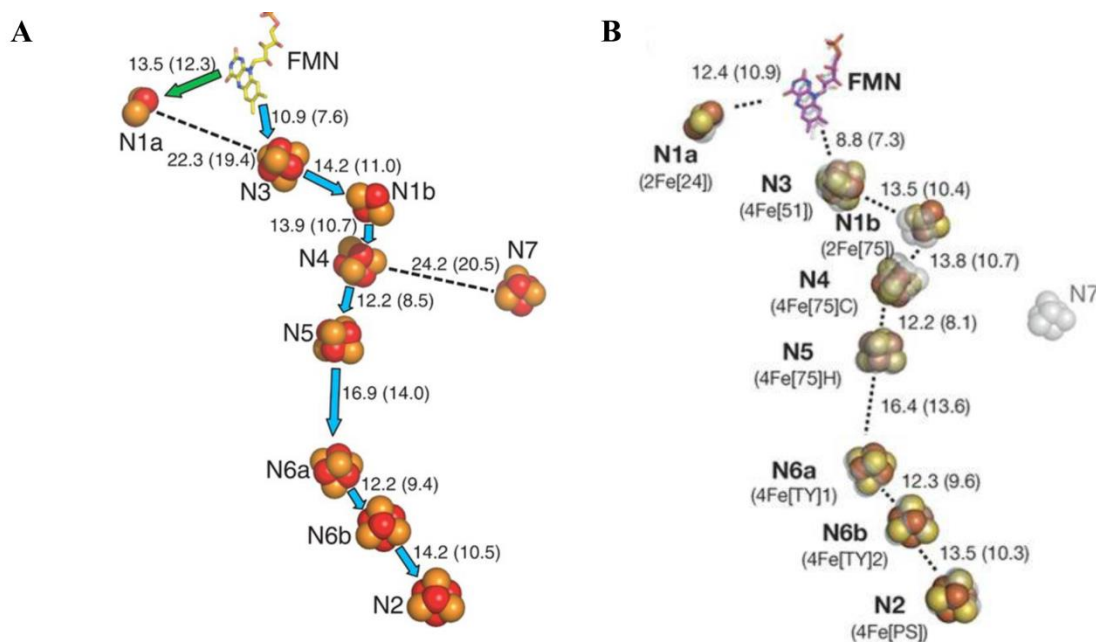
Figure 1.7. Arrangement of the redox centre of complex I.

Fe–S clusters are shown as spheres with centre-to-centre and edge-to edge (in brackets) distances indicated in Å. (A) Redox centres in *T. thermophilus* complex I. The main pathway of electron transfer is indicated by blue arrows, and a diversion to cluster N1a by a green arrow. (B) Redox centres in ovine complex I, overlaid with transparent grey depictions from *T. thermophilus*. Both the traditional and structure-based (in brackets) nomenclature for clusters is shown. Figure 1.3A and 1.3B were adapted from "*Structure of the Hydrophilic Domain of Respiratory Complex I from Thermus thermophilus*" [7], and "*Atomic structure of the entire mammalian mitochondrial complex I*", respectively. Copyright © 2018 Copyright Clearance Center, Inc. Adapted with permission.

1.2.4 Hydrophobic (membrane) domain of complex I

The core of complex I membrane arm comprises two modules, the proximal pump module (P_P) and the distal pump module (P_D), with a total of 64 transmembrane helices (Figure 1.8) [3, 41]. The P_P module contains subunits NuoH, NuoN, NuoA, NuoK, and NuoJ (known as ND1, ND2, ND3, ND4L, and ND6 in mitochondria, respectively), whereas the P_D module contains NuoM, and NuoL (ND4 and ND5). The P_P and P_D modules are bridged by a lateral helix (>60 Å long) of NuoL C-terminal extension, which lines NuoL, NuoM, and NuoN on the concave side of the arm close to the matrix side (Figure 1.8A). NuoL, M, and N, which are homologous to the Mrp Na^+/H^+ antiporter family (MrpA, MrpD and MrpD, respectively) [37, 39, 62], share a

structurally highly similar core of 14 TMHs with two repeats of five TMHs (1, TMH4–8; 2, TMH9–13) in inverted topology (Figure 1.9A). Each repeat features a discontinuous helix (TMH7 and TMH12) interrupted in the middle of the bilayer by an extended loop (5–7 residues) containing a proline conserved between all three antiporter-like subunits (Figure 1.9B). It was proposed that such helices are hallmarks of proton or ion translocation by introducing some flexibility and charge to the middle of the membrane [63-65]. NuoH is the most conserved membrane subunit of complex I, forming the interface with the hydrophilic domain and contributing to the Q site. The core fold of subunit NuoH (TM helices 2–6) is unexpectedly, similar to a half-channel of the antiporter-like subunits [3, 41]. The half-channel formed by NuoH is linked to the second half-channel formed by the adjacent subunits NuoA, K and J, NuoH (TM5) and NuoJ (TM3) with two discontinuous antiporter helices TM7 and TM12 (Figure 1.10). The antiporter-like subunits complete the four proton-translocation pathway.
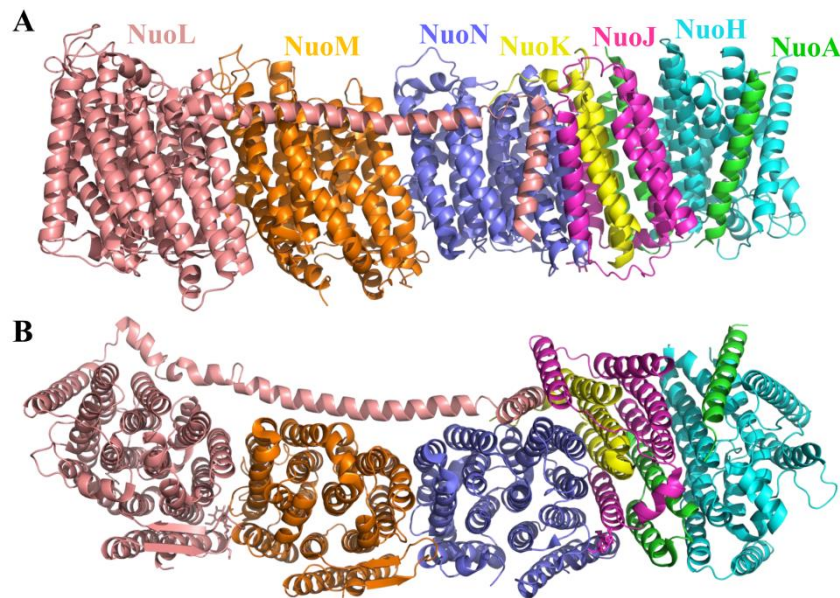
Figure 1.8. **Architecture of complex I membrane domain.**
Cartoon representation of the atomic model. (A) Side view in the membrane plane. (B) View from the periplasm into the membrane. Subunits are coloured as indicated by the labels, NuoA as green, NuoH as cyan, NuoJ as magenta, NuoK as yellow, Nuo N as blue, NuoM as orange, and NuoL as pink. The figure was made from the membrane domain of complex I from *T. thermophilus* (PDB code 4HE8) using Pymol.
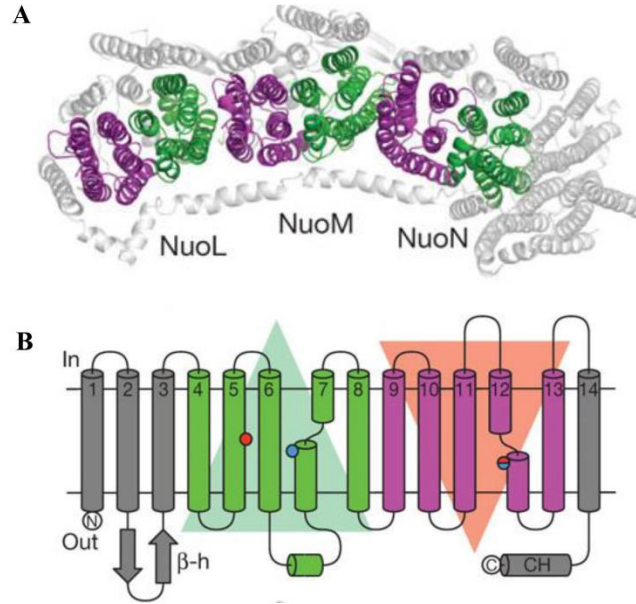
Figure 1.9. Fold of antiporter-like subunits NuoL, M and N.

(A) Structural repeats in the membrane domain, seen from the cytoplasm. TMs 4–8 are highlighted in green and TMs 9–13 in magenta. (B) Topology diagram of the antiporter-like subunits. Two inverted repeats of the conserved core are shown in green and magenta. The discontinuous helices (TMH7a/b, TMH12a/b) are hallmarks of ion-translocating membrane proteins [63-65]. The figure was adapted from "*Structure of the membrane domain of respiratory complex I*" [37]. Copyright © 2018 Copyright Clearance Center, Inc. Adapted with permission.
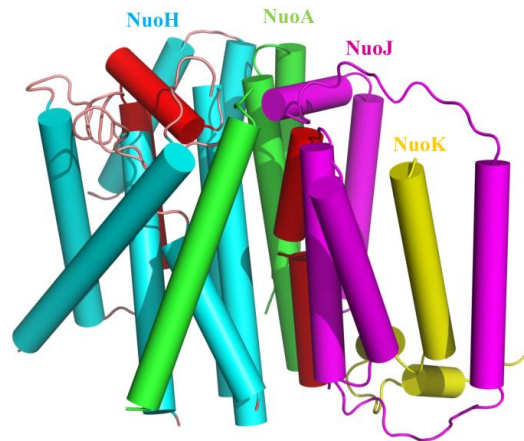
Figure 1.10. **Folds of subunits NuoH, A, J and K.**

The subunits were colored differently, NuoH as cyan, NuoA as green, NuoJ as magentas, and NuoK as yellow. The first half-channel is formed by NuoH and the second half-channel in small subunits NuoA,

NuoK and NuoJ. The discontinuous antiporter helices NuoH (TM5) and NuoJ (TM3) (in the roles of the key TM7 and TM12) were highlighted in red. The figure was made from the membrane domain of complex I from *T. thermophilus* (PDB code 4HE8) using Pymol.

As the substrate of complex I, the quinone analogs bind 15 Å away from the membrane surface, at the deep end of a long narrow cavity (30 Å long). In this position, the quinone headgroup is 12 Å (centre-to-centre) from the Fe–S cluster N2, which is appropriate for efficient electron transfer [3, 66]. The quinone binding site of bacterial complex I is shielded from the solvent, except for a narrow (approximately 2–3 × 4–5 Å) apparent entry point for the quinone, framed by helices TM1, TM6, the amphipathic AH1 from NuoH and TM1 from NuoA (Figure 1.11A). All residues facing the lipid bilayer are hydrophobic, but the interior of the chamber is unexpectedly lined by dominantly hydrophilic residues, especially in the area ('front') facing the tip of the membrane domain (to the left in Fig. 1.11B). However, the opposite side ('back') is coated by a hydrophobic patch mainly formed by residues from the hydrophilic subunits NuoCD and NuoB, which extends towards the entrance and is well designed to accommodate the quinone tail. The cavity front is mostly negatively charged, whereas the back is neutral and the 'top' (near cluster N2) is positively charged (Fig. 1.11B). The ionizable residues lining the chamber are highly conserved and their exchange by mutation frequently lead to a loss of complex I activity and to human diseases[59].
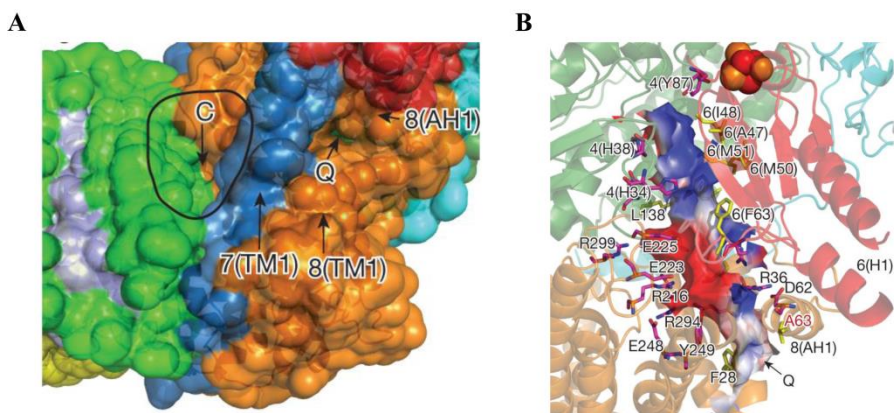
Figure 1.11. **Quinone-reaction chambers of *T. thermophilus* complex I.**
Fe–S cluster N2 is shown as red-orange spheres. (A) Surface (solvent-accessible) representation of the interface between the two domains. Nqo10 (NuoJ) is colored in green, Nqo7 (NuoA) in blue, Nqo8 (NuoH) in orange, and Nqo6 (NuoB) in red. The almost enclosed quinone binding site is built up by helices TM1, TM6 and amphipathic AH1 from Nqo8, as well as TM1 from Nuo7. The narrow entry point for the quinone is marked by a "Q". (B) Quinone-reaction chamber with its internal solvent-accessible wall colored in red for negative, white for neutral and blue for positive charges. Charged residues lining the cavity are shown with carbon in magenta and hydrophobic residues in yellow. Residues are labelled with a prefix indicating the subunit (omitted for Nqo8). Ala^{63}, the site of the primary Leber's hereditary optic neuropathy disease mutation [67], is labelled in red. The figure was adapted from "*Crystal structure of the entire respiratory complex I*" [3]. Copyright © 2018 Copyright Clearance Center, Inc. Adapted with permission.

1.2.5 The coupling mechanism

Various investigators proposed their own energy coupling mechanism of complex I [68-72]. The mechanism, presented here, based on a structural and evolutionary perspective is widely accepted in the field [3, 41, 72]. As shown detailed in Figure 1.12, ubiquinone reduction near cluster N2 triggers and drives proton pumping and structural rearrangements of the ubiquinone reduction site very likely play a key role in the coupling mechanism of complex I.
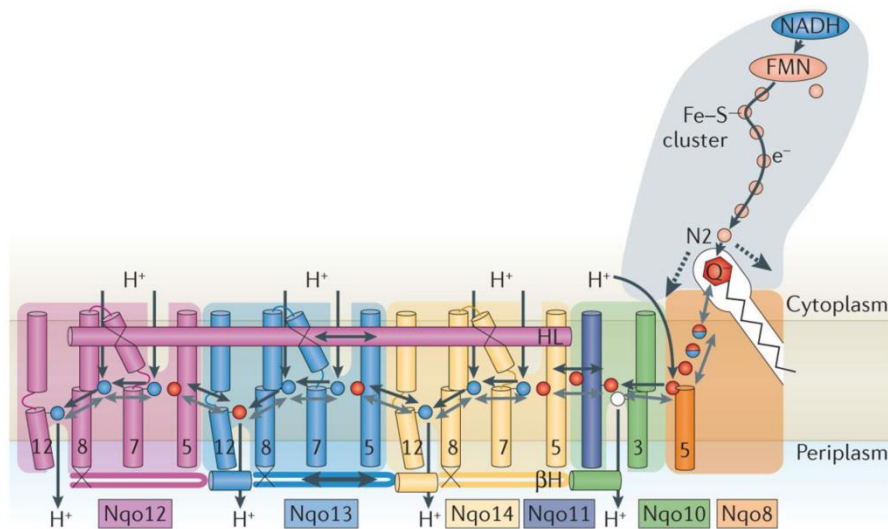
Figure 1.12. **Proposed coupling mechanism of complex I.**

Key helices and residues are depicted schematically. Upon electron transfer from the Fe–S cluster N2, negatively charged quinone (or charged residues nearby) initiates a cascade of conformational changes, propagating from the E-channel at Nqo8, Nqo10 and Nqo11 (NuoH, NuoJ and NuoK) to the antiporters via the central axis (indicated by grey arrows) comprising charged and polar residues that are located around flexible breaks in key transmembrane helices (TMHs). Cluster N2-driven shifts (dashed arrows) of Nqo4 (NuoD) and Nqo6 (NuoB) helices33 (not shown) are likely to assist overall conformational changes. Helix HL and the βH element help to coordinate conformational changes by linking discontinuous TMHs between the antiporters. Key charged residues can be protonated from the cytoplasm through several possible pathways, including inter-subunit transfer (indicated by black arrows). Following the reduction of quinone and completion of conformational changes, Lys or GluTM12 in the antiporters and Glu32 from Nqo11 (NuoK) in the E-channel each eject a proton into the periplasm. TMHs are numbered and key charged residues are indicated by red circles for Glu, blue circles for Lys or His, and white circle for Tyr. The involved charged residues are GluTM5, LysTM7, Lys or HisTM8 and Lys or GluTM12 from Nqo12–Nqo14 (NuoL, M and N), as well as Glu^{67} and Glu^{32} from Nqo11 (NuoK), which interacts with Tyr^{59} from Nqo10 (NuoJ), Glu^{213} from Nqo8 (NuoH) and some residues from the connection to the quinone cavity. The figure together with the figure legend was adapted with permission [8]. Copyright © 2018 Copyright Clearance Center, Inc.

1.2.6 Assembly of complex I

The biogenesis of complex I is multifaceted and sophisticated, involving the coordination of several crucial processes. The individual subunits and cofactors must be assembled correctly to executive function properly. Studies of the assembly process of complex I have been extremely challenging because of its enormous size, complexity, and lack of high resolution structure. Nevertheless, in recent years, extensive studies have provided new insights into the biosynthesis of mitochondrial complex I. The assembly of mitochondrial complex I is assisted by a number of specialized assembly factors [73], which are encoded by the nucleus and are absent in the final mature enzyme. Fourteen known and three putative assembly factors were detected by a dynamic complexome profiling approach in the latest study [74]. According to the current model of complex I assembly, six different modules (Figure 1.13A) assemble independently and associate afterwards with each other to form the final enzyme (Figure 1.13B) [74-77].
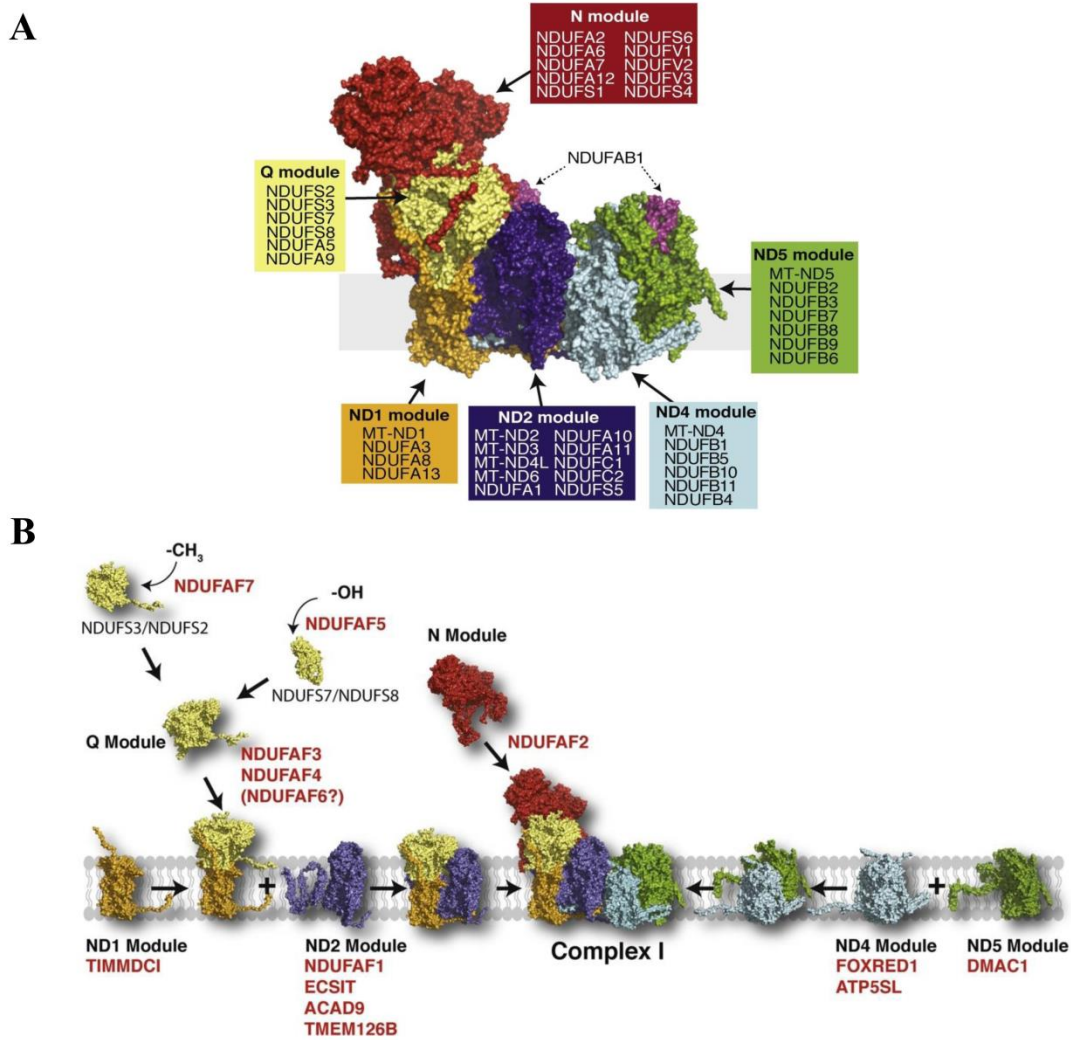
Figure 1.13. **Assembly of mitochondria complex I.**

(A) The modular composition of mitochondrial complex I. The different modules and their individual subunit composition are depicted using the structure of bovine heart complex I (PDB: 5LDW). (B) Assembly pathway of complex I. Complex I subunits assemble into discrete modules before joining to form the functional complex. The Q module integrates with the ND1 module and further associates with the ND2 module. The ND4 and ND5 modules associate together followed by integration with the Q/ND1/ND2 modules to form a late stage intermediate comprised of the Q-module and the completed membrane arm of complex I. The N-module then associates with this late stage intermediate to form the fully assembled complex I. The assembly factors associated with the biogenesis of each module are indicated. This Figure is adapted from "*Building a complex complex: Assembly of mitochondrial*

respiratory chain complex I" [76]. Copyright © 2018 Copyright Clearance Center, Inc. Adapted with permission.

Compared to the mitochondrial complex I, little is known about the assembly of the bacterial complex I. It was discovered in *E.coli* that NADH dehydrogenase fragment cannot be assembled by overexpression of *nuo*EFG, but can be obtained by overexpressing *nuo*BCDEFG in the cytoplasm [24]. Furthermore, the dehydrogenase fragment was enriched in the cytoplasm of single *nuo*-deletion mutants lacking either nuoA, CD, H, I, J, K, M or N, while the deletion of *nuo*L does not disturb the assembly of the residual complex [78]. In addition, the bioinformatics analysis of available bacterial genomes (~1,000) revealed that approximate half of the bacteria contain all 14 *nuo* genes. The *nuo* genes are organized as a polycistronic operon which is co-localized as *nuo*A to *nuo*N in 86% of these bacteria [79]. A possible assembly pathway of complex I for bacteria with the *nuo* gene co-localized as *nuo*A to *nuo*N was proposed by *Thorsten Friedrich et al.* (Figure 1.14) [80]. According to this model, NuoA was considered as an anchor protein which directs the *nuo*-mRNA to the Sec-translocon, thereby tethers the mRNA to the bacterial membrane and results in a translation of NuoB to NuoI close to the membrane. NuoA is involved in H-bonding with NuoB, which probably provides a platform together with NuoCD for further assembly of the dehydrogenase fragment. Then NuoH was added, which might tighten the interaction of the assembled peripheral arm by additional interactions to NuoB, CD and I. The connection between the peripheral and the membrane arm was further strengthened by acquisition of NuoJ and K. The membrane arm might be extended by another assembly intermediate which consists of NuoL, M and N. The proteins possibly involved in the maturation of *E. coli* complex I were listed in Table 1.5.
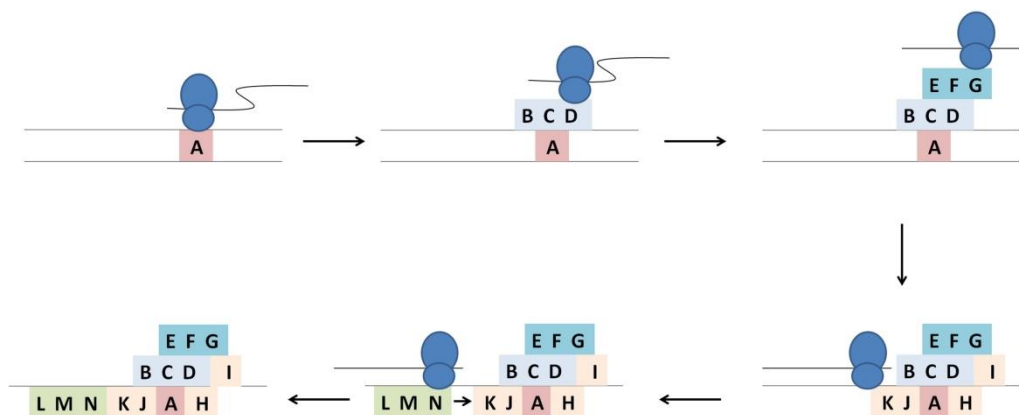
Figure 1. 14. **The proposed assembly pathway of bacterial complex I.**

The ribosome (not drawn in scale) is shown in blue; the mRNA is represented by the black line. The sequential addition of individual subunits is indicated by the letter of the respective *E. coli* complex I subunit. The figure was adapted with permission [80]. Copyright © 2018 Copyright Clearance Center, Inc.

Table 1.5. Probable proteins involved in assembly of *E. coli* complex I

Protein	Full name	Putative function in complex I assembly
NfuA	Fe/S biogenesis protein NfuA	Involvement in the maturation of complex I NuoG [81]
CyaY	Fe-S cluster assembly protein CyaY	Not essential to the assembly of complex I Transient interaction with complex I [82]
YajL	Protein/nucleic acid deglycase 3	Chaperonin activity; protect complex I against oxidative oxidative stress or keep NuoG in an assembly compotent state during insertion of Fe-S cluster [83]
LdcI/CadA	Lysine decarboxylase	Play a role in the insertion or the repair of N2 cluster of complex I together with RavA [78]
RavA	Regulatory ATPase variant A	Metal chelatase or chaperone-like protein [84, 85]
ViaA	VWA domain protein interacting with AAA ATPase	Stimulator of RavA ATPase activity [86]

1.3 *Aquifex aeolicus*

A. aeolicus is the model organism not only of the *Aquifex* genus but also of the *Aquificaceae* family and the order *Aquificales.* Isolated first in 1992 by R. Huber and K. O. Stetter [87], *A. aeolicus* is one of the most hyperthermophilic bacteria known. The maximum growth temperature is up to 95 ℃. As a microaerophilic obligate chemolithoautotroph, *A. aeolicus* does not grow on a number of organic substrates, including sugars, amino acids, yeast extract or meat extract. It is cultured at 85 ℃ under a $H_2/CO_2/O_2$ (79.5:19.5:1.0) atmosphere in a medium containing only simple inorganic components [87]. Albeit the oxygen amount is very low, it is sufficient to be used as an electron acceptor by the complex respiratory apparatus. Up to now, no alternative electron acceptor has been discovered. Phylogenetic analysis of 16S ribosomal RNA sequences indicated that *A. aeolicus* is one of the deepest-branching bacteria (Figure 1.15) [88-90].

A. aeolicus is the first hyperthermophilic eubacterium with known genome [91]. The length of the genomic DNA of *A. aeolicus* is 1,551,335 base pairs (bps), only one-third the size of the *E. coli* genome. The densely-packed genome exhibits striking features of the gene organization. Most genes are organized in polycistronic operons and many convergently transcribed genes overlap slightly. However, many genes grouped within operons in other organisms, such as genes for the biosynthesis of some amino acids (tryptophan or histidine), are dispersed throughout the *A. aeolicus* genome or appear in novel operons. Furthermore, the genes encoding subunits of the same enzyme are often separated on the chromosome. Additionally, there are often more than two genes encoding for one protein or subunit isoforms.
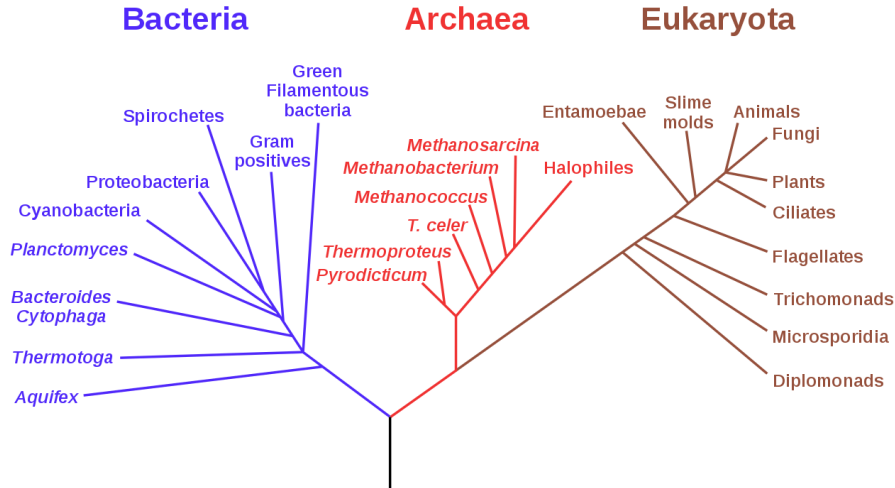
Figure 1.15. **Phylogenetic tree of life based on 16S RNA sequences, showing the three life domains: bacteria, archaea, and eukaryote.**

The black trunk at the bottom of the tree links the three branches of living organisms to the last universal common ancestor. The position of *Aquifex* is in the deepest known branch within the domain Bacteria (The figure was cited from NASA Astrobiology Institute)

The exceptional adaptation and metabolic capabilities as well as enzyme properties make *A. aeolicus* an excellent model system for studies of the respiratory pathway, protein structure and function, and the early evolution of metabolism.

1.4 Complex I from *A. aeolicus*

The analysis of the *A. aeolicus* genome reveals the presence of 24 *nuo* genes encoding subunits of complex I [91]. Seven genes are present in a duplicated form and two genes as triplicates. As found in *E. coli* and a few other bacteria, the genes *nuo*C and *nuo*D in *A. aeolicus* are fused to one gene $nuoD_2/nuoD_1$ [25, 91]. However, in contradiction to most of the other bacteria, these genes are not organized within one operon but are dispersed in different loci (Figure 1.16).
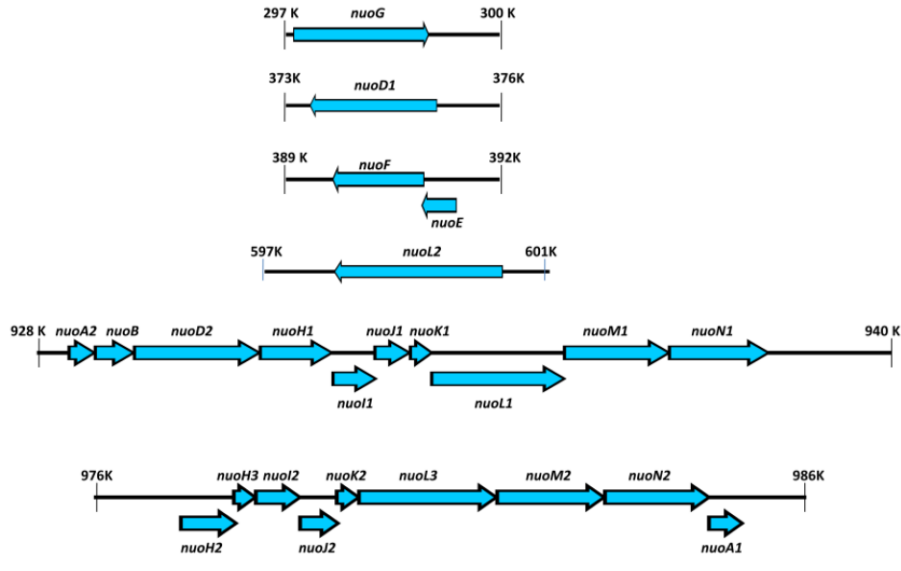
Figure 1.16. **Organization of *nuo* genes encoding complex I in the genome of *A. aeolicus*.**
Twenty four genes are separated into six DNA fragments located in both the plus and minus DNA strands.

A. aeolicus complex I was first purified and characterized as a highly stable and active complex in our lab by Peng *et al.* [92]. Seven subunits were identified and a typical L-shaped structure was observed by single particle cryo-EM. To date, 20 partially homologous subunits have been identified in complex I from *A. aeolicus* by combining MALDI-TOF and LILBID mass spectrometry methods [93]. These subunits could be assigned to two different isoforms, named NQOR1 and NQOR2, with molecular mass of 504.17 kDa and 523.99 kDa, respectively. NQOR1 consists of subunits $NuoA_2$, NuoB, $NuoD_2$, NuoE, NuoF, NuoG, $NuoI_1$, $NuoH_1$, $NuoJ_1$, $NuoK_1$, $NuoL_1$, $NuoM_1$ and $NuoN_1$; whereas NQOR2 comprises subunits $NuoA_1$, NuoB, $NuoD_1$, NuoE, NuoF, NuoG, $NuoH_2$, $NuoI_2$, $NuoJ_1$, $NuoK_1$, $NuoL_2$, $NuoM_2$ and $NuoN_2$ (Table 1.6). Sequence alignment shows that the sequence identity between the two homologous *A. aeolicus* complex I subunits is relatively high; normally $> 50\%$ with sequence coverages of $> 90\%$ (Table 1.7). The only exception is the NuoL homologues with an identity of only 36%, by sequence coverage of only 54%.

The crystal structure of the hydrophilic arm of NQOR1 has been determined at 2.9 Å, which is formed by NuoB, E, F, G, I_1 and D_2 (Figure 1.17A). The N-terminal helix of

NuoI$_1$ is unexpectedly inserted into the membrane, greatly contributing to the crystal packing. A novel tetranuclear Fe-S cluster N8 (between N4 and N7) is discovered in the structure (Figure 1.17B). It is generally believed that N7 is not involved in the electron transfer of complex I due to a long distance from the main chain (more than 20 Å). However, the N8 cluster is expected to connect N7 to N4 in the main electron transfer chain and may constitute an alternative electron transfer pathway. NuoG, NuoD$_2$ and NuoI$_1$ appear to provide an interface for another substrate binding. A small redox protein, alkyl hydroperoxide reductase C (AhpC2), was identified by MS in the purified complex I sample. It is a kind of thiol peroxidase that scavenges various peroxide substrates (hydrogen peroxide (H$_2$O$_2$), alkyl hydroperoxides, peroxinitrite etc.) and protects cells from oxidative damage. It is well to be reminded that complex I is considered to be one of the major sources of reactive oxygen species (ROS). On the basis of the information above, we proposed that AhpC2 is a potential substrate or binding partner of *A. aeolicus* complex I, which may be involved in an additional electron transfer pathway under oxidative stress.

Table 1.6. Subunits composition of *A. aeolicus* complex I

	NQOR1	NQOR2
Subunits composition	NuoA$_2$, B, D$_2$, E, F, G, H$_1$, I$_1$, J$_1$, K$_1$, L$_1$, M$_1$, and N$_1$	NuoA$_1$, B, D$_1$, E, F, G, H$_2$, I$_2$, J$_1$, K$_1$, L$_2$, M$_2$, and N$_2$
Molecular weight (kDa)	504	520

Table 1.7. Sequence alignments of isoform subunits of *A. Aeolicus* complex I

Subunits	Amino acid	Query cover (%)	Identity (%)	E value
NuoA$_1$/A$_2$	126/118	92	56	6e-40
NuoD$_1$/D$_2$	593/586	100	58	0.0
NuoH$_1$/H$_2$	336/259	75	57	3e-100
NuoI$_1$/I$_2$	201/208	96	55	3e-76
NuoL$_1$/L$_2$	622/787	54	36	2e-43
NuoM$_1$/M$_2$	491/501	96	40	1e-103
NuoN$_1$/N$_2$	464/488	96	35	1e-84
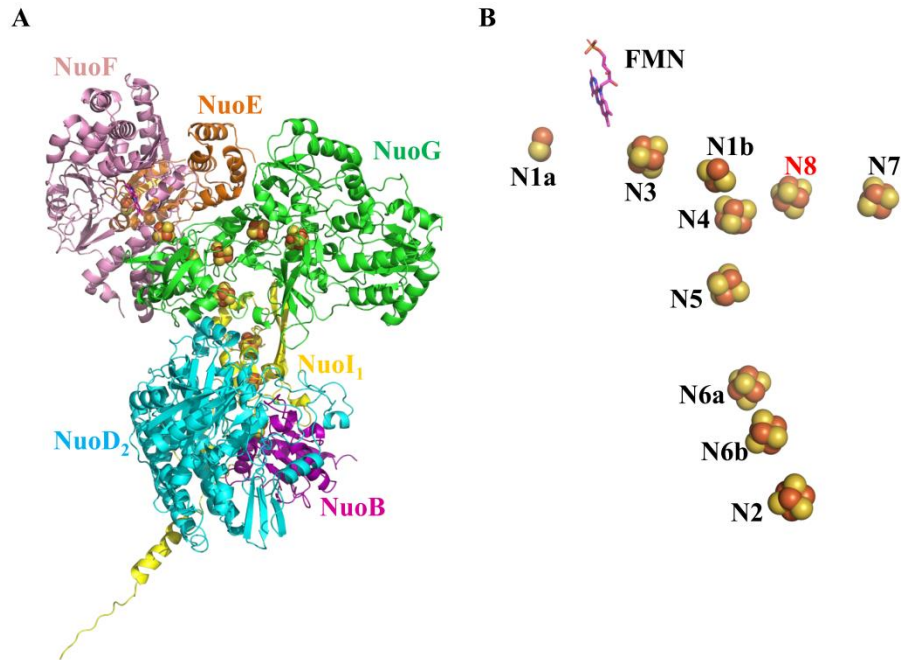
Figure 1.17. **Structure of hydrophilic arm of *A. aeolicus* complex I.**

(A) Side view. Each subunit is colored as its label (NuoB as magenta, NuoD$_2$ cyan, NuoI$_1$ as yellow, NuoE as orange, NuoF as purple, and NuoG as green); FMN is shown as purple sticks, metal sites as red spheres for Fe atoms and yellow spheres for S atoms. (B) Arrangement of Fe-S clusters. An additional Fe-S cluster N8 (between N4 and N7) was discovered and highlighted in red. The structure was solved by Peng *et al.* (Data unpublished)

1.5 Open questions

Compared to complex I from different organisms, complex I from the hyperthermophilic bacterium *A. aeolicus* exhibits some unique and interesting features. And thus a series of questions were raised.

(1) Two isoforms of complex I were assigned sample preparation from the native source. Why are there two isoforms present in *A. aeolicus*? What is the role of the individual isoform? What are the structural and functional variations between them? How does the cell regulate their expression to exploit such variations?

(2) Complex I from *A. aeolicus* carries a novel N8 cluster which may be involved in an alternative electron transfer pathway. What are its electron donor and acceptor and how does this pathway operate?

(3) A redox protein AhpC2 was detected in the purified complex I sample. What is the relationship between AhpC2 and complex I? Is it a substrate of *A. aeolicus* complex I as we postulated?

1.6 Aim of this work

Characterization of native *A. aeolicus* complex I had been performed by *Peng et al*. (See chapter 1.4). However, the lack of a genetic system hampered the more comprehensive studies of the complex I from *A. aeolicus*. To date, genetic manipulation in *A. aeolicus* is impossible. The cells have to be cultivated at 85 ℃ with a gas mixture of oxygen and hydrogen. Heterologous production of *A. aeolicus* complex I is also a challenging task. There are 13 genes encoding complex I, which are dispersed in different operons in the genomic DNA. The size of the complex I is large (> 500 kDa) and it harbors up to ten Fe-S clusters. No entire *A. aeolicus* complex I has been heterologously produced so far, only the expression of a subcomplex composed of NuoE and NuoF was reported [1]. Despite of these difficulties, our primary goal is to construct a heterologous genetic system for *A. aeolicus* complex I. Then with this platform, we aimed to (1) to separately prepare NQOR1 and NQOR2 for individual structural and functional studies; (2) to characterize the Fe-S clusters by EPR measurements with the special focus on the spectral properties of the novel N8 cluster; (3) to investigate the interactions between AhpC2 and complex I.

2. Materials and methods

2.1 Materials

2.1.1 Chemicals

The chemicals used in this work were obtained from Carl Roth GmbH (Karlsruhe Germany), Sigma-Aldrich (Taufkirchen Germany), New England Biolabs (Ipswich, USA), Thermo Scientific (Bonn, Germany), and Invitrogen (Carlsbad, USA) unless stated otherwise.

2.1.2 Primers

Primers were obtained from Eurofins MWG operon (Ebersberg, Germany) and Sigma-Aldrich (Taufkirchen, Germany). DNA Sequencing was performed by Eurofins MWG Operon (Ebersberg, Germany) and Seqlab (Göttingen, Germany).

2.1.3 Bacterial strains

A. aeolicus VF5 cells were purchased from the Archaeenzentrum (Regensburg, Germany) and stored at -20 °C before use.

Escherichia coli strains used in this work are listed in Table 2.1

Table 2.1. *E. coli* **strains used in the work**

Strains	Genotype	Reference
DH5α	F⁻*endA1 glnV44 thi1 recA1 relA1 gyrA96 deoR nupG purB20* φ80d*lacZ* ΔM15 Δ(*lacZYA-argF*) U169, hsdR17($r_K^- m_K^+$), λ⁻	[92]
C43(DE3)	*F⁻ ompT gal dcm hsdS$_B$ ($r_B^- m_B^-$) (DE3)*	[93]
TOP10	*F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ⁻*	Invitrogen
NM554	*F– araD139 Δ(ara-leu)7696 galE15 galK16 Δ(lac)X74 rpsL (StrR) hsdR2 (rK– mK+) mcrA mcrB1 recA13*	[94]

BA14 (Δ *nuo*)	*glnV44(AS) rfbC1 endA1 spoT1 thi-1*		[95]

BA14 is a NADH:Ubiquinone oxidoreductase deletion strain derived from *E. coli* 1100 strain. It was kindly provided by Professor Steven B. Vik from Department of Biological Sciences, Southern Methodist University, Dallas, TX 75275-0376 USA.

2.1.4 Plasmids

All empty vectors used in this work are listed in Table 2.2.

Table 2.2. Empty vectors used in the work

Plasmid	Feature	Application	Supplier
pJET1.2/blunt	pMB1 origin, T7 promoter, Amp^R	Cloning	Thermo Scientific™
pET30a	pBR322 origin, T7 promoter, Kan^R, N-terminal S-tag and His$_6$ tag, C-terminal His$_6$ tag	Expression	Novagen
pET32b	pBR322 origin, T7 promoter, Amp^R, N-Trx tag, S-tag and His$_6$ tag, C- terminal His$_6$	Expression	Novagen
pBAD-A2	pBR322 origin, *ara*BAD promoter, Amp^R, C-terminal His$_{10}$-tag	Expression	[96]
pBAD-C3	pBR322 origin, *ara*BAD promoter, Amp^R, N-terminal His$_{10}$-tag, C-terminal strepII tag	Expression	[96]
pBAD-CM1	pBR322 origin, *ara*BAD promoter, Amp^R, C-terminal His$_6$ and C-terminal strepII tag	Expression	[96]
pTTQ18-A2	pMB1 origin, Tac promoter, Amp^R, C-terminal His$_{10}$-tag	Expression	[96]
pTTQ18-C3	pMB1 origin, Tac promoter, Amp^R, N-terminal His$_{10}$-tag, C-terminal strepII tag	Expression	[96]

pQE-A2	pColE1 origin, T5$lac2^{\mathrm{f}}$ promoter, Amp$^{\mathrm{R}}$, C-terminal His$_{10}$-tag	Expression	[96]
pQE-C3	pColE1 origin, T5$lac2^{\mathrm{f}}$ promoter, Amp$^{\mathrm{R}}$, N-terminal His$_{10}$-tag, C-terminal strepII tag	Expression	[96]
pBAD33	pACYC184/p15A ori, *ara*C promoter, Cam$^{\mathrm{R}}$	Expression	[97]

All expression vectors constructed in this work are listed in Table 2.3

Table 2.3. Lists of expression vectors generated in this work

Name	Parent vector	Target gene	Tag
pET32b-*nuo*B	pET32b	*nuo*B	N-Trx, N-His$_6$
pET32b-*nuo*D$_2$	pET32b	*nuo*D$_2$	N-Trx, N-His$_6$
pET32b-*nuo*I$_1$	pET32b	*nuo*I$_1$	N-Trx, N-His$_6$
pET32b-*nuo*G	pET32b	*nuo*G	N-Trx, N-His$_6$
pET32b-*nuo*D$_1$	pET32b	*nuo*D$_1$	N-Trx, N-His$_6$
pET32b-*nuo*I$_2$	pET32b	*nuo*I$_2$	N-Trx, N-His$_6$
pBAD-A2-*nuo*I$_2$	pBAD-A2	*nuo*I$_2$	N-His$_{10}$
pBAD-C3-*nuo*I$_2$	pBAD-C3	*nuo*I$_2$	C-His$_{10}$
pTTQ18-A2-*nuo*I$_2$	pTTQ18-A2	*nuo*I$_2$	N-His$_{10}$
pTTQ18-C3-*nuo*I$_2$	pTTQ18-C3	*nuo*I$_2$	C-His$_{10}$
pQE-A2-*nuo*I$_2$	pQE-A2	*nuo*I$_2$	N-His$_{10}$
pQE-C3-*nuo*I$_2$	pQE-C3	*nuo*I$_2$	C-His$_{10}$
pBAD-CM$_1$-*nuo*EFG	pBAD-CM$_1$	*nuo*EFG	*nuo*G C-Strep II
pBAD-CM$_1$-*nuo*BD$_2$I$_1$-*nuo*EFG	pBAD-CM$_1$	*nuo*BD$_2$I$_1$-*nuo*EFG	*nuo*G C-Strep II
pBAD-CM$_1$-*nuo*BD$_1$-*nuo*EFG	pBAD-CM$_1$	*nuo*BD$_1$-*nuo*EFG	*nuo*G C-Strep II
pBAD-A2-BD$_2$EFGI$_1$	pBAD-A2	*nuo*BD$_2$EFGI$_1$	*nuo*F C-Strep
pBAD33-R1	pBAD33	*nuo*A$_2$BD$_2$H$_1$I$_1$J$_1$K$_1$L$_1$M$_1$N$_1$	No tag

pBAD33-R2	pBAD33	*nuo*$H_2I_2J_2K_2L_3M_2N_2A_1$	No tag

2.1.5 Medium, buffer and solution

The bacterial media used in this work (Table 2.4) were prepared in deionized water and autoclaved at 121 °C for 20 mins. To prepare agar plates, 1.5% (w/v) agarose was added to the liquid medium before autoclaving. Medium supplements (Table 2.5) were sterilized by filtration through filters of 0.2 μm membrane (Sarstedt Aktiengesellschaft & Co.). Antibiotics were added to the cooled medium (around 50 °C) after autoclaving to prepare agar plates supplemented with antibiotics. The stock solutions of antibiotics were stored at -20°C. The solutions of supplements for complex I expression were prepared just before use.

All buffers and solutions used in this work are prepared in Millipore water and are shown within the texts of the corresponding sections.

Table 2.4. Bacterial media used in the work

Medium	Composition
Luria-Bertani (LB) broth medium	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl
Terrific Broth (TB) medium	1.2% (w/v) tryptone, 2.4% (w/v) yeast extract,
	0.4% (v/v) glycerol, 0.17 M KH$_2$PO4, 0.72 M K$_2$HPO4
SOC medium	2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl,
	2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4, 20 mM glucose

Table 2.5. Supplements of Media

Additives	Stock solution	Working concentration
Antibiotics		
Carbenicillin (Carb)	100 mg/ml	100 μg/ml
Chloramphenicol (Cam)[a]	34 mg/ml	34 μg/ml
Kanamycin (Kan)	50 mg/ml	50 μg/ml
Supplements for expression of complex I		

Ferric ammonium citrate	25% (w/v)	0.1 g/L
Ferrous sulfate	20% (w/v)	0.1 g/L
Sodium sulfide	5% (w/v)	50 μM
L-Cysteine	Powder	0.12 g/L
Riboflavin (Vitamin B2)	Powder	20 mg/ml
Inducer		
IPTG	1 M	0.3 mM
L-arabinose	20 %	0.05%

[a] Cam is dissolved in pure ethanol

2.1.6 Enzymes, proteins, inhibitors and kits

The enzymes, proteins, inhibitors, markers and kits used in this work are summarized in Table 2.6.

Table 2.6. List of enzymes, proteins, inhibitors, markers and kits

Name	Application	Supplier
Phusion High-Fidelity DNA Polymerase	Polymerase chain reaction (PCR)	Thermo Scientific
Taq DNA Polymerase	Colony PCR	NEB
CloneJET^{TM} PCR Cloning Kit	Molecular cloning	Thermo Scientific
Restriction enzyme	DNA ligation	Thermo Scientific
Quick Ligation^{TM} Kit	DNA modification	NEB
In-Fusion® HD Cloning Kit	DNA ligation	Takara Bio USA
QuikChange Lightning Site-Directed Mutagenesis Kit	Mutation	Agilent Technologies
QIAprep® Spin Miniprep Kit	Plasmid extraction	Qiagen
Zymoclean^{TM} Gel DNA Recovery Kit	DNA purification	ZYMO Research
The E.Z.N.A.® Bacterial DNA Kit	DNA extraction	Omega Bio-tek
1 kb DNA Ladder	DNA standard	NEB
100 bp DNA Ladder	DNA standard	NEB

cOmplete™ Protease Inhibitor Cocktail	Protease inhibitor	Roche
SIGMAFAST™ Protease Inhibitor Tablets	Protease inhibitor	Merck KGaA
Pierce™ BCA Protein Assay Kit	Protein concentration	Thermo Scientific
Albumin fraction V	Western blot	Carl Roth
Albumin fraction V (biotin-free)	Western blot	Carl Roth
Avidin (egg white)	Avidin-biotin interaction	Gerbu
NativeMark™ Protein Standard	Protein standard	Invitrogen
PageRuler™ Prestained Protein Ladder	Protein standard	Thermo Scientific
SilverQuest™ Staining Kit	Protein staining	Invitrogen
SERVAGel™ IEF Starter Kit	Electrophoresis	SERVA Electrophoresis GmbH
Gel Filtration Calibration Kits (HMW)	Protein standards	GE Healthcare
The JCSG Core Suite I	Crystallization screen	Qiagen
The JCSG Core Suite II	Crystallization screen	Qiagen
The JCSG Core Suite III	Crystallization screen	Qiagen
The JCSG Core Suite IV	Crystallization screen	Qiagen
MemGold CF	Crystallization screen	Molecular Dimension
The MbClass II Suite	Crystallization screen	Jena Bioscience
The MbClass Suite	Crystallization screen	Jena Bioscience
JBScreen HTS I	Crystallization screen	Jena Bioscience
JBScreen HTS II	Crystallization screen	Jena Bioscience
JBScreen Pentaerythritol HTS	Crystallization screen	Jena Bioscience
MPI pH Screen	Crystallization screen	MPIBP-MMB

2.1.7 Antibodies

The antibodies used for western blot (WB) and immunoprecipitation (IP) in this work are listed in Table 2.7

Table 2.7. List of Antibodies used in this work

Type	Name	Target	Supplier
Primary Antibodies	Monoclonal Anti-polyHistidine-Alkaline Phosphatase antibody produced in mouse	His-tagged proteins	Sigma-Aldrich
	Alkaline phosphatase conjugated streptavidin	Strep II-tagged proteins	BIO-RAD
	Anti-NuoA$_2$	NuoA$_2$	Thermo Fisher
	Anti-NuoB	NuoB	
	Anti-NuoD$_2$	NuoD$_2$	
	Anti-NuoE	NuoE	
	Anti-NuoF	NuoF	
	Anti-NuoG	NuoG	
	Anti-NuoH$_1$	NuoH$_1$	
	Anti-NuoI$_1$	NuoI$_1$	
	Anti-NuoJ$_1$	NuoJ$_1$	
	Anti-NuoK$_1$	Subunit NuoK$_1$	
	Anti-NuoL$_1$	NuoL$_1$	
	Anti-NuoM$_1$	NuoM$_1$	
	Anti-NuoN$_1$	NuoN$_1$	
Secondary Antibodies	Monoclonal mouse anti-rabbit IgG conjugated with alkaline phosphatase	IgG from rabbit	Sigma-Aldrich

2.1.8 Chromatography columns and matrices

The columns and matrices used for protein purification and sample analyzation are listed in Table 2.8

Table 2.8. Columns and matrices used for protein purification

Column name	Type	Manufacture
Ni-NTA agarose	IMAC affinity chromatography	Qiagen
HisTrap HP 1 ml	IMAC affinity chromatography	GE Healthcare
Strep-Tactin® Superflow® high capacity 50% suspension	IMAC affinity chromatography	IBA GmbH
PD-10 Desalting Columns	Size-exclusion	GE Healthcare
Zeba™ Spin Desalting Columns	Size-exclusion	Thermo Fisher Scientific
Mono Q 10/100 GL	Anion ion exchange	GE Healthcare
Superdex 200 10/300 GL	Size-exclusion	GE Healthcare
TSK-GEL G4000SW	Size-exclusion	TOSOH Bioscience

2.1.9 Database and software

Table 2.9. List of database and web servers used in the works

Name	URL
EMBL	http://www.embl.org/
ProtParam	https://web.expasy.org/protparam/
NCBI: National Center for Biotechnology Information	http://www.ncbi.nlm.nih.gov/
PDB: Protein Data Bank	https://www.rcsb.org/
UniProt: Universal Protein Resource	http://www.uniprot.org/
Web of Knowledge	http://www.webofknowledge.com/
EMBL-EBI	https://www.ebi.ac.uk/
ExPASy: SIB Bioinformatics Resource Portal	http://www.expasy.org/
PubMed	http://www.ncbi.nlm.nih.gov/pubmed
Clustal Omega	https://www.ebi.ac.uk/Tools/msa/clustalo/
ESPript 3.0	http://espript.ibcp.fr/ESPript/ESPript/
Dali server	http://ekhidna2.biocenter.helsinki.fi/dali/
PDBePISA	http://www.ebi.ac.uk/pdbe/pisa/
PDBsum	http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html

The software used in this work is listed in Table 2.10.

Table 2.10 List of software

Software	version	Application
Clone Manager	Professional 9	Cloning simulation
Endnote	X7	Reference management
Microsoft Office	2010	Microsoft office software
Origin	9.0	Data processing and analysis
Photoshop	CS5 extended	Image editor
PyMOL	2.07	3D molecular visualizer
Unicorn	5.11	Äkta control system
Phenix	1.9-1692	Collection of crystallography programs
REFMAC	5.0	Structure refinement
CCP4	6.4.0	Collection of crystallography programs
XDS	3.1.15	Processing of diffraction images
XSCALE	3.1.15	Scaling and merging of X-ray datasets
Coot	0.72	Model building

2.2 Methods

2.2.1 Molecular biological method

2.2.1.1. Isolation of genomic DNA from *A. aeolicus*

For gene cloning, *A. aeolicus* genomic DNA was extracted from fresh cells of *A. aeolicus* by using the E.Z.N.A.® Bacterial DNA Kit (Omega Bio-tek, Inc) according to the manufacturer's protocol for Gram-negative bacteria. The genomic DNA was stored in TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) at -20 ℃ for long-term storage.

2.2.1.2 Isolation of plasmid DNA

A single bacterial colony was picked up from the plates and inoculated 4ml LB media containing appropriate antibiotics. After incubation at 37 ℃ (*E. coli*) with shaking at 220 rpm overnight, plasmid DNA was isolated using the QIAprep® Spin Miniprep Kit (Qiagen) according to the manufacturer's protocols. The plasmid DNA was stored in ddH$_2$O at -20 ℃ for long-term storage.

2.2.1.3 DNA amplification

The DNA sequences encoding for each subunit of complex I were obtained from NCBI (NC_000918) and used for primer design. Primers were custom synthesized by Sigma-Aldrich and are listed in Table 2.11. The stock and working concentrations of primer solutions were 100 μM and 10 μM, respectively. The target genes were amplified from *A. aeolicus* genomic DNA by PCR using the Phusion High-Fidelity DNA Polymerase (Thermo Scientific) in a T-Gradient themocycler (Biometra). Gradient PCR was performed to determine the optimal annealing temperature. The volumes of 10 μl and 50-μl were applied for analytical and preparative PCR cycles, respectively. The PCR reactions were set up on ice. To avoid degradation of primers caused by 3'→5' exonuclease activity, Phusion DNA Polymerase is the last component added to the PCR mixture. As a negative control, all components except the DNA template were used. The components of a typical reaction are listed in Table 2.12 and the program listed in Table 2.13.

Table 2.11. Primers used in this work

	Sequence (5'-3')	Amplification / vector
p1f	GAGCTCCATGGTTGCAATAAATTCTAACGGTT	*nuo*B / pET32b
p1r	AAGCTTCACCTTTAGTTCTCTGGGA	
p2f	GGATCCGATGAAATGGGTCAATAAGGGAAC	*nuo*D$_2$ / pET32b
p2r	GTCGACCCTATCCGTCTCACCAACAACC	
p3f	GAATTCGATGTCCGAAAAGGTGAAGATTTACA	*nuo*G / pET32b
p3r	AAGCTTAACACTCCTTTCGTAGTAAACCG	
p4f	GAATTCGATGGGTGTGAAGAAGTTAAGCAGGA	*nuo*I$_1$ / pET32b
p4r	AAGCTTACCACTCCACTCTATCTTCCCC	
p5f	GGATCCGATGCCGTGGGCTAAGGAAGGAG	*nuo*D$_1$ / pET32b
p5r	GTCGACTCTGTCCGTTTCTCCCAC	
p6f	GAATTCGATGATTAAGAAGGTAGCTGCAAAGCC	*nuo*I$_2$ / pET32b
p6r	AAGCTTCACCTCCTCGGGTTTAG	
p7f	GGATCCTCATGATTAAGAAGGTAGCTGCAAAGCC	*nuo*I2 / pBAD-A2/C3, pTTQ-A2/C3 and pQE-A2/C3
p7r	GAATTCGACACCTCCTCGGGTTTAGGTAATTT	
p8f	CAGGAGGAATTAACCATGGTTGCAATAAATTCTAACGG	*nuo*BD2 / pBAD-A2
p8r	CAGGTTTTCGGAATTCCTATCCGTCTCACCAACAACC	
p9f	GGTGAGACGGATAGGTAATTGTGGTTTATGCATTTTAGGAG	*nuo*I$_2$ / pBAD-A2-nuoBD$_2$
p9r	CAGGTTTTCGGAATTACCACTCCACTCTATCTTCCCCC	
p10f	CCTATCCGTCTCACCAACAACC	
p10r	AATTCCGAAAACCTGTACTTTCAAGGTG	
p11f	TTCGGGTGTATGTGTATAGTTCCGACGGTAGG	*nuo*G / pJET1.2
p11r	GAAGAGGCACTTTCGTGTTGCGATTTGG	
p12f	GCTTGGGAAGCAGGTGCTGGTGGTAGGAAGGAGATGGTG	*nuo*EF / pJET1.2-*nuo*G
p12r	GTAGCTCAGAGGTAGTTAAAGGGGAAGGGAAGCAG	
p13f	TTACCTATCCGTCTCACCAACAACC	*nuo*EFG / pBAD-A2-nuoBD$_2$I$_1$
p13r	TTGTGGTTTATGCATTTTAGGAG	
p14f	GAGACGGATAGGTAAGGTGGTAGGAAGGAGATGG	
p14r	ATGCATAAACCACAACTAAACACTCCTTTCG	
p15f	TAATTCCTCCTTCCTCGAGTCATTAATGGTG	*nuo*EFG / pBAD-CM1
p15r	TCTAGACCCGGGGTCGACTCAGC	
p16f	AGGAAGGAGGAATTAATGTTTAAAACGGAGTTTGAATTTCCCGAAG	
p16r	CGACCCCGGGTCTAGAAACACTCCTTTCGTAGTAAACC	
p17f	CTATGTAAAGATATACCAGTCGACTCTAGAGGATCCC	*nuo*A$_2$BD$_2$H$_1$I$_1$J$_1$K$_1$L$_1$M$_1$N$_1$ / pBAD33
p17r	GATGCGAAGAGGATTAAGCATGCAAGCTTGGCTGTTTTGG	

| p18f | TGGTATATCTTTACATAGTTTTTGC | |
| p18r | CTACCTCTACCAGCAGCGAGGATTATTGCC | |

Table 2.12. PCR reaction mixture

Components	Volume (μl) / 50 μl	Final concentration
H_2O	Add to 50	-
5×Phusion® HF Buffer	10	1×
10 mM dNTPs mixture	1	200 μM
primer F (10 μM)	1	0.2 μM
primer R (10 μM)	1	0.2 μM
template DNA	X	10-100 ng
Phusion® HF DNA Polymerase	0.5	0.02 U/μl

Table 2.13. PCR cycle conditions

Step	Temperature (°C)	Time	Cycle
Initial denaturation	98	1 min	1
Denaturation	98	10 s	
Annealing	Depends (55-72)	30 s	29
Extension	72	10-30s/kb	
Final Extension	72	10 min	1
	4	Hold	

2.2.1.4 DNA Electrophoresis and DNA purification

Prior to the electrophoresis, 1% agarose gel was prepared with the agarose powder (NEEO ultra quality, Carl Roth) dissolved in 1×TAE buffer and DNA sample was prepared by addition of DNA Loading buffer (Thermo Scientific). Afterwards, the gel was submerged in 1×TAE running buffer hold by a self-made horizontal gel chamber and the samples were loaded on the gel. The run was performed using PowerPacTM Basic power suppliers (Bio-Rad) at room temperature (RT). The DNA gel was run at 100 V (5V/cm) for 60-90 mins and the progress of the separation was monitored using colored dyes in the loading buffer. 100 bp or 1 kb DNA ladders (NEB) were used to identify the size of the DNA fragments. After electrophoresis, the gel was stained in ethidium

bromide (Roth) solution for 5-10 mins at RT and destained in water for 5 - 10 mins. The gels were subsequently visualized and photographed using a Bio-Rad gel documentation system with a 302 nm UV transilluminator or visualized under UV light (312 nm, Biometra TI1 transilluminator). The buffer used for electrophoresis of DNA is listed in Table 2.14.

Table 2.14. TAE buffers for agarose gel electrophoresis

Buffer	Components	Preparations
TAE (50×)	2 M Tris-HCl	242 g
	1 M glacial acetic acid	57.1 mL
	50 mM EDTA	100 mL 0.5 M EDTA (pH8.0)
	H_2O	Add to 1 L

DNA bands of interest were excised from the gel with a clean scalpel and purified using ZymocleanTM Gel DNA Recovery Kit (ZYMO RESEARCH) according to manufacturer's protocol. The purified DNA was dissolved in nuclease-free ddH$_2$O and stored at -20 °C before use.

2.2.1.5 Absorbance based DNA Quantification

DNA concentration was determined using a NanoDrop Spectrophotometer (Thermo Scientific). 1 - 2 μl of DNA sample was enough for a measurement. The quality of DNA sample was evaluated by the ratio of absorbance at 260 nm vs 280 nm and 260- nm vs 230 nm. Pure DNA preparations have an A260/A280 ratio of greater than or equal to 1.8 and A260/A230 of approximately 2.0-2.2.

2.2.1.6 Gene cloning

PCR products were cloned into the positive selection cloning vector pJET1.2/blunt using CloneJET^{TM} PCR Cloning Kit (Thermo Scientific) according to the manufacturer's protocol. The resulting product was stored at -20 °C before transformation.

2.2.1.7 Restriction digestion and ligation

The double digestion of DNA was set up in a 1.5 ml Eppendorf tube containing 20 to 50-µl reaction mixtures (Table 2.15). The large scale protocol was used to prepare DNA for downstream applications (i.e. ligation), and the small scale protocol was used for screening. The reaction was performed using the Fast Digest Enzyme (Thermo Scientific) according the recommended protocol at 37 °C for 15-30 mins. The resulting DNA fragments were analyzed by gel electrophoresis or further purified for ligation.

The ligation of DNA insertions into vectors was done using the Quick Ligation kit (NEB) for 5 to 15 mins at RT. The reaction mixture was prepared as recommended protocol (Table 2.16).   Normally a molar ratio of 3:1 and 5:1 (insert:vector) was used for cohesive-end ligations and blunt-end ligations, respectively.

Table 2.15. Reaction Conditions for FastDigest

Components	Large scale		Small scale
	Plasmid DNA	PCR product	Plasmid DNA
Nuclease-free water	Add to 50 µl	Add to 50 µl	Add to 20 µl
10×FastDigest® buffer or 10×FastDigest® Green buffer	5	5	2
Enzyme 1	2.5	2.5	1
Enzyme 2	2.5	2.5	1
DNA	Up to 5 µg	~500 ng	Up to 1 µg
Total volume	50 µl	50 µl	20 µl

Table 2.16 Ligation reaction conditions

Components	Volume (μl)
Quick ligation Buffer ($2\times$)	5
Vector DNA	50-100 ng
Target inserts	50-100 ng
Quick T4 DNA ligase	0.5
H_2O	Add to 10 μl

2.2.1.8 Ligation independent cloning

Ligation independent cloning was performed for directional cloning without restriction digestion and ligation [98-100]. The target insertions were amplified using the primers containing 20 to 35 bps with the addition of 15-20 bps overlapping at the 3'- end, which are gene-specific and homologous to the vector ends. The insertion was amplified and the vector was linearized by PCR using Phusion HF DNA polymerase. The amplified insertion and vector were analyzed by gel electrophoresis and purified from agarose gel. The ligation independent cloning reaction was carried out as described before [98] or using the In-Fusion HD cloning kit (Clontech) following the recommended protocol. The resulting ligation product was stored at -20 °C before transformation.

2.2.1.9 Preparation of chemically competent cells and transformation

The chemically competent cell was prepared according to the rubidium chloride method [101]. 4 ml of LB medium was inoculated with a single colony and incubated at 37 °C, 220 rpm overnight as preculture. 1 ml of preculture was subcultured in 100 ml SOC/LB medium in a 250 ml flask and shaken (180 rpm) at 37 °C to an OD_{600} of 0.4 – 0.6. The culture was then chilled on ice. Cell suspensions were pelleted at 4,400 g for 5 mins at 4 °C (Sigma 4K15 centrifuge). Cell pellets were gently resuspended in 40 ml of cold TFB I buffer and incubated on ice for 5 mins. Afterwards, the cells were re-pelleted at 4 °C, 4400 g for 5 mins, gently resuspended in 4 ml of cold TBF II buffer and incubated on ice for 15-60 mins. The resulting competent cells were transferred into sterile and pre-chilled 1.5 ml microcentrifuge tubes with 100 μl/aliquot, and flash frozen in liquid nitrogen and

stored at -80 °C. All buffers and solutions used for competent cell preparation are listed in Table 2.17.

The competent cell was thawed on ice. 1-5 µl ligation product or plasmid was added to 50 µl competent cells as soon as the last bit of ice in the tube disappeared. After incubation for 30 mins on ice, the mixture was subjected to a heat shock at 42 °C for 45 s without shaking. Afterwards, the solution was incubated on ice for 5 mins. 950 µl of RT SOC medium was added and the resulting mixture incubated in a 10 ml tube at 37 °C shaking vigorously (220 rpm) for 1 h. An appropriate amount of the mixture (around 1/10 for plasmid DNA and 100% for ligation mix) was spread on LB agar plates containing the appropriate antibiotics. The plates were incubated at 37 °C overnight.

Table 2.17 Buffers used to prepare chemically competent cells

Buffer	Components	Preparation
TFB I	30 mM KOAc, pH 5.8	12.3 ml 1 M KOAc
	100 mM RbCl	5 g RbCl
	50 mM $MnCl_2$	20.5 ml 1 M $MnCl_2$
	10 mM $CaCl_2$	4.1 ml $CaCl_2$
	15 % glycerol (v/v)	61.5 g Glycerol
	H_2O	Add H_2O to 100 ml
		Adjust to pH 5.8 with acetic acid and filter to sterilize
TFB II	100 mM MOPS, pH 6.8	1.5 ml 1 M MOPS
	10 mM RbCl	1.5 ml 1 M RbCl
	75 mM $CaCl_2$	11.25 ml 1 M $CaCl_2$
	15 % glycerol (v/v)	22.5 g glycerol
	H_2O	Add water to 150 ml
		Adjust to pH 6.5 with KOH and filter to sterilize

2.2.1.10 DNA verification

The positive colony was verified by colony PCR or restriction digestion and further confirmed by DNA sequencing. Typically 5-10 single colonies were tested.

Taq DNA polymerase (NEB) was used for colony PCR and the reaction conditions was shown in Table 2.18. The individual colony was added directly to the PCR reaction and lysed during the initial heating step. The released plasmid DNA from the cell served as template for the amplification reaction. Primers targeting vector DNA flanking the insert were used to determine the specificity and size of the insert DNA. Presence or absence of a PCR amplicon and size of the product are determined by DNA electrophoresis alongside a DNA size marker on an agarose gel.

Table 2.18 Colony PCR conditions

Step	Temperature (°C)	Time	Cycle
Initial denaturation	95	10 min	1
Denaturation	95	30 s	
Annealing	60	30 s	24
Extension	68	1 min/kb	
Final Extension	68	5 min	1
	4	Hold	

DNA sequencing was conducted by Microsynth Seqlab and Eurofins Genomics. DNA was prepared as recommended instructions. Sequencing primers are listed in Table 2.19

Table 2.19 List of primers used for DNA sequencing

Name	Primers	Sequencing Target
T7 promoter	TAATACGACTCACTATAGGG	pET vectors
T7 terminator	GCTAGTTATTGCTCAGCGG	
pBADseqfor7	AGATTAGCGGATCCTACCT	pBAD vectors
pBADseqrev8	CAGACCGCTTCTGCGTTCTG	
pQ5-seqfor	TGAGCGGATAACAATTTCAC	pQE vector
pQ6-seqrev	GAGTTCTGAGGTCATTACTG	
pTTQ-3 seqfor	CATCATAACGGTTCTGGCA	pTTQ vector
pTTQ-4 seqrev	CTATTACGCCAGCTGGCGA	
*nuo*EFG seqfor1	TCTACGAACACGAAACCTGC	*nuo*EFG
*nuo*EFG seqfor2	GTGCGACGAGGTCGTAGGTA	

*nuo*EFG seqrev1	CCTGCTCTGTAGCTTCACCC	
SP1	CGTTCCCAGAGAACTAAAGGTG	*Nuo*BD$_2$EFGI$_2$
SP2	AGAACGGAAAAGGCGAGACT	
SP3	GGAGAACTAAGGATGTTGGCG	
SP4	AAGGTGGTAGGAAGGAGATGG	
SP5	CCAGAGTTCATTCCATAGATGAGT	
SP6	TATAATTTCCATGGGCTGGG	
SP7	GAAATCTGCTTCCCTTCCCC	
SP8	ACGAGAGCCCTTTACGTTGA	
SP9	ACGCTTTCGTGTTCATAGGC	
SP10	GCCTTCTGGAAGTACCTCGG	
SP11	GTCTCTTCTGTGGGCTCTGC	
SP12	GAGACAGGCTGACGTTCTCA	*nuo*A$_2$BD$_2$H$_1$I$_1$J$_1$K$_1$L$_1$M$_1$N$_1$
SP13	CCTGAGGAAAACGAAAGGGT	
SP14	GTAGAGGTTCCCGAGAAGGC	
SP15	TAAACAAGGAACACCCTGCG	
SP16	GATGGCTTTAGTTCCGTCCA	
SP17	TGTTCAAGACATTTGCACTGG	
SP18	CGGATATCTTTGAACTCGCA	
SP19	GGTGGGTGCAATCCTTCTT	
SP20	CACTCTCTTCCTTGATGGCA	
SP21	ATCGGAGCGATAACCATGAC	
SP22	GCTTGATAGGAGGGGAGGAG	
SP23	TTGAAACATACGTTCCGTGG	
SP24	GATTTTCCCTGCCCTTCTTC	
SP25	ATAATTGGCATGGCTCTCGT	
SP26	ACTCGGTTCAGCACTCGTTT	
SP27	GATAGGTATCCCTCCCGCTG	

2.2.1.11 Site-directed mutagenesis

Site-directed mutagenesis was performed using QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies, Inc) following the manufacturer's specifications. The primers were generated by QuikChange Primer Design online (https://www.genomics.agilent.com/primerDesignProgram.jsp). The primers for mutation

were listed in Table 2.20. After site-directed mutagenesis, the resulting constructs were confirmed by DNA sequencing.

Table 2.20 List of primers for mutation

Oligonucleotide	Sequence (5'-3')	Purpose
MP1	GACTTCACACCTGTGTCTACTACCGAGTTCGTAGC	C49S
MP2	GCTACGAACTCGGTAGTAGACACAGGTGTGAAGTC	
MP3	CTGGTACATCTCTAAGAAGAAGGTTGTCGACAAG	C212S
MP4	CTTGTCGACAACCTTCTTCTTAGAGATGTACCAG	
MP5	CAAGAAGGATACGAGTCTGCAGACTGGTACATCTG	C218S
MP6	CAGATGTACCAGTCTGCAGACTCGTATCCTTCTTG	

2.2.1.12 Storage of bacteria strain

E. coli strains were stored at -80 ℃ in 25% glycerol (v/v). Typically, 500 μl of pre-sterilized glycerol (50%) were mixed with 500 μl of a late logarithmic-phase *E. coli* culture suspension in LB medium. The glycerol stocks were then flash frozen in liquid nitrogen and stored at -80 ℃.

2.2.2 Protein Biochemistry

2.2.2.1 Inducible expression in *E. coli*

The inducible expression in *E. coli* was performed as follows:

(1) A single colony was picked from a freshly streaked plate of the expression host containing the recombinant vector and inoculated 4 ml LB medium containing the appropriate antibiotic. The 4 ml culture was incubate at 37 ℃ with vigorous shaking at 220 rpm till OD600 reached 0.6-0.8 for inducible expression test or till OD600 was approximate 1 to inoculate the starter culture by 1% (v/v).

(2) The starter culture was prepared by inoculating 1 ml precultured *E. coli* cell in a 250 ml unbaffled Erlenmeyer flask containing 100 ml LB medium. After growing the OD600 around 1, the culture was stored at 4 ℃ overnight.

(3) 6 × 2 L of LB medium (in 5 L unbaffled Erlenmeyer flask), supplemented with appropriate antibiotic and additives (see table 2.6) , were inoculated with 6 × 20 ml of the starter culture and incubated at 37 °C with vigorous shaking at 180 rpm till OD600 reached 0.6-0.8. The cells were then induced with 0.3 mM IPTG (for vectors pTTQ18, pQE, pET) or 0.05% (w/v) arabinose (for vectors pBAD).

The cells were harvested by centrifugation at 10,500 g at 4 °C for 15 min (Centrifuge Avanti J-26XP, Rotor JLA-8.1000, Beckman Coulter), flash frozen in liquid nitrogen and stored at -80 °C until use.

2.2.2.2 Expression test and Determination of protein cellular localization

Protein expression was tested in whole cell lysate under denaturing condition. The cell pellets of 1 ml culture were resuspended in 100 μl of 20% (w/v) SDS in water and incubated for 15 min at 37 °C. The suspension was then centrifuged at 16,000 × g for 20 min (Centrifuge 5415D, Eppendorf) at RT to remove cell debris and genomic DNA. The supernatant was analyzed by SDS-PAGE and Western blot.

To determine the protein cellular localization (membrane-inserted, cytoplasm, or inclusion bodies), frozen cells were suspended in the lysis buffer containing DNase I (Roche) and protease inhibitor (Roche) and disrupted by French press (SLM Aminco®) twice at 1000 psi. The insoluble cytoplasmic fraction (cell debris and inclusion bodies) was collected by centrifugation at 4 °C, at 23,000 × g for 60 min (Rotor GSA/SS34, Sorvall RC5B superspeed centrifuge). The membrane fraction was collected by centrifugation at 4 °C at 200,000 × g for 60 min (Rotor 70 Ti) or at 150,000 × g for 90 min (Rotor 45 Ti) (Ultracentrifuge Optima L-90K, Beckman Coulter). The soluble cytoplasmic fraction was collected as the supernatant of the latter centrifugation step. The protein concentration of the three fractions was determined by BCA methods and the fractions were analyzed by western blot.

2.2.2.3 Membrane preparation

The membrane preparation was performed at 4 °C. Cell pellets from 1–12 L culture were suspended in cell lysis buffer (20 mM Tris-HCl, pH 7.4, 1 mM EDTA, 2 mM DTT, 10%

Glycerol, DNAase grade II, protease inhibitor, 2 mM $MgCl_2$) in a ratio of 1 g cells to 6-ml buffer. The cell suspension was homogenized and then disrupted using a French Press at a pressure of 19,000 psi (40K cell, Thermo Fisher Scientific) for at least 2 cycles. Alternatively, large amount of cell suspension (200 mL or more) was lysed using a microfluidizer at a pressure of 12,000 psi (Microfluidics Corp) at least for 3 times. After cell disruption, the cell lysate was centrifuged at 23, 000 g for 60 min to remove the cell debris. The supernatant was then ultracentrifuged at 200,000 g for 60 min (Rotor 70 Ti) or at 150,000 $\times$ g for 90 min (Rotor 45 Ti) (Ultracentrifuge Optima L-90K, Beckman Coulter). The resulting membrane pellet was suspended in the buffer containing 20 mM Tris-HCl, pH 7.4, 1 mM EDTA, 2 mM DTT, 10% Glycerol and flash frozen in aliquots using liquid nitrogen before storage at -80 ℃. The total protein concentration of the membrane was evaluated using BCA assay.

2.2.2.4 Solubilization of membrane proteins and detergent screening

The frozen membranes solution was thawed and diluted to concentration of 10 mg/ml using the same buffer. The initial solubilization was performed by adding 1% n-Dodecyl β-D-maltoside (DDM) to the membrane solution dropwise and rotated top-to-bottom for 2 hours at 4 °C. Protease inhibitor was added to avoid protein degradation. The protein solution was subsequently ultracentrifuged at 200,000 g for 60 min (Rotor 70 Ti) or at 150,000 $\times$ g for 90 min (Rotor 45 Ti) (Ultracentrifuge Optima L-90K, Beckman Coulter). The obtained supernatant containing solubilized membrane proteins was filtered through a 0.2 μm polyethersulfone (PES) membrane prior to further purification. If DDM was not suitable for the target protein, a small scale solubilization screening was implemented to find an optimal detergent and condition. The detergents used in this screening procedure are listed in Table 2.21.

Table 2.21 Detergent used for membrane solubilization

Detergent	Abbreviation	CMC% (w/v)	Concentration % (w/v)	
			Solubilization*	Purification
n-decyl-β-D-maltoside	β-DM	0.087	1% - 2%	0.1
n-dodecyl-β-D-maltoside	β-DDM	0.0087	1% - 2%	0.05
cyclohexyl-hexyl-β-Dmaltoside	Cymal6	0.028	1% - 2%	0.05
2,2-dioctylpropane-1,3-bis-β-D-maltopyranoside	DMNG	0.0034	1% - 2%	0.03
2,2-didecylpropane-1,3-bis-β-D-maltopyranoside	LMNG	0.001%	1% - 2%	0.01
2,2-dihexylpropane-1,3-bis-β-D-glucopyranoside	OGNG	0.058	1% - 2%	0.1
Dodecyl Octaethylene Glycol Ether	C12E8	0.0048	1% - 2%	0.02
n-dodecyl-N, Ndimethylamine-N-oxide	LDAO	0.026	1% - 2%	0.1
n-Dodecylphosphocholine	FOS12	0.053	1% - 2%	0.1

* Concentration of the detergent for 10 mg/ml membrane.

2.2.2.5 Protein purification

2.2.2.5.1 Immobilized metal affinity chromatography (IMAC)

Ni-NTA affinity chromatography Ni-NTA agarose (Qiagen) was used in a batch procedure for His-tagged protein purification. The cell lysate was filtered using 0.22 μm filters (Sarstedt) and then incubated with pre-equilibrated Ni-NTA beads with lysis buffer containing 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 30 mM imidazole) for 2 h at 4 ℃ with constant rotation at 6 rpm. After incubation, the mixture was loaded onto a gravity column and first washed with lysis buffer to wash off the unbound proteins and then washed by buffer containing 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, and 70 mM imidazole to remove the non-specific bound proteins. Finally, the protein was eluted with buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl and 200 mM imidazole. Purification of the membrane protein was performed in a similar manner, but with the buffer containing 0.1% FOS12. The eluted fraction was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Strep-Tactin Purification Strep-tagged protein was purified using Strep-Tactin®
Superflow® high capacity 50% suspension (IBA) following a batch purification protocol
according to the manufacturer's instructions. The membrane fraction was supplemented
with 0.2 mg/ml avidin and incubated with appropriate amount of beads at 4 °C for 2
hours. The unbound proteins were washed off by washing buffer (20 mM Tris-HCl, pH
7.4, 150 mM NaCl, 1 mM EDTA, 2mM DTT and 0.05% DDM), while the target protein
was eluted with elution buffer (washing buffer with 5 mM desthiobiotin).

2.2.2.5.2 Ion exchange chromatography

The ion exchange chromatography was performed using the Mono Q 10/100 GL column
(GE healthcare) on Äkta purifier systems (GE Healthcare). The components of the start
buffer and elution buffer were shown in Table 2.22. The solubilized membrane fractions
were filtered using the 0.22 μm filters and loaded on the pre-equilibrated column.
Different proteins were separated using gradient elution.

Table 2.22 Buffer used in Ion exchange chromatography

	Components
Start buffer	20 mM Tris-HCl, 2 mM EDTA, 2 mM DTT, 10% glycerol, 0.1% DDM
Elution buffer	20 mM Tris-HCl, 2 mM EDTA, 2 mM DTT, 10% glycerol, 1 M NaCl, 0.05% DDM

2.2.2.5.3 Size exclusion chromatography

Size exclusion chromatography was routinely used as the last step of purification or to
check the homogeneity of the purified proteins. The Superdex 200 10/300 GL column
(GE healthcare) or the TSK-GEL G4000SW (TOSOH Bioscience) connected to an Äkta
purifier systems (GE Healthcare) were used for preparative purpose. Superdex 200 PC
3.2/30 column (GE healthcare) or Yarra 3u SEC-4000 column (300 × 4.6 mm,
Phenomenex®) connected to the Äkta Micro system was used for analytical purposes.
Prior to application of the protein sample, the SEC column was equilibrated with proper
buffer and the eluted fractions from affinity purification were concentrated to proper
amount using Amicon concentrators (Ultra-4 & Ultra-15, 50 kDa cut-off, Millipore) and
filtered through a 0.2 μm membrane. The detailed conditions are specified in "Result"
chapter.

2.2.2.6 Determination of protein concentration

The bicinchoninic acid (BCA) protein assay [102] was used to determine the concentration of total protein during the solubilization and purification precisely. The assay was performed using the Pierce™ BCA Protein Assay Kit (Thermo Scientific) according to the User Guide. If the amount of the protein was very limited, protein concentration was roughly evaluated by measuring the absorbance at 280 nm in a NanoDrop ND-1000 spectrophotometer (one A280 unit = 1 mg/ml).

2.2.2.7 SDS-PAGE

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) separates proteins and protein subunits by size under denaturing conditions and can be used to check the purity of proteins and analyze the composition of protein complexes. SDS-PAGE was performed using the NuPAGE® precast 4-12% Bis-Tris gel in a commercial gel chamber (XCell SureLock Mini-Cell, Invitrogen) containing $1\times$ MES running buffer at constant 200 V, for 45 min in cold room (4 ℃). The buffers and solutions used for SDS-PAGE were prepared according to NuPAGE® Technical Guide (Invitrogen).

2.2.2.8 Non denaturing gel electrophoresis (Native PAGE)

Native PAGE was performed using precast NativePAGE™ Novex® 4-16% Bis-Tris Mini in an XCell SureLockTM Mini-Cell chamber (Invitrogen) according to the manufacturer's specifications. The runs were conducted at 4 °C, at 150 V constant for 120 min (Pharmacia LKB ECPS 3000/ 150). Blue Native PAGE (BN-PAGE) was performed using 50 mM BisTris and 50 mM Tricine as the anode buffer and the same buffer supplemented with 0.02% (w/v) or 0.002% Coomassie brilliant blue G-250 as the dark or light blue cathode buffer, respectively. The samples with detergents were supplied with 5% (w/v) Coomassie brilliant blue G-250 as an additive prior to loading.

2.2.2.9 Gel staining

Gels were stained by either one of the following two methods:

Coomassie blue staining - After electrophoresis, the gels were directly stained with Coomassie blue staining solution (detection range 5 – 10 µg protein). After 30 - 60 min at RT, the gels were destained using destaining solution.

Table 2.23 Solutions used for Coomassie blue staining

	Components
Coomassie staining solution	0.04% (w/v) Coomassie brilliant blue R-250
	40% (v/v) ethanol
	10% (v/v) acetic acid
	H_2O
Destaining solution	30% (v/v) ethanol
	10% (v/v) acetic acid
	H_2O

Silver staining - Silver staining was used for sensitive detection (minimum detection limits in the range of 0.3 to 10 ng) of proteins separated by SDS-PAGE / Native PAGE. The silver staining was performed with the SilverQuest^{TM} silver staining Kit (Invitrogen) following the manufacturer's specifications.

2.2.2.10 Antibody production

In this work, 15 polyclonal antibodies were generated against subunits of *A. aeolicus* complex I (NuoA$_2$, NuoB, NuoD$_2$, NuoE, NuoF, NuoG, NuoH$_1$, NuoI$_1$, NuoJ$_1$, NuoK$_1$, NuoL$_1$, NuoM$_1$, NuoN$_1$, NuoD$_1$, and NuoI$_2$). Correspondingly, the antibodies were named anti-NuoA$_2$, anti-NuoB, anti-NuoD$_2$, anti-NuoE, anti-NuoF, anti-NuoG, anti-NuoH$_1$, anti-NuoI$_1$, anti-NuoJ$_1$, anti-NuoK$_1$, anti-NuoL$_1$, anti-NuoM$_1$, anti-NuoN$_1$, anti-NuoD$_1$, and anti-NuoI$_2$, respectively. Four of the antibodies, anti-NuoD$_1$, anti-NuoD$_2$, anti-NuoI$_1$, and anti-NuoI$_2$, were used for immunoprecipitation. Other antibodies were used for western blot. All the antibodies were produced following a 2-rabbit 90-day protocol in Thermo Scientific Pierce custom antibody service. The reaction of crude sera to the target protein was tested by western blot. The selected crude sera will be used for affinity purification to obtain polyclonal antibodies.

2.2.2.11 Western blot analysis

Western blot was used to identify target proteins by specific antibodies during gene expression, protein location and purification. The poly-histidine-tagged proteins were detected using a monoclonal α-poly-histidine-alkaline phosphatase conjugated antibody produced in mouse (Sigma-Aldrich) at a dilution of 1:2000. The Strep-tag fusion proteins were detected by Precision Protein™ StrepTactin-AP Conjugate (BIO-RAD) at a dilution of 1:5000. The subunits composed of complex I were analyzed using specific custom peptide polyclonal antibodies as the primary antibodies (1:500 dilution) and monoclonal Anti-Rabbit Immunoglobulins-Alkaline Phosphatase antibody produced in mouse (Sigma-Aldrich) as the secondary antibody. The proteins separated on the SDS-PAGE / Native PAGE gels were transferred to the PVDF membrane using the iBlot® 7-minute Blotting System (Invitrogen). After that, the membrane was incubated in blocking buffer for 1 hour at RT or overnight at 4 ℃. Subsequently, the membrane was washed with $1\times$ TBST buffer for 5 min with gentle agitation and incubated with the primary antibodies for 1 - 2 h. After the incubation, the membrane was washed 3 times with $1\times$ TBST buffer, followed by the incubation with the second antibodies conjugated to alkaline phosphase (AP) for another 1 hour. After washing for 5 min (3 times) with $1\times$ TBST buffer, and with AP buffer 3 times for 5 min, the signals were detected using the a SIGMAFAST™ BCIP®/NBT tablet (Sigma-Aldrich) dissolved in 10 ml deionized water.

Table 2.24 Lists of solutions used for western blot analysis

Solutions	Components
TBST buffer ($10\times$)	100 mM Tris-HCl, pH 8.0, 1.5 M NaCl 0.5% (v/v) Tween-20
AP buffer ($10\times$)	100 mM Tris/HCl, pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$
Blocking buffer	1% (w/v) BSA in $1\times$ TBST buffer
Blocking buffer (Strep-Tactin)	1% (w/v) Biotin-free BSA in $1\times$ TBST buffer

2.2.2.12 Immunoprecipitation

In this study, immunoprecipitation (IP) is used for the small-scale affinity purification of individual isoform of complex I using a specific polyclonal antibody that is immobilized to a solid support such as magnetic particles or agarose resin. Anti-NuoD$_2$ and anti-NuoI$_1$ are NQOR1-specific antibodies and are utilized to precipitate NQOR1, while anti-NuoD$_1$ and anti-NuoI$_2$ are NQOR2-specific antibodies and are aimed to precipitate NQOR2. The immunoprecipitation was first performed using the Pierce™ Classic IP Kit (Thermo Scientific™). To eliminate the coprecipitated contaminants and improve the binding efficiency, purified complex I S1 sample was used as antigen. The complex I S1 sample (cells were grown at 1% oxygen concentration and 6 h before harvesting the oxygen concentration was increased to 3%) was prepared as described before [103]. The antibodies were affinity-purified for optimal results. The IP followed the instruction of the Pierce™ Classic IP Kit with slight modification. Considering that complex I is membrane protein, all the buffer used here are containing 0.05% DDM. In addition, the IP washing buffer was change to be the same with S1 sample buffer containing 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, and 0.05% DDM. A total amount of 26 μg S1 sample was combined with 8 μg of each antibody (the molar ratio of Ab/Ag is 1:1). The Ab/Ag mixture was diluted to 300 μl by washing buffer and rotated at 4 ℃ for 1 hour. The immune complex I was added to Protein A/G Plus Agarose in the spin column. The column was incubated with gentle end-over-end mixing for 1 hour. Afterwards, the column was washed by IP washing buffer (3 times) and 1×conditioning buffer (once). The protein was eluted with 50 μl Low-pH elution buffer. 5 μl of 1 M Tris, pH 9.5 was added to the collection tube to neutralize the pH. The result of immunoprecipitation was subsequently analyzed by western blotting. To improve the Ab-Ag binding efficiency, the molar ratio of Ab/Ag (1:1, 2:1, 5:1 and 10:1) was optimized. Pierce™ Co-Immunoprecipitation Kit was used to eliminate co-elution of antibodies.

2.2.2.13 Enzyme activity assay

In gel activity assay In-gel histochemical staining of proteins separated by native electrophoresis is a valuable tool during solubilization and purification of respiratory

complexes. It can be used for the detection of protein complexes in different purification fractions after chromatography. In gel activity of complex I is based on its NADH hydrogenase activity in which complex I oxidizes NADH, with hydrogen transferred to nitro-blue tetrazolium (NBT). The reduction of NBT results in formazan which precipitates as blue-purple crystals visualizing the complex I bands on the electrophorsis gel [104]. Following the protocol by Wittig *et al.* [105], the gels of BN-PAGE were incubated in complex I assay solution (5 mM Tris-HCl, pH 8.0, 2.5mg/mL NBT and 0.1mg/mL NADH) immediately after electrophoresis for 30 mins at 37 °C. The assay was stopped using fixing solutions containing 40% ethanol and 10% acetic acid.

NADH:ubiquinone oxidoreductase activity assay The NADH:ubiquinone oxidoreductase activity assay of complex I was performed according to the protocol described by Estornell *et al.* [106] and Peng *et al.* [103]. Three different quinones (Table 2.26) were tested in the assay, including the commercial available ubiquinone analog decylubiquinone (DQ, Sigma Aldrich), and two anologs of the natural *A. aeolicus* quinone 2-demethylmenaquinone-7 (DMK) [107], DMK-S2 and DL01, which are designed and kindly provided by the collaborator Dr. Hamid Nasiri (guest researcher in the group of Prof. Schwalbe, Institute for Organic Chemistry and Chemical Biology, University of Frankfurt) and Dana Lashley (visiting assistant professor at the department of Chemistry at the College of William & Mary, Virginia). 2–4 μg/ml of purified complex I sample together with 80 μM of the respective quinone were premixed and incubated in assay buffer containing 50 mM KCl, 10 mM Tris-HCl, 1 mM EDTA, 2 mM KCN (pH 7.4). The reaction was started by adding 150 μM NADH. The oxidation of NADH was monitored the absorption change of 340 nm over 5 mins at 80 °C, using a DW2000 UV-VIS spectrophotometer (SLM AMINCO).

Table 2.25. Quinone analogs used in this work

Name	Formula	Molecular weight
Decylubiquinone	$C_{19}H_{30}O_4$	322.45
DMK-S2	$C_{19}H_{24}O_2$	284.39
DL01	$C_{12}H_{14}O_2$	190.24

2.2.3 Biophysical method

2.2.3.1 Mass spectrometry

Peptide mass fingerprinting mass spectrometry (PMF-MS)

The protein samples can be derived from the SDS-PAGE or BN-PAGE gels or directly prepared in solution. For the gel-separated protein, the gel band of target protein was excised, de-stained and dried. Afterwards, the gel band was soaked into a trypsin solution for digestion overnight at 37 ℃. The digested peptides were extracted from the gels by sonication and the solution was acidified with TFA and dried under vacuum. For protein prepared in solution, the protein was denatured by urea before digestion. The peptides were separated by high performance liquid chromatography (HPLC) and then were transferred from the LC column into the MS device for measurement. After measurement, the peptide masses are compared to a database containing known protein sequences. For cysteines-rich protein, reduction and alkylation prior to in-gel digestion the reduction and alkylation step may increase the protein coverage. The mass spectrometry was performed by Fiona Rupprecht from the group of Membrane and crystal MS in Max-Planck-Institute of Biophysics.

2.2.3.2 Electron paramagnetic resonance spectroscopy

EPR samples (1-5 mg/ml) were prepared by reaction with 2.4 mM NADH for 30-60 s in buffer containing 20mM Tris-HCI, pH 7.4, 150mM NaCI and 0.05% DDM. After incubation, the reduced protein sample was transferred into an EPR tube and immediately frozen in a mixture of 2-Methylbutane (Fluka) and Methycyclohexane (Fluka) in a ratio of 80%: 20% (v/v) that was cooled down using liquid nitrogen to approx. 120 K. The frozen samples were stored in liquid nitrogen until EPR measurements. The EPR spectra were record on a Bruker ESP 300E spectrometer at X-band using an Oxford Instrument ESR900 helium flow cryostat. Predominant reduction of cluster N2 was achieved upon addition of a few grains of solid sodium dithionite and freezing the solution after short reaction time (5–10 s). Simulations of spectra were performed using the Bruker Simfonia Software package. Redox titrations were performed essentially as described by Garofano et al. [108]. The protein sample S2 (1.2 mg/ml) was titrated under anaerobic conditions at

RT, pH 7.4, in the range from −450 to +240 mV. Evaluation of EPR signals from tetranuclear clusters was done at 12 K and of the binuclear cluster at 40 K. The EPR measurement was conducted within collaboration with Dr. Klaus Zwicker, Dr. Alberto Collauto, and Prof. Dr. Thomas Prisner from Johann Wolfgang Goethe Universität.

2.2.3.3 Ultraviolet-visible spectroscopy

Ultraviolet–visible (UV-VIS) spectroscopy was used to analyze the cofactors of *A.* - *aeolicus* complex I under different redox state. The measurement was performed using a Lambda 35 UV/Vis spectrometer (Perkin Elmer) and 10 mm quartz micro-cuvettes (104.002-B, Hellman). The purified complex I was dissolved in buffer containing 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM DTT, 1 mM EDTA, 0.05% DDM at 0.91 mg/ml. Absorption spectra were measured between 400 and 700 nm with a scan rate of 120 nm/min (data interval: 0.2 nm; slit width: 0.5 nm). The spectroscopic changes in different redox states, air-oxidized (as purified) complex I and fully-reduced complex I by adding 10 mM NADH were recorded at RT.

2.2.3.4 Surface Plasmon Resonance

The interaction of complex I with AhpC2 was analyzed by (Jonsson et al. 1991) SPR measurements [109] using a BIAcoreT100 or BIAcore 3000 system with CM5 chips (GE Healthcare) at RT. Prior to the measurement, the protein samples were exchanged to running buffer containing 50 mM HEPES, pH 7.4, 500 mM NaCl, and 0.05% (w/v) DDM using desalting column. One flow cell of CM5 sensor chip was activated with a 1:1 mixture of 0.2 M 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and 0.05 M N-Hydroxysuccinimide (NHS) in water as described by the manufacturer. To detect the binding of complex I to different concentration of AhpC2, native complex I was immobilized on the chip at a concentration 50 μg/ml; approximately 10076 RU of CPI was immobilized. The remaining binding sites were blocked by 1 M ethanolamine. Gradient concentrations of AhpC2 (treated by 1 M NaCl overnight) were then flowed over the chip surface at a flow rate 30 μl/min. The non- specific binding to a blank flow cell was subtracted to obtain corrected sensorgrams. The binding kinetics were analysed with Biacore T200 Evaluation Software using a w/o WTL model. Another measurement

was performed on BIAcore 3000 using CM5 chip using same method with the except that the sample was not treated with 1 M NaCl and a buffer containing 10mM HEPES, 150mM NaCl, pH7.4 and 0.05% DDM was used as running buffer. The SPR measurement was collaborated with the group of Prof. Zihe Rao in the Institute of Biophysics, Chinese Academy of Science.

2.2.3.5 Single particle electron microscopy

The negative stain grids were prepared using a droplet method as described [110] with a slight modification. In brief, before sample application, the carbon-coated grid (Beijing Zhongjingkeyi Technology Co., Ltd) was glow discharged for 1 min. The purified protein sample was diluted to a concentration of 0.01mg/ml with sample buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, with 0.1% DDM added for membrane protein. A 2.5 μl drop of protein solution was placed on the holey carbon grid. After 1 min incubation, the grid was continuously washed with two droplets of water and one droplet of 2% uranyl acetate (Beijing Zhongjingkeyi Technology Co., Ltd). Then the grid was incubated with one drop of 2% uranyl acetate for 1 min. The excess stain was removed using a filter paper and the grid was air-dried. The images were recorded at a nominal magnification of 57,000× (1.57 Å/pixel) using an FEI Talos F200C field emission electron microscope operated with an acceleration voltage of 200kV. Particles were picked using EMAN2 [111], CTF-corrected by CTFFIND3 [112] and used for generating 2D average using relion 1.4 [113]. Totally 1037 particles were extracted with a box size of 100 pixels. The extracted particle images were normalized and utilized for 2D classification. Particle coordinates were classified according to 2D classification with manual adjustment. Finally, each particle coordinates class was used for generating the 2D average picture, respectively. The EM was collaborated with the group of Prof. Zihe Rao in the Institute of Biophysics, Chinese Academy of Science.

2.2.3.6 Protein crystallization

Initial protein crystallization screening was performed automatically using the CrystalMation system (Rigaku, Carlsbad, California, USA) [114]. Crystallization was set up in 96-well plates by using the sitting-drop vapor diffusion method at 291 K. A

versatile set of screening conditions was used (see Table 2.7). Further optimization was carried out using hanging-drop vapor diffusion by varying the concentration of protein precipitant and salt on 24-well plates manually. 1 μl protein sample at 8- mg/ml in 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, was mixed with the same volume of reservoir solution and equilibrated against 400 μl reservoir solution at 291 K.

2.2.3.7 Data collection, Structure determination, and refinement

Before data collection, the crystals were harvested using cryoloops and soaked quickly with reservoir solution containing 20% (v/v) glycerol as cryoprotectant solution. The crystals were subsequently flash-frozen and stored in liquid nitrogen. X-ray diffraction data were collected using the synchrotron beamline X10SA at the Swiss Light Source (SLS). Diffraction data were processed with the XDS program package [115]. The initial phases were determined by molecular replacement using PHASER [116]. PhPrx (PDB ID: 3W6G) [117] was used as a search model. Initial refinement was performed with the program REFMAC5 [118] and further refinement was done with PHENIX REFINE [119]. Necessary model improvements, as well as the search for solvent molecules, were carried out using COOT [120] and 'update water' in PHENIX REFINE. Anisotropic thermal displacement factors were refined at 1.8 Å resolution, otherwise using the TLS (translation, libration, and screw) model. Data collection, structure determination and refinement statistics are summarized in Appendix Table S1. The final atomic coordinates and structure factor amplitudes have been deposited in the RCSB Protein Data Bank, with the accession code ID 5OVQ.

3. Results

3.1 Isolation and purification of individual isoform of *A. aeolicus* complex I for further structural and functional study

Separation of these two isoforms has been very difficult due to their higher homology. In the native preparations from *A. aeolicus*, they are always co-eluted as a mixture in the same chromatographic fractions and appeared as one band on the IEF and BN-PAGE gel. Here two strategies were adopted with the attempt to separate two isoforms.

3.1.1 Separation of two isoforms from the native source using immunoprecipitation (IP)

To eliminate the co-precipitated contaminants and improve the binding efficiency, pure complex I obtained from *A. aeolicus* S1 sample was used as antigen. Polyclonal antibodies after affinity purification were used to precipitate the antigen. Anti-NuoI$_1$, or anti-NuoD$_2$ is expected to target NQOR1, and anti-NuoI$_2$, or anti-NuoD$_1$ is for NQOR2. The precipitated antigen was validated by western blot using these four antibodies. The expected results are shown in Table 3.1.

Table 3.1 The expected results of immunoprecipitation

IP / WB	NQOR1		NQOR2	
	Anti-NuoD$_2$	Anti-NuoI$_1$	Anti-NuoD$_1$	Anti-NuoI$_2$
Anti-NuoD$_2$	+	+	-	-
Anti-NuoI$_1$	+	+	-	-
Anti-NuoD$_1$	-	-	+	+
Anti-NuoI$_2$	-	-	+	+

(+) Positive, (-) Negative

Initial experiment was performed using the Pierce Classic IP Kit. Due to the co-elution of antibody with the immunoprecipitated antigen occurring with this kit, there could be at least three protein bands on a reducing SDS-PAGE or by western blot, they were the antibody heavy chain (50kDa), light chain (25kDa) and the antigen. The molecular mass of the four subunits used for testing were 23.41 kDa for NuoI$_1$, 24.57 kDa for NuoI$_2$,

68.69 kDa for $NuoD_1$, and 67.89 kDa for $NuoD_2$, respectively. Due to the fact that the molecular mass of $NuoI_1$ and $NuoI_2$ are very close to that of the antibody light chain, the antibody may mask the immunoprecipitated antigen. Therefore, Non-reducing LDS Sample Loading Buffer (Thermo Scientific) was used for SDS-PAGE sample preparation, which will just result in one band of antibody with molecular weight around 75 kDa. A molar ratio of Ab/Ag at 1:1 was tried first, with a total amount of 26 μg S1 sample mixed with 8 μg of each antibody. As shown in Figure 3.1, the band of 75 kDa antibody unexpectedly masked the band of $NuoD_1$ (68.69 kDa) and $NuoD_2$ (67.89 kDa). In addition, no binding was detected between anti-$NuoI_2$ and complex I, only weak binding was detected using anti-$NuoI_1$ (Figure 3.1).
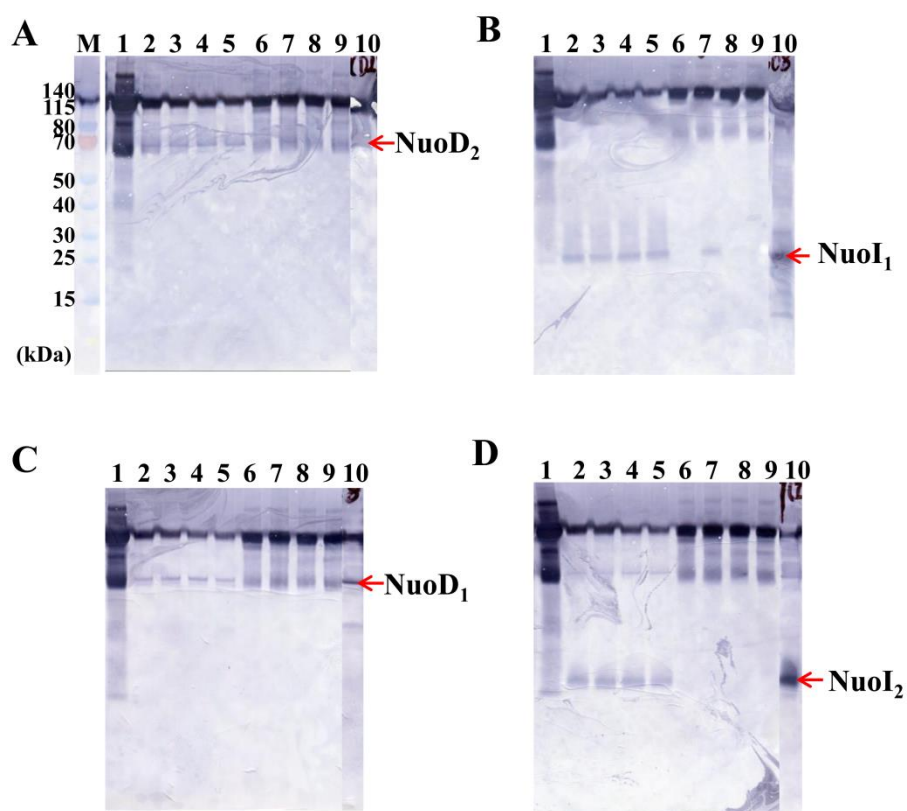
Figure 3.1. Western blot validations of complex I immunoprecipitation reactions using the Pierce Classic IP Kit.

The molar ratio of Ab/Ag is 1:1. The location of the individual subunit was labeled by red arrow on the right side of the gel. (A) Anti-$NuoD_2$ western blot, (B) Anti-$NuoI_1$ western blot, (C) Anti-$NuoD_1$ western blot, and (D) Anti-NuoI2 western blot. Lane M. PageRuler™ Prestained Protein Ladder, 10 to 180 kDa (Thermo Scientific). Lane 1. Antibodies as negative control, two major bands appeared around 70 kDa and

140 kDa, are approximately the molecular weight (75 kDa) or 2-fold molecular weight (150 kDa) of the antibody, respectively. Lane 2. Flow though from IP using anti-NuoD$_2$, lane 3. Flow though from IP using anti-NuoI$_1$, lane 4. Flow though from IP using anti-NuoD$_1$, lane 5. Flow though from IP using anti-NuoI$_2$; Lane 6. Elute fraction of IP using anti-NuoD$_2$, lane 7. Elute fraction of IP using anti-NuoI$_1$, lane 8. Elute fraction of IP using anti-NuoD$_1$, lane 9. Flow though of IP using anti-NuoI$_2$. Lane 10. Purified complex I S1 sample was used as positive control. Weak binding was detected using anti-NuoI$_1$.

To eliminate the co-elution of antibodies, the Pierce™ Co-Immunoprecipitation Kit was used for further experiments. An initial optimization was performed using anti-NuoI$_1$ and anti-NuoI$_2$. Different molar ratio of Ab/Ag (2:1, 5:1 and 10:1) was tested (Figure 3.2). The optimal Ab/Ag ratio was proved to be 10:1. Nevertheless, the binding was still too weak and most of the antigen appeared in the flow through. In addition, the protein precipitated by anti-NuoI$_1$ could be detected by all the four antibodies unexpectedly. The protein precipitated by anti-NuoI$_2$ could be detected by anti-NuoI$_2$ and anti-NuoD$_1$ as expected. No signal was detected by anti-NuoI$_1$. Extremely weak signals were detected by anti-NuoD$_2$. Immunoprecipitation using anti-NuoD$_1$ and anti-NuoD$_2$ with Ab/Ag molar ratio of 10:1 was also tested; however, no signal was detected by western blot. The experiment result was summarized in Table 3.2. It was promising to utilize IP to separate the individual isoform using anti-NuoI$_2$, nevertheless, further optimization is needed to improve the binding efficiency and specificity of the antibody to complex I.

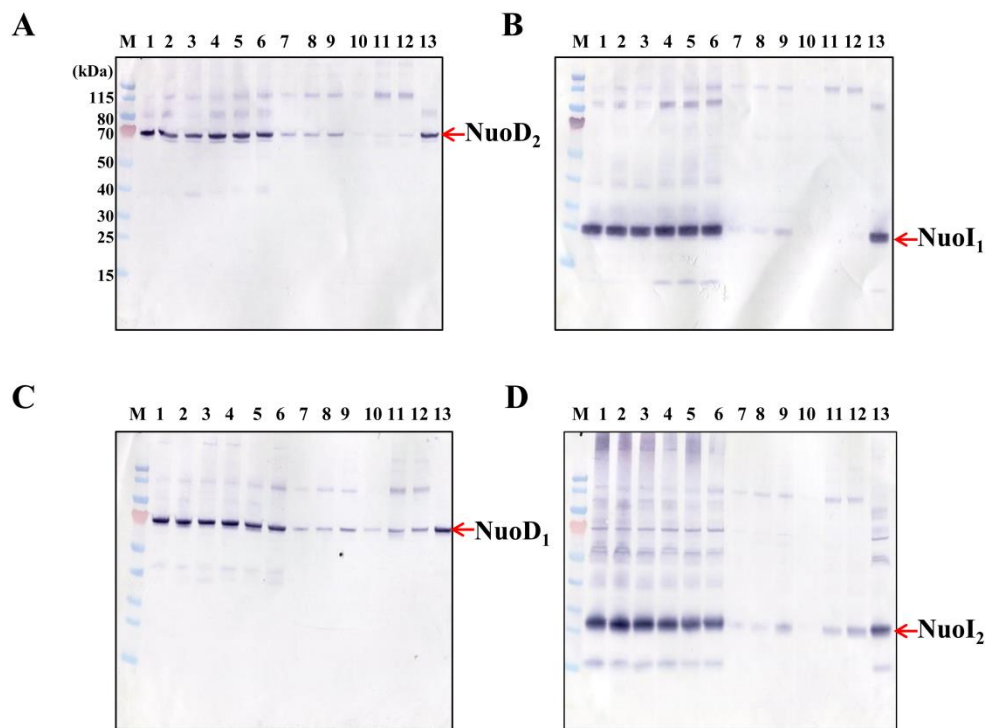**Figure 3.2**. **Western blot validations of complex I immunoprecipitation reactions after optimization.**
The location of the individual subunit was labeled. (A) Western blot using anti-NuoD$_2$. (B) Western blot using anti-NuoI$_1$. (C) Western blot using anti-NuoD$_1$. (D) Western blot using anti-NuoI$_2$. Lane M. PageRuler™ Prestained Protein Ladder, 10 to 180 kDa (Thermo Scientific). Lane 1-3. Flow through of IP using anti-NuoI$_1$ with Ab/Ag ratio of 2:1 for lane 1, 5:1 for lane 2, and 10:1 for lane 3. Lane 4-6. Flow through of IP using anti-NuoI$_2$ with Ab/Ag ratio of 2:1 for lane 4, 5:1 for lane 5, and 10:1 for lane 6. Lane 7-9. Elute fraction of IP using anti-NuoI$_1$ with Ab/Ag ratio of 2:1 for lane 7, 5:1 for lane 8, and 10:1 for lane 9. Lane 10-12. Elute fraction of IP using anti-NuoI$_2$ with Ab/Ag ratio of 2:1 for lane 10, 5:1 for lane 11, and 10:1 for lane 12. Most of the antigen appeared in the flow through. In addition, the protein precipitated by anti-NuoI$_1$ could be detected by all the four antibodies unexpectedly. The protein precipitated by anti-NuoI$_2$ could be detected by anti-NuoI$_2$ and anti-NuoD$_1$ as expected. No signal was detected by anti-NuoI$_1$. Extremely weak signals were detected by anti-NuoD$_2$.

Table 3.2 The experiment results of immunoprecipitation after optimization

IP WB	NQOR1		NQOR2	
	Anti-NuoD$_2$	Anti-NuoI$_1$	Anti-NuoD$_1$	Anti-NuoI$_2$
Anti-NuoD$_2$	-	-	-	-
Anti-NuoI$_1$	+	+	+	+
Anti-NuoD$_1$	-	-	-	-
Anti-NuoI$_2$	-	-	+	+

(+) Positive, (-) Negative

3.1.2 Heterologous production and characterization of individual isoform of *A. aeolicus* complex I in *E. coli*

Due to the high similarity of isoforms NQOR1 and NQOR2, it is nearly impossible to obtain them in a pure manner by purification from the natural source. In addition, genetic manipulations in *A. aeolicus* are not feasible yet, which considerably limits the research on the *A. aeolicus* complex I. Therefore, the heterologous production of *A. aeolicus* complex I in *E. coli* is strategically inescapable. The conception for producing the *A. aeolicus* complex I in *E. coli* is summarized in Table 3.3. In the first step, single subunits (NuoB, D_1, D_2, G, I_1, and I_2) were produced for characterization and comparison of the respective subunits from NQOR1 and NQOR2 (NuoD$_1$ vs NuoD$_2$, and NuoI$_1$ vs NuoI$_2$). In the second step, multi-gene operons were generated to produce complex I subcomplexes and finally the entire complex I for functional and structural analysis.

Table 3.3 The strategy used for production of *A. aeolicus* complex I

	Gene assembly	Cloning and Sequencing	Expression and Purification	MS identification	Characterization
Single gene	*nuo*B	✓	✓	✓	✓
	*nuo*D$_2$	✓	✓	✓	✗
	*nuo*I$_1$	✓	✓	✓	✓
	*nuo*G	✓	✓	✓	✓
	*nuo*D$_1$	✓	✓	✓	✗
	*nuo*I$_2$	✓	✓	✓	✓
Single operon	*nuo*BD$_1$	✓	✗	✗	✗
	*nuo*BD$_2$	✓	✗	✗	✗
	*nuo*BD$_2$I$_1$	✓	✗	✗	✗
	**nuo*EFG	✓	✗	✗	✗
	*nuo*BD$_2$I$_1$EFG	✓	✓	✓	✗
	**nuo*A$_2$BD$_2$H$_1$I$_1$J$_1$K$_1$L$_1$M$_1$N$_1$	✓	✓	✓	✓

* *nuo*A$_2$BD$_2$H$_1$I$_1$J$_1$K$_1$L$_1$M$_1$N$_1$ and *nuo*EFG were co-expressed to obtain the complex I NQOR1

3.1.2.1 Heterologous production and characterization of single subunits

3.1.2.1.1 Cloning and expression vector construction

The primer sequences used for gene cloning and vector construction for single subunits are shown in Table 2.12. Empty vectors and the corresponding expression vector are listed in Table 2.3 and 2.4.

The expression vectors used for single subunits (NuoB, D_1, D_2, G, I_1, and I_2) production was generated based on pET32b (Figure 3.3A). The genes *nuo*B, D_1, D_2, G, I_1, and I_2 were amplified from *A. aeolicus* genome by PCR using primers *p1f / r, p2f / r, p3f / r, p4f / r, p5f / r*, and *p6f / r*, respectively. The genes were further inserted into multiple cloning sites (MCS) of pET32b with *nuo*B at SacI / HindIII, *nuo*D_1, D_2 at BamHI / SalI, *nuo*G, I_1, and I_2 at EcoRI / HindIII, respectively. Additional expression vectors for NuoI$_2$ production were constructed on the basis of in-house modified pTTQ18 / pBAD / pQE expression vectors (Figure 3.3B). The gene *nuo*I_2 was amplified using primer *p7f / r*, and was inserted into pBAD-A2/C3, pTTQ-A2/C3 and pQE-A2/C3 at BamHI / EcoRI site.
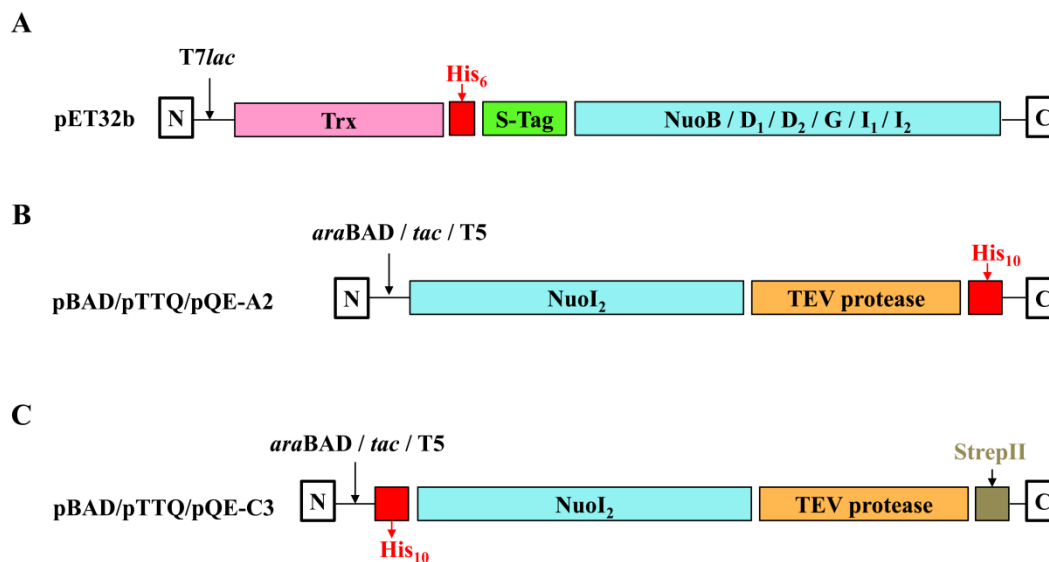
Figure 3.3. Vectors for single-gene expression.
Tags fused to the N- or C- termini of target proteins are colored in red for His$_{10}$ or His$_6$-Tag, tan for StrepII-Tag, pink for Trx-Tag, and green for S-Tag, respectively. TEV protease recognition site is shown in orange. (A) Based on the plasmid pET32b, expression vectors were constructed for NuoB, D_1, D_2, G, I_1, and I_2. (B) Based on pBAD / pTTQ / pQE-A2, expression vectors were constructed for NuoI$_2$. T7*lac* indicates T7

promoter / lac operator, while araBAD / tac / T5 an araBAD promoter for pBAD vectors (Invitrogen), indicates a moderately strong hybrid trp-lac (tac) promoter for pTTQ18 vectors [121], a strong T5 promoter for pQE vector (Qiagen) respectively.

3.1.2.1.2 Expression test and Purification of the single subunit

3.1.2.1.2.1 Production and purification of individual subunit using pET32b vector in BL21 (DE3)

Expression of single gene was verified by western blot and SDS-PAGE analysis with the expression of empty vector as a negative control. As shown in Figure 3.4, all six subunits have been expressed successfully.
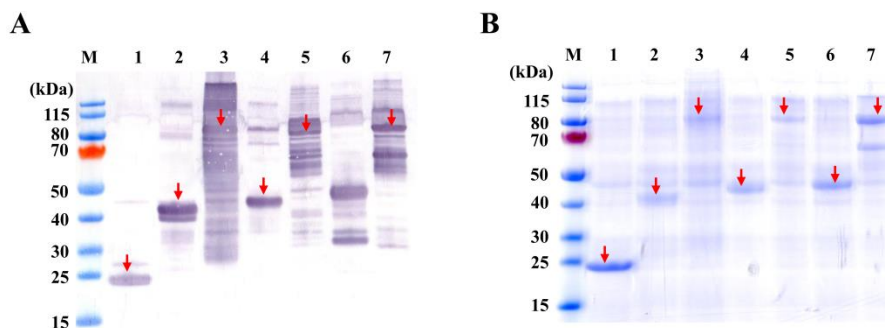
Figure 3.4. Single-gene expression tests.

(A) Western blot analysis using anti-His antibody. (B) SDS-PAGE analysis. M, PageRuler™ Prestained Protein Ladder, 10 to 180 kDa (Thermo Scientific). Lane 1, expression of empty vector pET32b with Trx-N-His detected, lane 2, expression of *nuo*B, lane 3, expression of *nuo*G, lane 4, expression of *nuo*I$_1$, lane 5, expression of *nuo*D$_1$, lane 6, expression of *nuo*I$_2$, and lane 5, expression of *nuo*D$_2$. The location of each subunit was labeled by red arrow.

The single subunit was further isolated and purified by a combination of Ni-NTA purification and size-exclusion chromatography. The quality of the protein was evaluated by SDS-PAGE and gel filtration. Pure and homogeneous NuoB could be obtained in the initial purification (Figure 3.5).
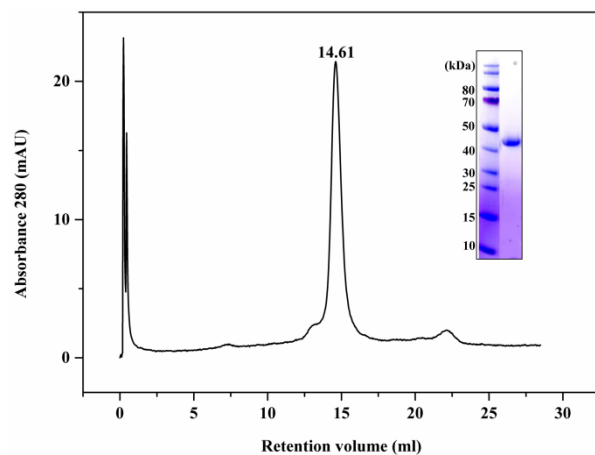
Figure 3.5. Purification profiles of NuoB.

The purity and homogeneity of NuoB was checked by gel filtration using the Superdex 200 10/300 GL column and SDS-PAGE analysis.

However, $NuoD_1$, D_2, G and I_1 were discovered to be aggregated during production (Figure 3.6), while most of $NuoI_2$ protein appeared in flow-through during His-trap purification (Figure 3.7).
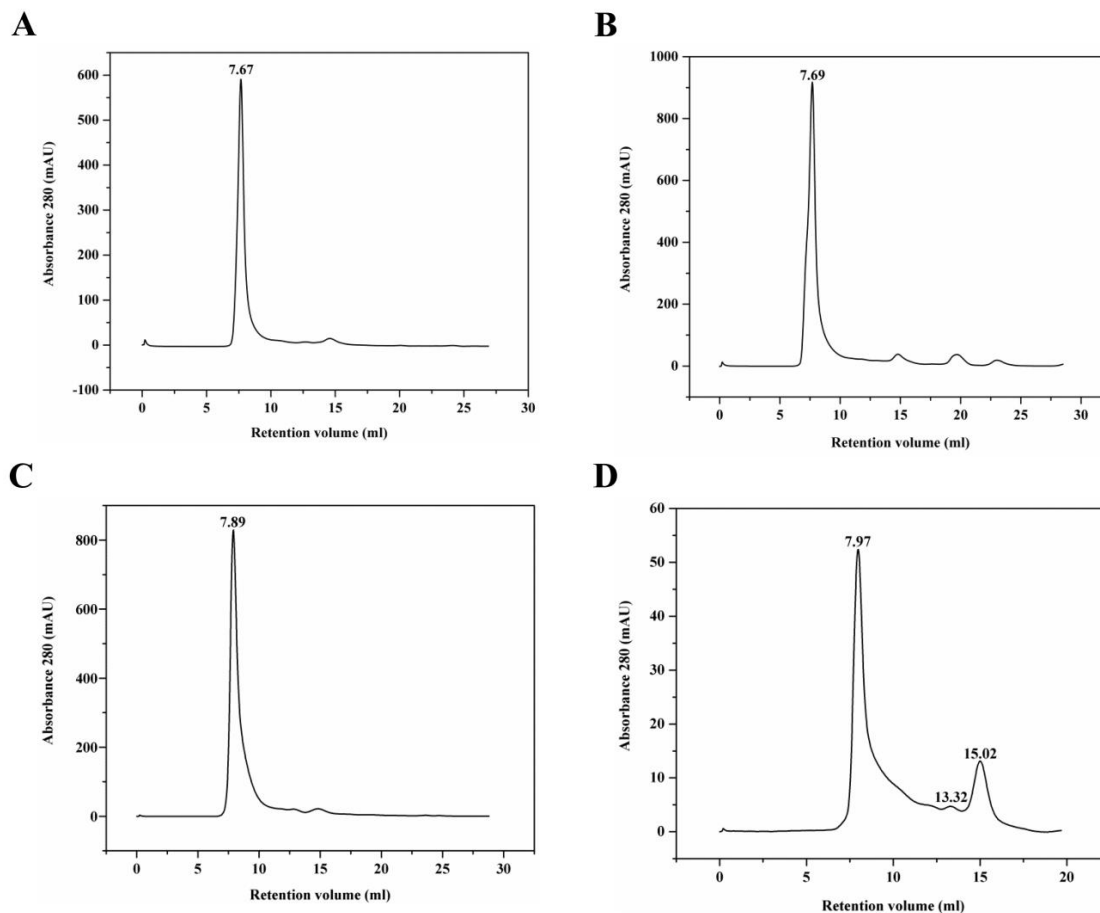
Figure 3.6. SEC profiles using Superdex 200 10/300 GL column.

(A) $NuoD_1$, (B) $NuoD_2$, (C) NuoG, and (D) $NuoI_1$. The proteins were eluted in the void volume which indicated the aggregation of proteins.
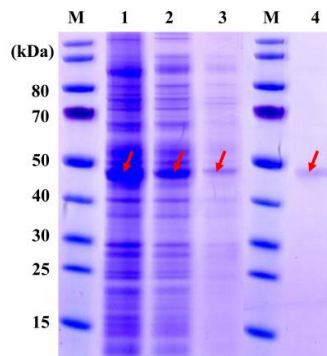
Figure 3.7. Ni-NTA purification of $NuoI_2$.

The location of $NuoI_2$ was labelled by red arrows. Lane 1, whole cell lysate, lane 2, flow through, lane 3, washing fractions, and lane 4, elution fractions. Most of the protein appeared in the flow through suggesting that the binding of $NuoI_2$ to the beads was not effective.

Considering that $NuoD_1$, D_2, I_1 and I_2 are components of connecting module of complex I, and are membrane-attached, 0.05% DDM was added to the buffer during purification. After optimization, pure and homogeneous $NuoI_1$ could be obtained (Figure 3.8). Most of the $NuoD_1$, D_2 still aggregated and there was no improvement for the binding of $NuoI_2$ to the Ni-NTA agarose as well.
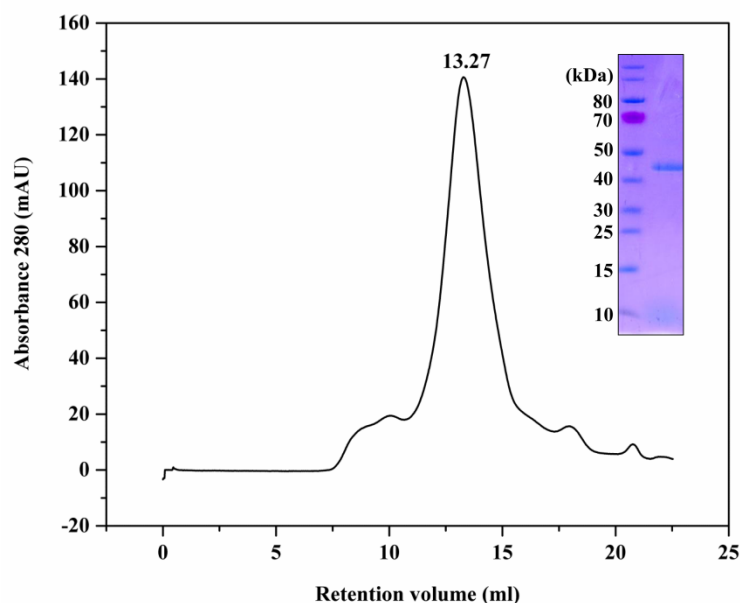
Figure 3.8. Purification profiles of $NuoI_2$.

The purified sample was analyzed by SEC using the Superdex 200 10/300 GL column and SDS-PAGE. The elution buffer used here contains 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 0.05% DDM.

3.1.2.1.2.1 Heterologous production, purification and crystallization of $NuoI_2$

$NuoI_2$ is a subunit of hydrophilic arm of *A. aeolicus* complex I NQOR2. The sequences of $NuoI_1$ and $NuoI_2$ share 50.7% identity. It was mentioned above (see 3.3.1.2.1), most of $NuoI_2$ appeared in flow through during Ni-NTA purification (see Figure 3.3), and therefore six other expression vectors were constructed with the attempt to obtain pure and homogeneous $NuoI_2$. Expression of the gene $nuoI_2$ was verified by western blot analysis (Figure 3.9 A), which indicated that $nuoI_2$ could be expressed using all six expression vectors. The pBAD-A_2-*nuo*I_2 / TOP10 system was utilized in the following steps. Due to the fact that the N-terminal helix of subunit I_1 was discovered to unexpectedly extend into the interface of the membrane in the crystal structure, the

location of NuoI$_2$ was verified by centrifugation and western blot analysis (Figure 3.9B). It was proved that NuoI$_2$ was located in the membrane fractions.

Figure 3.9. Expression test and protein location determination of NuoI$_2$.
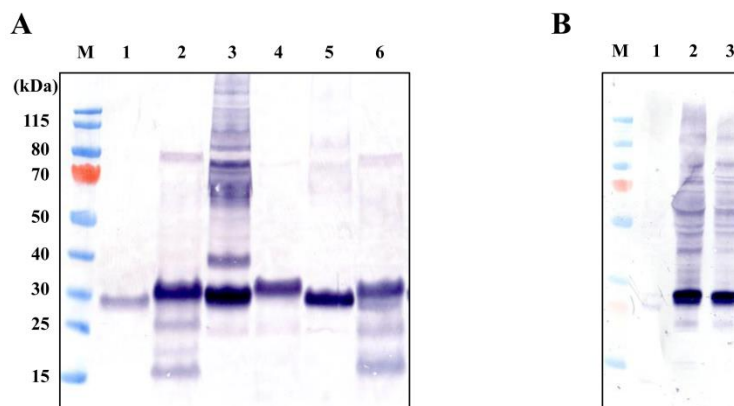(A) Western blot analysis of NuoI$_2$ expression. Lane 1, pQE-A2-*nuo*I$_2$ / C43 (DE3), lane 2, pQE-C3-*nuo*I$_2$ / C43 (DE3), lane 3, pBAD-A2-*nuo*I$_2$ / TOP10, lane 4, pBAD-A2-*nuo*I$_2$ / TOP10, lane 5, pTTQ-A2-*nuo*I$_2$ / NM554, and lane 6, pTTQ-C3-*nuo*I$_2$ / NM554. (B) Determination of the protein location. Lane 1, supernatant, lane 2, membrane, lane- 3, cell pellet.

The membrane of pBAD-A$_2$-*nuo*I$_2$ / TOP10 cell (5 mg/ml) was solubilized using 1% (w/v) DDM at RT for 2 hour and ultra-centrifuged to remove undissolved debris. NuoI$_2$ was purified by a combination of Ni-NTA and SEC purification. The purified protein sample dissolved in the buffer containing 50 mM Tris-HCl, 150 mM NaCl, and 0.05% DDM was analyzed by SDS-PAGE and western blot. An evident contaminate was detected, which was proved to be a chaperonin protein by MS identification (Figure 3.10). To eliminate contaminate and obtain pure NuoI$_2$, various detergents (see chapter 2.2.2.4) were checked during membrane solubilization and Ni-NTA purification (Figure 3.11). Contaminate could be removed by using detergent FOS12 to solubilize and purify the protein. The purity and homogeneity of NuoI$_2$ was further checked by gel filtration using the Superdex 200 10/300 GL column and SDS-PAGE analysis (Figure 3.12). It indicated that pure and homogeneous NuoI$_2$ could be obtained.
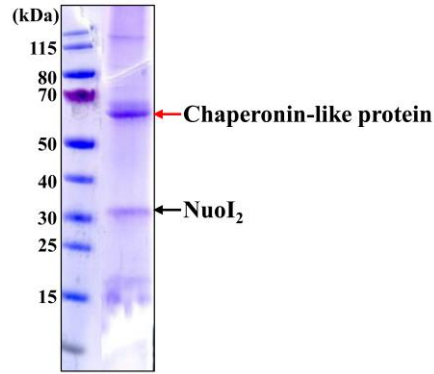
Figure 3.10. SDS-PAGE analysis of purified NuoI$_2$.

Membrane solubilization and protein purification was performed using detergent DDM. The location of NuoI$_2$ and contaminant were labeled with black and red arrow, respectively.
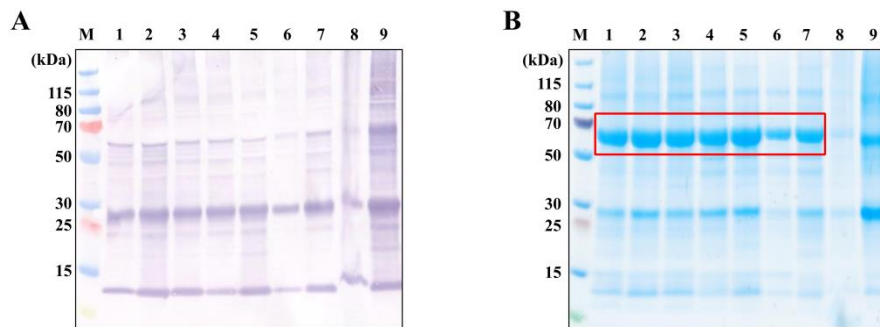
Figure 3.11. Detergent screening.

(A) Western blot analysis of Ni-NTA purification of NuoI$_2$ using buffer with various detergents. (B) SDS-PAGE analysis of Ni-NTA purification of NuoI$_2$ using buffer with various detergents. (A) and (B), Lane 1, β-DM, lane 2, β-DDM, lane 3, cymal6, lane 4, DM-NG, lane 5, LM-NG, lane 6, OG-NG, lane 7, C12E8, lane 8, LDAO, lane 9, FOS12. Contaminant could be detected in the SDS-PAGE gel using Coomassie staining (highlighted in the red frame), but could not be detected by western blot.
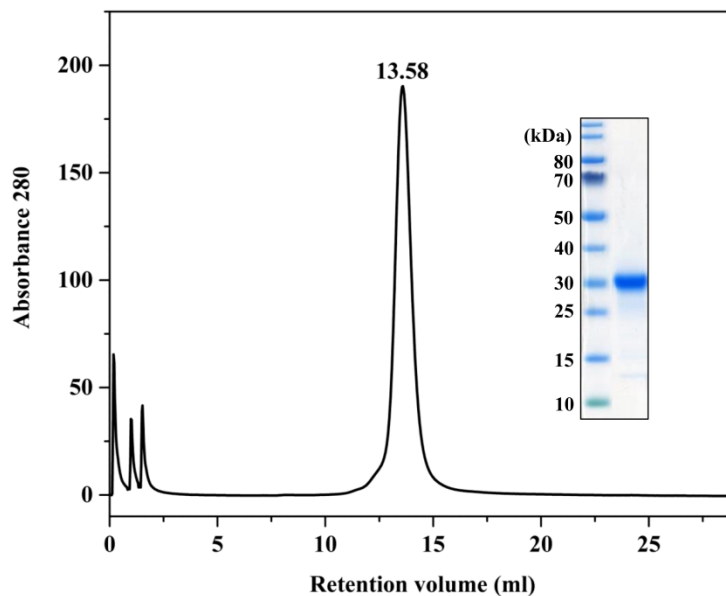
Figure 3.12. Purification profiles of NuoI$_2$.
The purity and homogeneity of NuoI$_2$ was checked by gel filtration using the Superdex 200 10/300 GL column and SDS-PAGE analysis.

3.1.2.2 Heterologous production of subcomplexes of NQOR1

It was reported before that the fully assembled NADH dehydrogenase fragment can be overproduced in *E. coli* when the genes nuoE, F, and G were simultaneously overexpressed with the genes nuoB, C, and D [24]. Here we perform heterologous coexpression of *A. aeolicus nuo*B, D$_2$, E, F, G and I$_1$ in *E. coli* with the attempt to obtain a fully assembled NADH dehydrogenase fragment (NuoEFG) to conduct functional analysis.

3.1.2.2.1 Construction of the artificial operon and expression vector

The genes *nuo*BD$_2$ (including 37 bp of *nuo*B upstream) were cloned into pBAD-A2, resulting in the expression vector pBAD-A2-BD$_2$. The gene *nuo*I$_1$ was further insert into pBAD-A2-BD$_2$, generating pBAD-A2-BD$_2$I$_1$. Meanwhile the genes *nuo*EF and *nuo*G were amplified from the genome DNA individually and cloned into pJET1.2 successively. The resulted operon *nuo*EFG was sub-cloned into pBAD-A2-BD$_2$I$_1$, generating the final

expression vector containing an artifitial operon nuoBD$_2$EFGI$_1$ (Figure 3.13). The gene of StrepII-Tag was inserted at the C-teminus of nuoF.
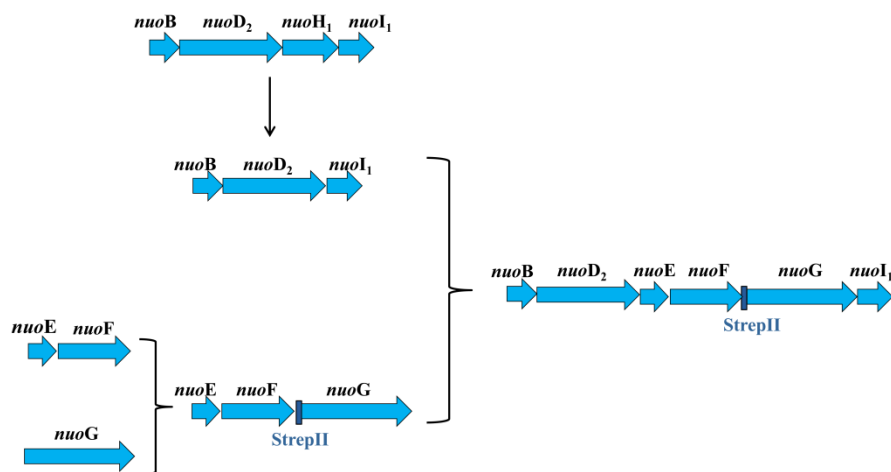
Figure 3.13. Construction of the artificial operon nuoBD$_2$EFGI$_2$.

3.1.2.2.2 Purification of sub-complexes produced by pBAD33-nuoBD$_2$EFGI$_1$

Purification was performed using Strep-Tactin purification followed by a gel filtration. The purified protein was analysed by SDS-PAGE, western blot and BN-PAGE (Figure 3.14). The composition of the sub-complexes was further confirmed by MS identification. Unexpectedly, a sub-complexes composed by NuoE and NuoF was obtained from expression of pBAD33-nuoBD$_2$EFGI$_1$ in $E.coli$ TOP10 cell which has been produced by overexpressed the $A.$ $aeolicus$ genes nuoE, F, and G heterologously in $E.$ $coli$ and reported previously [122]. NuoG was not detected in the purified protein sample.
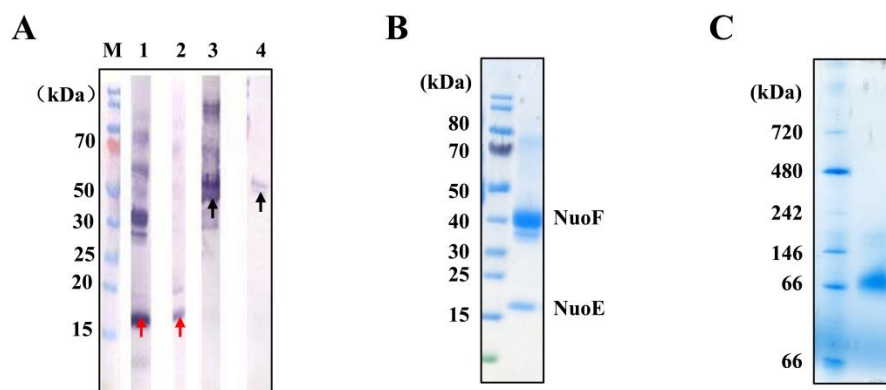
Figure 3.14. Purification of sub-complexes produced by pBAD-A2-*nuo*BD$_2$EFGI$_1$.

(A) Western blot analysis. Lane 1, 2 western blot detection by Anti-NuoE, lane 3, 4, western blot detection by Anti-NuoF. Lane 1, 3, sample produced and purified using pBAD33-NuoBD$_2$EFGI$_1$/Top10. Lane 2, 4, purified native complex I sample from *A. aeolicus*. The location of NuoE was labelled by red arrows, NuoF labelled by black arrows. (B) SDS-PAGE analysis by Coomassie staining. The location of NuoE and NuoF was labelled. (C) BN-PAGE analysis.

3.1.2.3 Heterologous production and characterization of entire *A. aeolicus* complex I NQOR1

3.1.2.3.1 Construction of expression vector

Given that the genes encoding complex I NQOR1 are dispersed in three loci (see chapter 1.4) and the total size of the *nuo* genes is large (~ 14 kb), the strategy of co-expression using two expression vectors was adopted. The information of the basic vectors was summarized in Table 3.4

Construction of the expression vector was shown in Figure 3.15. The artificial operon of *nuo*EFG was sub-cloned into ORF2 of pBAD-CM1 from pJET-*nuo*EFG. The 3' terminus of *nuo*G was fused with the StrepII-Tag of pBAD-CM1. The operon of *nuo*A$_2$BD$_2$H$_1$I$_1$J$_1$K$_1$L$_1$M$_1$N$_1$ was amplified from *A. aeolicus* genome DNA and inserted into pBAD33 by In-Fusion reaction. The resulting expression vectors were verified by sequencing and co-transformed into *nuo* deletion strain BA14 for inductive expression.

Table 3.4 Information of the basic vectors used

Plasmid	Replicon	Promoter	Resistance
pBAD33	f1_origin/ p15A_origin	*Ara*C-promoter	Chloramphenicol
pBAD-CM1	pBR322 origin	*Ara*C-promoter	Ampicillin
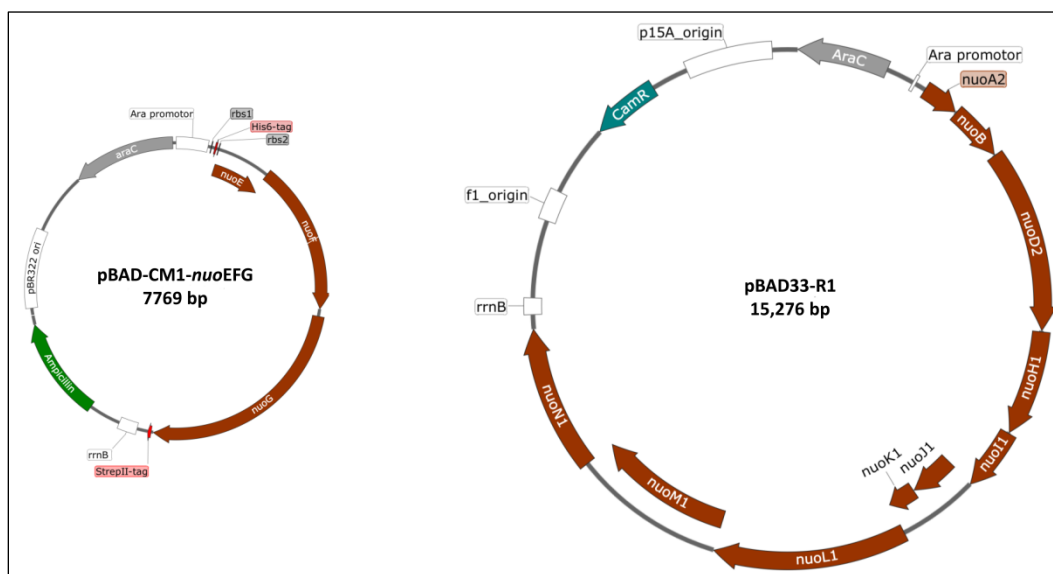
Figure 3.15. Expression vectors generated for production of NQOR1.

The genes labelled on the vectors are colored differently, with *nuo* genes as brown, *Ara*C as grey, Amp as green and CamR as teal. The artificial operon of *nuo*EFG was sub-cloned into the ORF2 of pBAD-CM1, resulting in the expression vector named pBAD-CM1-*nuo*EFG (left side). The operon of $nuoA_2BD_2H_1I_1J_1K_1L_1M_1N_1$ amplified from the genome of A. aeolicus was inserted into pBAD33, generating expression vector named pBAD33-R1 (right side). The ribosome binding site (RBS) of *nuo* genes are original from *A. aeolicus*.

3.1.2.3.2 Purification of heterologously produced NQOR1

The purification of heterologously produced NQOR1 involved two successive purification steps, anion exchange chromatography and Strep-Tactin purification. The anion exchange chromatography was carried out in a similar manner to the established protocol for native *A. aeolicus* complex I using a mono Q 10/100 GL column [103]. The anion exchange chromatograms are shown in Figure 3.16A. The fractions from the mono Q column were tested using NADH dehydrogenase activity staining (Figure 3.16B). The

fractions with NADH dehydrogenase activity were collected, combined and concentrated for Strep-Tactin purification. The contaminants were removed by washing buffer containing 20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 2 mM DTT, 1 mM EDTA and 0.05% DDM, and the target protein was eluted using elution buffer additionally containing 5 mM d-Desthiobiotin. Afterwards, the purified strep-tagged complex I was verified by BN-PAGE and in gel activity (Figure 3.17). Unexpectedly, six bands were present in the native gel, which all exhibited NADH dehydrogenase activity.
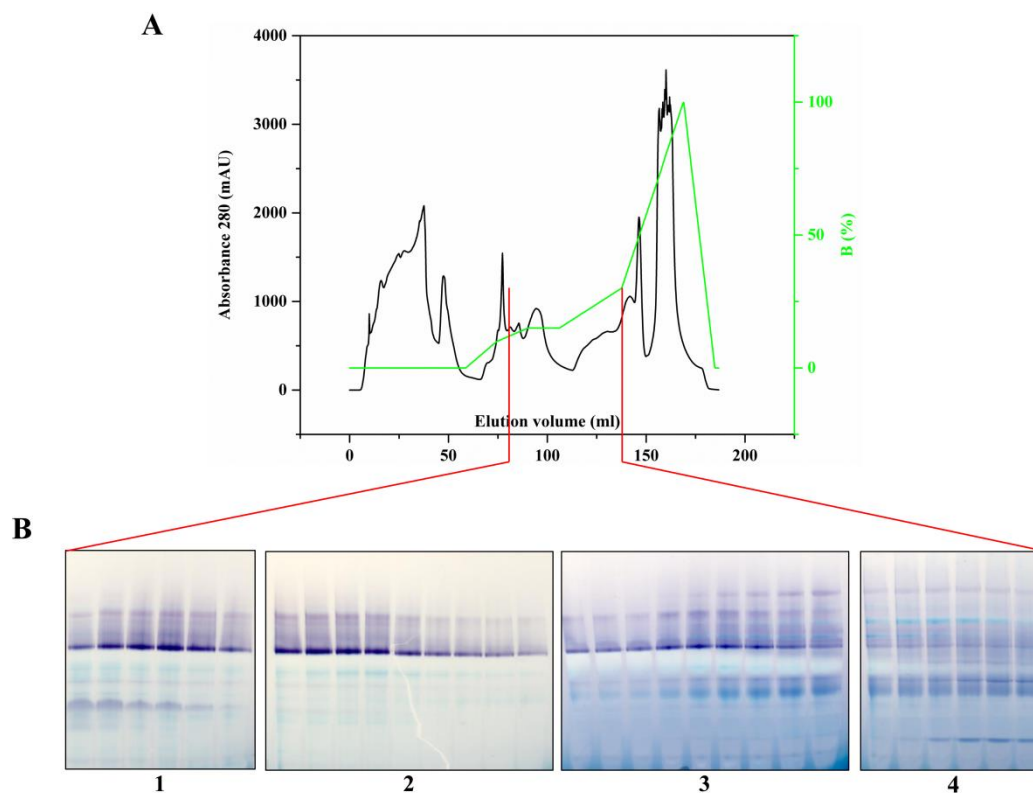

Figure 3.16. Purification of NQOR1 using mono Q 10/100 GL column.
(A) Anion exchange chromatograms (B) Fractions eluted from mono Q column with in gel activity. Proteins having NADH dehydrogenase activity could be stained, distinct from the ones without activity in color.
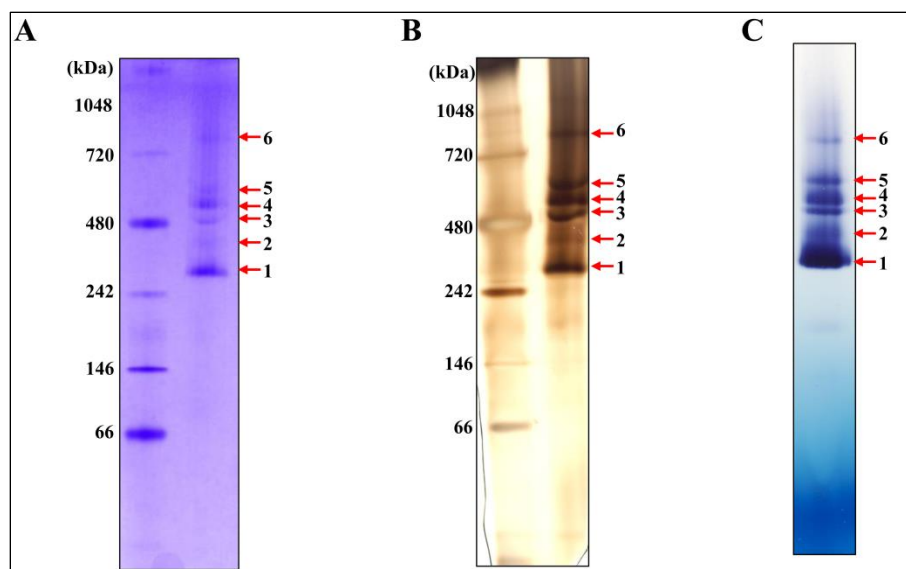
Figure 3.17. Purified NQOR1 validated by BN-PAGE and NADH dehydrogenase activity staining.
(A) BN-PAGE of purified NQOR1 stained by Coomassie Brilliant Blue staining solution. (B) Silver staining of BN-PAGE. (C) In gel activity of purified NQOR1. The bands present in the gel are shown by red arrows and labeled. The labels on (A), (B), (C) are corresponding to each other.

3.1.2.3.3 MS identification

The bands on the native gel, named band 1-6 with the increase of molecular weight, were sliced and the composition of each band was determined using mass spectrometry (see chapter 3.3.2.4). Several unique peptides corresponding to all known complex I subunits were identified from the Native Page slices. Six subunits (NuoB, D_2, E, F, G, and I_1) composed of *A. aeolicus* complex I hydrophilic arm were detected in band 1 (Table 3.5). Seven subunits were detected in band 2, in addition to the subunits from hydrophilic arm; $NuoH_1$ from membrane domain was identified as well (Table 3.6). In band 3, 11 subunits were detected, including 6 subunits from hydrophilic arm and 5 subunits ($NuoH_1$, A_2, J_1, K_1, and N1) from membrane domain (Table 3.7). In band 4, 12 subunits were identified, with all of 11 subunits detected in band 3 as well as an extra subunit $NuoL_1$ (Table 3.8). All 13 subunits of *A. aeolicus* complex I were identified in band 5, which indicated a full-assembled complex (Table 3.9). There was no any protein from *E. coli* was detected in band 1-5. The compositions of band 6 was complicated and interesting (Appendix Table S2), in which 7 subunits (NuoB, D_2, E, F, G, I_1 and H_1) from *A. aeolicus* complex I were

identified, together with proteins from *E. coli*, including ribosomal proteins, chaperone proteins, *et al*.

Table 3.5 MS identification of BN-PAGE gel band 1

Description	Coverage [%]	Peptides	PSMs	Unique Peptides	AAs	MW (kDa)
NADH dehydrogenase I subunit D [OS=*Aquifex aeolicus* VF5]	81	51	284	51	586	67.9
NADH dehydrogenase I subunit G [OS=*Aquifex aeolicus* VF5]	75	53	224	53	632	72.7
NADH dehydrogenase I subunit F [OS=*Aquifex aeolicus* VF5]	76	35	210	35	426	47.5
NADH dehydrogenase I subunit I [OS=*Aquifex aeolicus* VF5]	58	15	34	15	201	23.4
NADH dehydrogenase I subunit B [OS=*Aquifex aeolicus* VF5]	87	12	32	12	179	19.9
NADH dehydrogenase I subunit E [OS=*Aquifex aeolicus* VF5]	60	11	32	11	160	18.5

Total mass: 249.9 kDa

Table 3.6 MS identification of BN-PAGE gel band 2

Description	Coverage [%]	Peptides	PSMs	Unique Peptides	AAs	MW [kDa]
NADH dehydrogenase I subunit G [OS=*Aquifex aeolicus* VF5]	72	42	119	42	632	72.7
NADH dehydrogenase I subunit D [OS=*Aquifex aeolicus* VF5]	72	39	95	39	586	67.9
NADH dehydrogenase I subunit F [OS=*Aquifex aeolicus* VF5]	54	23	42	23	426	47.5
NADH dehydrogenase I subunit B [OS=*Aquifex aeolicus* VF5]	53	8	12	8	179	19.9
NADH dehydrogenase I subunit E [OS=*Aquifex aeolicus* VF5]	48	7	10	7	160	18.5
NADH dehydrogenase I subunit I [OS=*Aquifex aeolicus* VF5]	25	5	9	5	201	23.4
NADH dehydrogenase I chain H [OS=*Aquifex aeolicus* VF5]	11	2	2	2	336	36.9

Total mass: 286.8 kDa

Table 3.7 MS identification of BN-PAGE gel band 3

Description	Coverage [%]	Peptides	PSMs	Unique Peptides	AAs	MW [kDa]
NADH dehydrogenase I subunit G [OS=*Aquifex aeolicus* VF5]	77	50	287	50	632	72.7
NADH dehydrogenase I subunit D [OS=*Aquifex aeolicus* VF5]	78	41	164	41	586	67.9
NADH dehydrogenase I subunit F [OS=*Aquifex aeolicus* VF5]	63	26	57	26	426	47.5
NADH dehydrogenase I subunit B [OS=*Aquifex aeolicus* VF5]	72	10	20	10	179	19.9
NADH dehydrogenase I subunit I [OS=*Aquifex aeolicus* VF5]	49	11	21	11	201	23.4
NADH dehydrogenase I subunit E [OS=*Aquifex aeolicus* VF5]	48	7	16	7	160	18.5
NADH dehydrogenase I chain H [OS=*Aquifex aeolicus* VF5]	28	6	11	6	336	36.9
NADH dehydrogenase I subunit N1 [OS=*Aquifex aeolicus* VF5]	25	6	10	6	464	51.3
NADH dehydrogenase I chain A [OS=*Aquifex aeolicus* VF5]	19	2	3	2	118	13.3
NADH dehydrogenase I chain J [OS=*Aquifex aeolicus* VF5]	35	6	7	6	162	17.8
NADH dehydrogenase I subunit K1 [OS=*Aquifex aeolicus* VF5]	15	3	5	3	100	10.7

Total mass: 379.9 kDa

Table 3.8 MS identification of BN-PAGE gel band 4

Description	Coverage [%]	Peptides	PSMs	Unique Peptides	AAs	MW [kDa]
NADH dehydrogenase I subunit G [OS=*Aquifex aeolicus* VF5]	73	51	262	51	632	72.7
NADH dehydrogenase I subunit D [OS=*Aquifex aeolicus* VF5]	80	46	183	46	586	67.9
NADH dehydrogenase I subunit F [OS=*Aquifex aeolicus* VF5]	71	30	79	30	426	47.5
NADH dehydrogenase I subunit I [OS=*Aquifex aeolicus* VF5]	56	17	29	17	201	23.4
NADH-quinone oxidoreductase subunit B [OS=*Aquifex aeolicus* VF5]	72	12	17	12	179	19.9
NADH-quinone oxidoreductase subunit E [OS=*Aquifex aeolicus* VF5]	48	6	11	6	160	18.5
NADH dehydrogenase I chain H [OS=*Aquifex aeolicus* VF5]	29	7	9	7	336	36.9
NADH dehydrogenase I subunit N1 [OS=*Aquifex aeolicus* VF5]	25	4	6	4	464	51.3
NADH dehydrogenase I chain A [OS=*Aquifex aeolicus* VF5]	19	2	3	2	118	13.3
NADH dehydrogenase I chain J [OS=*Aquifex aeolicus* VF5]	35	6	7	6	162	17.8
NADH dehydrogenase I K1 [OS=*Aquifex aeolicus* VF5]	43	2	3	2	100	10.7
NADH dehydrogenase I chain L [OS=*Aquifex aeolicus* VF5]	4	1	2	1	622	68.9

Total mass: 448.8 kDa

Table 3.9 MS identification of BN-PAGE gel band 5

Description	Coverage [%]	Peptides	PSMs	Unique Peptides	AAs	MW [kDa]
NADH dehydrogenase I subunit G [OS=*Aquifex aeolicus* VF5]	72	47	335	47	632	72.7
NADH dehydrogenase I subunit D [OS=*Aquifex aeolicus* VF5]	75	39	144	39	586	67.9
NADH dehydrogenase I subunit F [OS=*Aquifex aeolicus* VF5]	73	28	58	28	426	47.5
NADH dehydrogenase I subunit I [OS=*Aquifex aeolicus* VF5]	47	10	19	10	201	23.4
NADH dehydrogenase I subunit B [OS=*Aquifex aeolicus* VF5]	72	10	17	10	179	19.9
NADH dehydrogenase I subunit E [OS=*Aquifex aeolicus* VF5]	48	7	13	7	160	18.5
NADH dehydrogenase I chain H [OS=*Aquifex aeolicus* VF5]	31	6	10	6	336	36.9
NADH dehydrogenase I chain L [OS=*Aquifex aeolicus* VF5]	19	9	13	9	622	68.9
NADH dehydrogenase I chain M [OS=*Aquifex aeolicus* VF5]	22	8	11	8	491	55
NADH dehydrogenase I subunit N1 [OS=*Aquifex aeolicus* VF5]	41	9	10	9	464	51.3
NADH dehydrogenase I subunit K1 [OS=*Aquifex aeolicus* VF5]	65	6	9	6	100	10.7
NADH dehydrogenase I chain A [OS=*Aquifex aeolicus* VF5]	19	2	4	2	118	13.3
NADH dehydrogenase I chain J [OS=*Aquifex aeolicus* VF5]	44	4	5	4	162	17.8

Total mass: 503.8 kDa

Table 3.10 summary of the MS identification

Nuo Band	Hydrophilic domain						Hydrophobic domain						
	B	D_1	E	F	G	I_2	A_1	H_1	J_1	K_1	L_1	M_1	N_1
1*	+	+	+	+	+	+	-	-	-	-	-	-	-
2*	+	+	+	+	+	+	-	+	-	-	-	-	-
3*	+	+	+	+	+	+	+	+	+	+	-	-	+
4*	+	+	+	+	+	+	+	+	+	+	+	-	+
5*	+	+	+	+	+	+	+	+	+	+	+	+	+
6**	+	+	+	+	+	+	-	+	-	-	-	-	-

* Band 1-5, no protein from *E. coli* was detected.

** Band 6, proteins from *E. coli* was detected.

3.1.2.3.4 NADH:ubiqunione oxidoreductase activity assay

After strep-tactin purification, NQOR1 was changed to buffer 20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 2 mM DTT, 1 mM EDTA and 0.05% DDM to remove the desthiobiotin. The activity of *A. aeolicus* complex I NQOR1 was measured as described with slight modification [103]. The reaction buffer containing 20 mM Tris-HCl, 100 mM NaCl, 10% glycerol, 0.1% CHAPS and 500 μM KCN was used here. The ubiquinone analog DL01 was used as the electron acceptor. The decrease of Absorbance 340 nm was monitored (Figure 3.18).
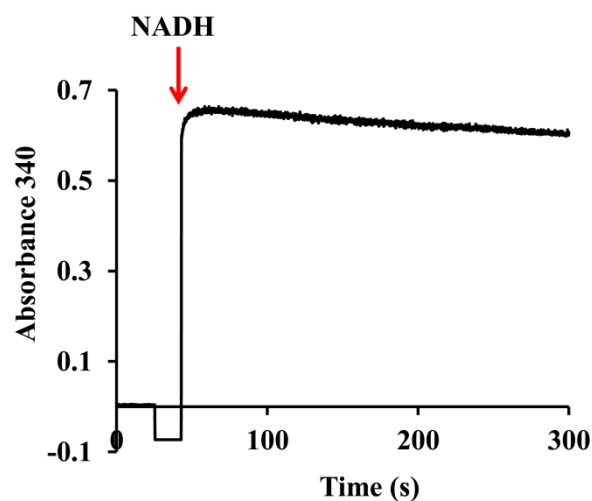

Figure 3.18. Time-course oxidation of NADH by the electron acceptors DL01.

3.1.2.3.5 UV-VIS spectra of complex I NQOR1 under different redox conditions

NQOR1 was characterized by UV-VIS spectrophotometer under different redox state (Figure 3.19)
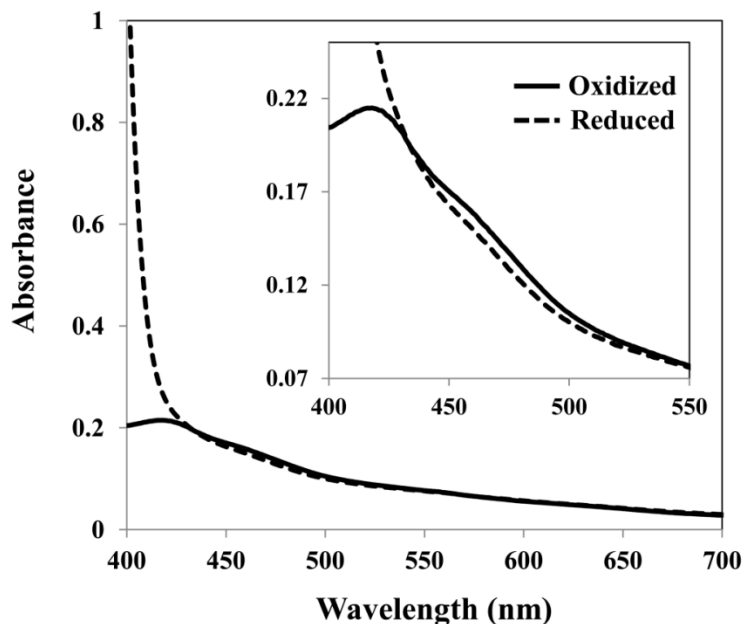
Figure 3.19. UV-VIS absorption spectra of purified complex I NQOR1.

Complex I was dissolved in 20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 2 mM DTT, 1 mM EDTA and 0.05% DDM at 0.91 mg/ml. Inset, the enlarged absorption spectra from 400 nm to 550 nm. The oxidized form, thick solid line and the reduced form by the addition of 10 mM NADH, dotted line.

3.1.2.3.6 Electron Microscopy of the Negatively Stained Particles and image processing

The purified NQOR1 was analyzed by electron microscopy. L-shaped particles could be observed in the electron microscope. A gallery of particles, selected for image processing, was shown in Figure 3.20A.

A

500 Å

B

6.2%

171Å

142Å

450Å

Figure 3.20. Electron micrographs of heterologously produced NQOR1.

(A) Images obtained by negative staining. A gallery of L-shaped particle selected for image processing. (B) Representative 2D class ave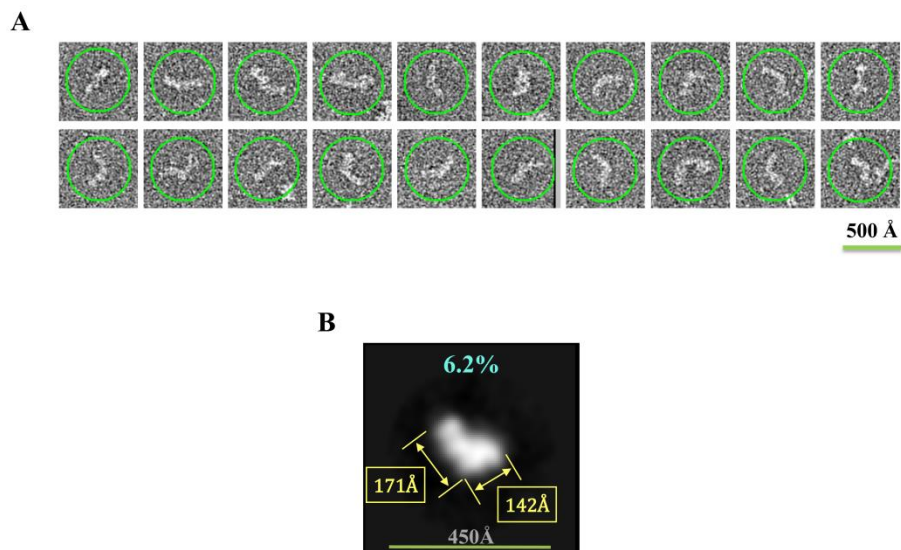rage views. The length of two arms of complex I was labelled. The figure was made by Hailong Gao in the group of Prof. Zihe Rao, Institute of Biophysics, Chinese Academy of Science.

3.2 Identification of the Fe-S clusters by EPR measurement

Complex I purified from *A. aeolicus*, single subunits and entire NQOR1 produced in *E.coli* were used for EPR measurement to obtain the properties of Fe-S clusters.

3.2.1 Identification of the Fe-S clusters in complex I purified from *A. aeolicus*

The purification of complex I from *A. aeolicus* membrane was performed according to the protocol described as before [103], which involve an anion-exchange step followed by size exclusion chromatography. The purified protein was dissolved in the buffer containing 20- mM Tris-HCl, pH 7.4, 150 mM NaCl and 0.05% DDM. The quality of the native preparations was evaluated by in gel activity assay (Figure 3.21).
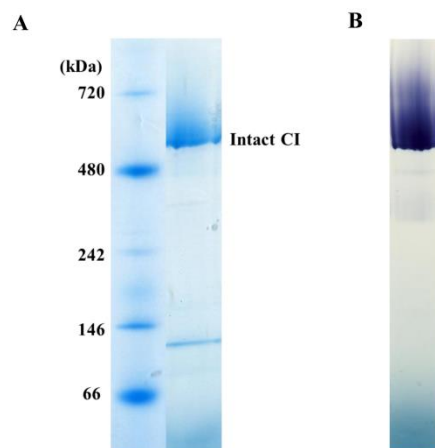
Figure 3.21. **BN-PAGE and NADH dehydrogenase activity staining.**

(A) BN-PAGE using NativePAGE™ 4-16% Bis-Tris Protein Gels (Thermo Scientific). The location of the intact complex I (CI) with a molecular weight around 500 kDa was labeled on the right side. (B) NADH dehydrogenase activity staining. BN-PAGE gel incubated with 2.5 mg/ml NBT and 150 μM NADH for 30 mins at RT.

The activity of the purified *A. aeolicus* complex I was assayed by monitoring electron transfer from NADH to ubiquinone at 80 ℃ according the protocol described previously [103]. Purified complex I was obtained from three different sample preparations (S1, S2, and S3) [123] was test accordingly. In addition to the commercially available decyubiquinone (DQ, Sigma Aldrich) which is ubiquitously used, two of *A. aeolicus* quinone (2-demethylmenaquinone-7) [107] analogs, DMK-S2 and DL01 were also used as electron acceptors. The inhibition of the enzyme activity by rotenone is evaluated as well. As shown in Table 3.11, DL01 exhibited a 70–80% higher specific activity compared to the other two quinones. The activity of complex I using DMK-S2 was unexpected lower, which may be due to its poor solubility. Furthermore, complex I from sample S2 present a higher specific activity compared to S1 and S3. The activity assay was performed by Jana Juli from Johann Wolfgang Goethe Universität

Table 3.11 Specific activity of *A. aeolicus* complex I using different electron acceptors

Sample	Acceptors	Specific activity ($\mu mol \cdot min^{-1} mg^{-1}$)	Inhibition by rotenone (%)
S1	DQ	0.46 ± 0.02	71.5%
	DMN-K2	0.32 ± 0.01	48.7%
	DL01	1.60 ± 0.05	54.7%
S2	DQ	0.44 ± 0.02	66.3%
	DMN-K2	0.46 ± 0.03	43.6%
	DL01	2.33 ± 0.10	66.3%
S3	DQ	0.29 ± 0.03	66.6%
	DMN-K2	0.23 ± 0.02	66.0%
	DL01	0.86 ± 0.09	71.1%

The table was adapted from the thesis "Solubilization experimental studies on the supercomplex formation of the respiratory complex I in a hyperthermophilic bacterium *Aquifex Aeolicus*" by Jana Juli.

Intact complex I from sample S2 revealed a highly overlapped EPR spectrum at 12 K originating from several paramagnetic centers (Figure 3.22A). At higher temperature (40 K) the axial spectrum of a binuclear cluster (N1b) could be identified ($g_{z,y,x}$ = 2.026, 1.938, 1.931). The EPR signal of cluster N2 ($g_{z,y,x}$ = 2.051, 1.909, 1.897) was obtained after addition of dithionite to a sample of isolated complex I and freezing after short reaction time (10s). Signal intensities at g = 2.077 and g = 1.95 revealed similar temperature dependence (Figure 3.22B). These signals were assumed to belong to a third paramagnetic center which we would, tentatively assign to cluster N4 according to its temperature behavior. The temperature and microwave power dependent EPR spectra (Figure 3.22B and C) did not reveal any clear additional signals that could be attributed to further paramagnetic centers. Especially there were neither indications for a second binuclear cluster (N1a, $g_{z,y,x}$ = 2.00, 1.95, 1.92) at higher temperature like in the enzyme from *E. coli* [52] nor a fast relaxing cluster resulting in an N5 signal ($g_{z,y,x}$ = 2.07, 1.93, 1.90) as known from bovine or *Yarrowia lipolytica* complex I [49, 124]. Obviously the prominent signal in the field range 350–360mT contains contributions of several iron sulfur clusters but unfortunately the relative low concentrations of actually available

complex I samples did not allow a more elaborated deconvolution of spectra by varying reducing conditions, microwave power and temperature.

Redox titration of complex I from *A. aeolicus* S2 sample at pH 7.4 (Figure 3.23), resulted in redox midpoint potentials for the binuclear cluster N1b, $E_{m,7.4} = -274$ mV and for the tetranuclear N2, $E_{m,7.4} = -184$ mV. The second identified EPR signal of a tetranuclear cluster at g = 2.077 was hard to evaluate in redox titration experiments due to strong background effects in this field region, but a rough estimation yielded a midpoint potential of approximately -257 mV at pH 7.4. These values are in the same potential range as those reported in the literature for the *E. coli* enzyme [52, 125].
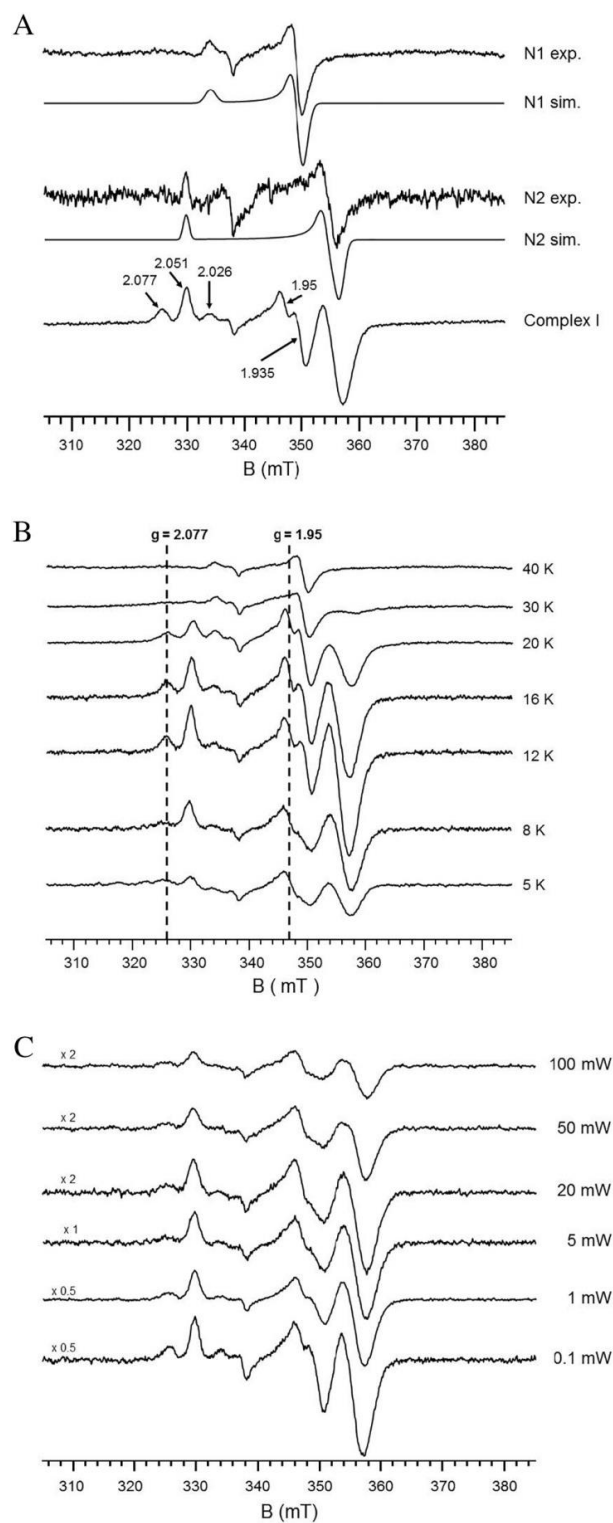
Figure 3.22. **EPR spectra of native *A. aeolicus* complex I.**

(A) The spectrum of a 2Fe-2S cluster (N1 exp.) was recorded with a NADH reduced sample at 40 K and an almost pure cluster N2 spectrum (N2 exp.) at 12 K was obtained after dithionite reduction of complex I and

freezing the sample after short reaction time (10 s). (Complex I) EPR spectrum of complex I reduced by NADH at 12 K. Prominent g-values are indicated. (N1 sim) and (N2 sim) represent simulated spectra of N1 and N2, respectively. Experimental parameters: N1 exp. and N2 exp. mircrowave power 5 mW, modulation amplitude 0.64 mT; Complex I, microwave power 1 mW, modulation amplitude 0.64 mT. Simulation parameters: N1 sim. g (z,y,x)=2.026, 1.938, 1.931, L (z,y,x)=1.6, 1.8, 1.8 mT; N2 sim. g (z,y,x)=2.051, 1.909, 1.897, L (z,y,x)=1.0, 1.8, 1.6 mT. (B) Temperature dependence of EPR spectra from *A. aeolicus* complex I reduced by NADH. From the similar temperature dependence of signal intensities at g=2.077 and g=1.95 (marked by dotted lines) these signals were assumed to originate from one iron sulfur cluster that was tentatively assigned to N4. There were no other clear additional signals at lower temperatures than those already identified in the 12 K spectrum (panel A). Experimental parameters: microwave power 5 mW, modulation amplitude 0.64 mT. (C) Power dependence of EPR signals form NADH reduced *A. aeolicus* complex I (exemplary representation of spectra at 8 K). No additional EPR signals were detectable at low temperature and varying microwave powers. The same analysis was done for temperatures: 40, 30, 20, 16.12, and 5 K. (Spectra were normalized in matters of scan number, receiver gain and power. For better visibility spectra were scaled by indicated factors on the left). The figure was provided by the collaborator, Dr. Klaus Zwicker in Universitätsklinikum Frankfurt.
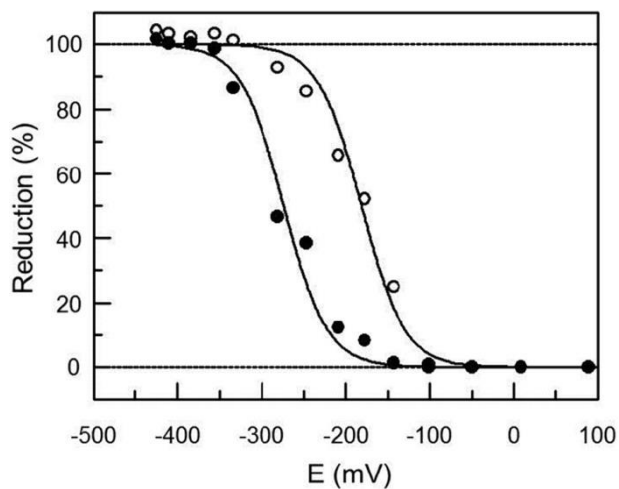
Figure 3.23. **Redox titrations of *A. aeolicus* complex I at pH 7.4 followed by EPR spectroscopy.** (Filled circles) Fitting according the Nernst equation of the potential dependent intensity of the EPR signal of the binuclear cluster at 40 K resulted in a midpoint potential $E_{m,7.4}$=−273 mV and (open circles) for the potential dependency of the g=2.05 EPR signal $E_{m,7.4}$=−184 mV. Intensity values were normalized to the calculated maximum intensity of the reduced form. The figure was provided by the collaborator, Dr. Klaus Zwicker in Universitätsklinikum Frankfurt.

3.2.2 Characterization of Fe-S clusters in single subunits by EPR measurement

EPR samples of NuoB, I_1, I_2, and G were prepared under strict anaerobic conditions. The sample was incubated with 5 mM Na-dithionite for 30 s in the EPR tube under argon atmosphere. The Fe-S clusters of the preparations were characterized by X-band EPR spectroscopy under miscellaneous conditions, e.g. reduction state, EPR parameters, and temperatures.

There was no EPR signal of Fe-S clusters detected in the sample of NuoB, $NuoI_1$ and $NuoI_2$, neither in an oxidized nor in reduced state. While in the sample of NuoG, a signal for 3Fe-4S clusters appeared in oxidized NuoG which normally does not occur in complex I (Figure 3.24A) and there was an EPR signal in the reduced state that might originate from a 2Fe-2S cluster which is similar to N1b in the holoenzyme, but quite weak, broader and more inhomogeneous than that of N1b (Figure 3.24B).
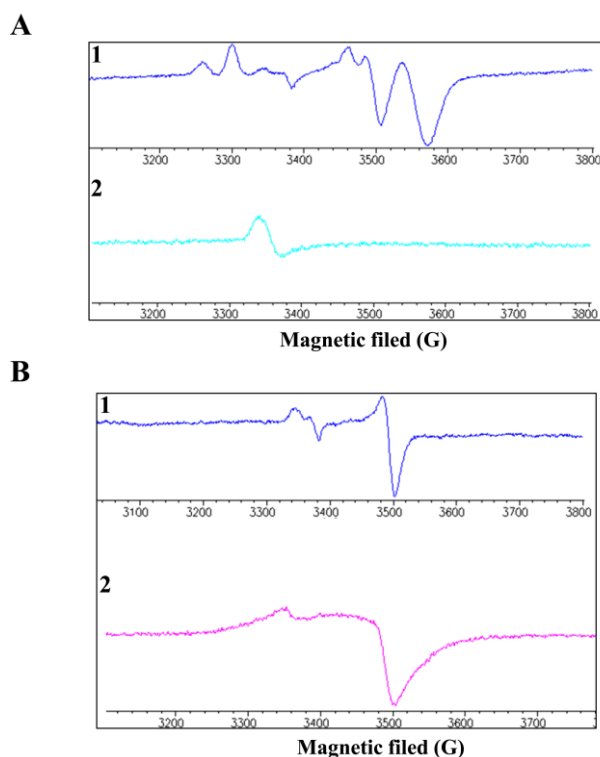

Figure 3.24. EPR spectra of NuoG.

(A) EPR signal characteristic for 3Fe-4S cluster observed in the air-oxidized NuoG at 12 K and 1 mW. (B) EPR signals which might originate from 2Fe-2S cluster observed in reduced NuoG at 40 K and 5 mW. The figure was provided by the collaborator, Dr. Klaus Zwicker in Universitätsklinikum Frankfurt.

It was assumed that the iron sulfur clusters were lost or damaged during isolation and purification of the subunits. In case of NuoB, $NuoI_1$ and $NuoI_2$, this seems to happen completely and in case of NuoG almost completely except a small portion that seems distorted, possibly a damaged remainder of a 4Fe-4S or distorted 2Fe-2S cluster.

3.2.3 Identification of the Fe-S clusters in heterologously produced NQOR1

EPR measurement of NQOR1 was performed as described (see 2.2.3.2) with slight changes. 10 μl 20 mM NADH was added to 90 μl 5 mg/ml heterologously produced NQOR1 in a buffer containing 20 mM Tris-HCl, 150 mM NaCl, 2 mM DTT, 1 mM EDTA and 0.05% DDM. After incubation at RT for 30 s, the sample was transferred to EPR tube and frozen as described. The EPR spectra were record on a Bruker Elexsys E500 spectrometer at X-band using an Oxford Instrument ESR900 helium flow cryostat.
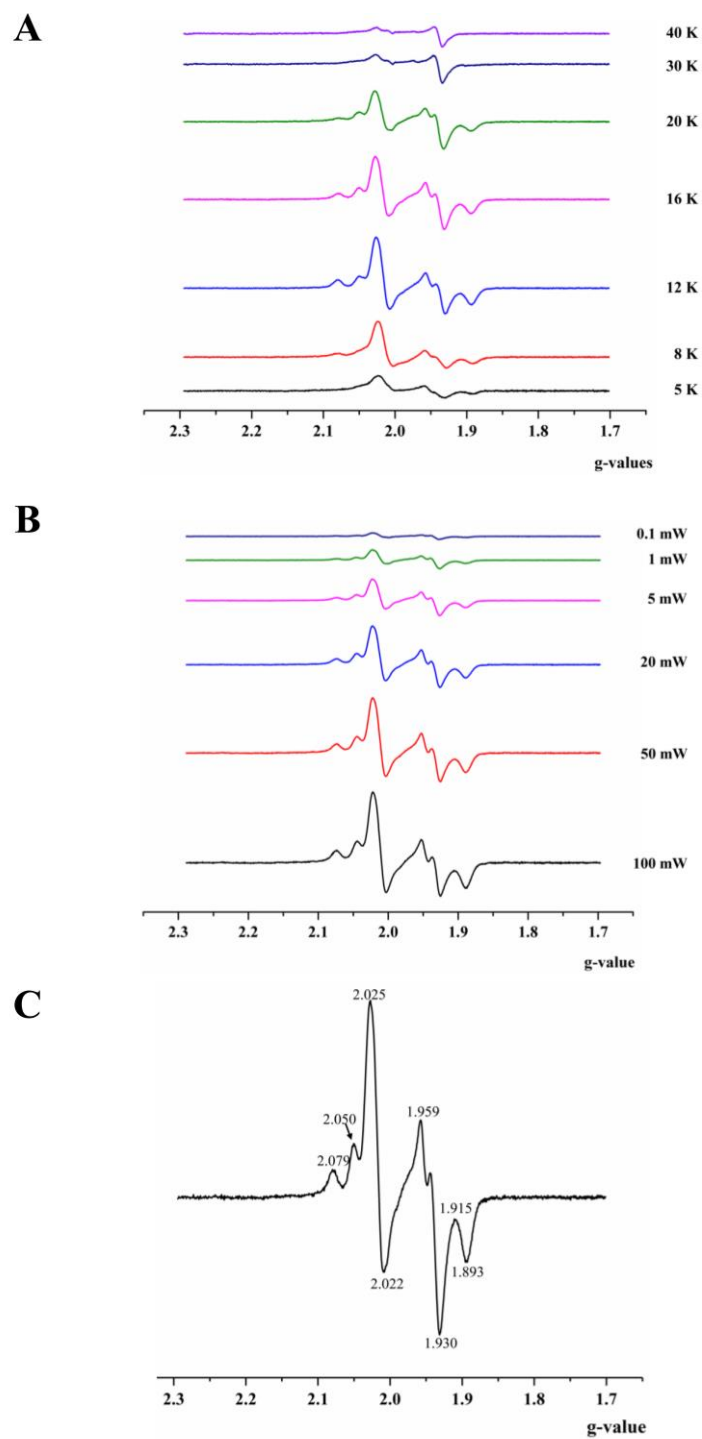
Figure 3.25. **EPR spectra of heterologously produced NQOR1.**

(A) Temperature dependence of EPR spectra. (B) Power dependence EPR spectra. (C) EPR spectra under 16 K, 5.0 mW.

3.3 Investigation the relationships between complex I and AhpC2

3.3.1 Heterologous production and characterization of AhpC2

The protein AhpC2 from *A. aeolicus* consists of 222 amino acid residues. It is assigned to the 1-Cys Prxs class when searching in the PeroxiRedoxin classification indEX (PREX) database [126]. AhpC2 has three Cys residues, at positions of 49, 212, and 218. Structure-based sequence alignment of AhpC2 against other peroxiredoxins representing the six different subfamilies (Figure 3.26) indicates that Cys^{49} is absolutely conserved in all Prxs and serves as the peroxidatic cysteine (C_P). Interestingly, Cys^{212} and Cys^{218} are found in a "CXDWXXC" motif (Figure 3.26), which is also present in some archaeal 1-Cys Prxs. In addition, residues Pro^{42}, Thr^{46}, Ser^{74}, and Arg^{126} are also highly conserved in the Prxs family.
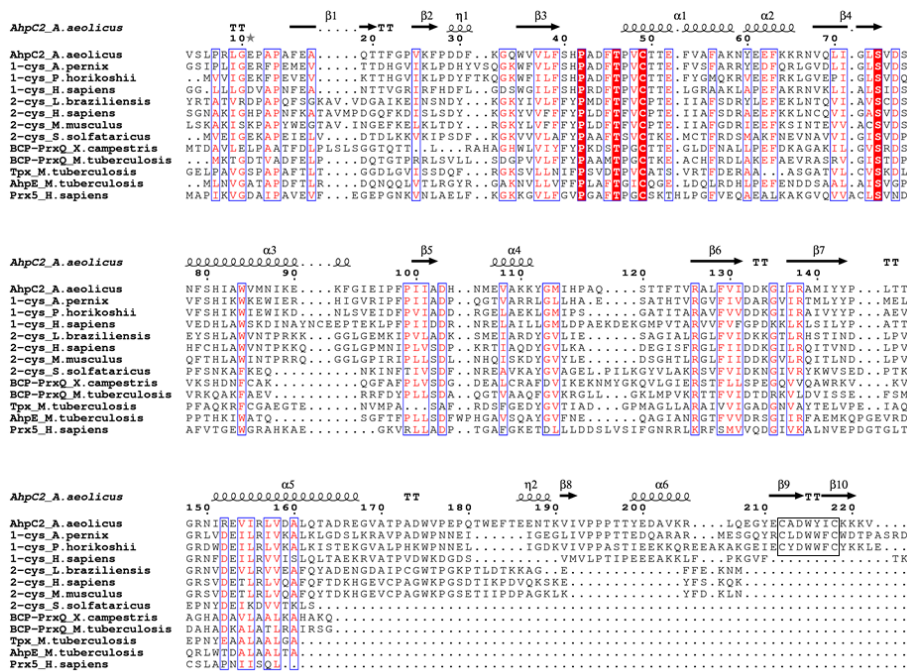
Figure 3.26. **Amino acid sequence alignment of AhpC2 from *A. aeolicus* VF5 against other Prxs, representing major six subfamilies of the Prx superfamily.**

All sequences were downloaded from the NCBI database (www.ncbi.nlm.nih.gov). The sequences of 1-Cys Prxs are from *Aeropyrum pernix* (WP_010866908.1), *Pyrococcus horikoshii* (WP_010885304.1), *Homo sapiens* (NP_004896.1), 2-Cys Prxs are from *Leishmania braziliensis* (XP_001562236.1), *Homo sapiens* (NP_001189360.1), *Mus musculus* (AAH03349.1), *Sulfolobus solfataricus* (WP_009989334.1), BCP_PrxQ

prxs are from *Xanthomonas campestris* pv. campestris str. (NP_637105.1), *Mycobacterium tuberculosis* H37Rv (NP_216124.1), Tpx is from *Mycobacterium tuberculosis* H37Rv (NP_216448.1), AhpE is from *Mycobacterium tuberculosis* H37Rv (NP_216754.1), Prx5 is from *Homo sapiens* (NP_036226.1). The multi-sequence alignment was performed by programs Clustal Omega [127] and ESPript 3.0 [128]. The secondary structural elements of AhpC2 are displayed above the sequences. The totally conserved residues are highlighted in red. The motif of "CXDWXXC" is indicted by a black box.

AhpC2 was expressed in *E. coli* and purified to homogeneity using a combination of Ni-NTA purification and size exclusion chromatography. The purified protein appeared as a single band with a molecular mass around 27 kDa on SDS-PAGE under reducing conditions (Figure 3.27).
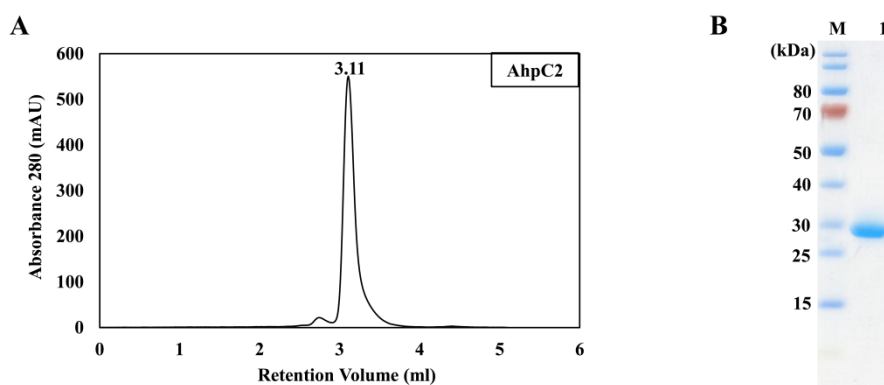
Figure 3.27. **Purification profiles of AhpC2.**
(A) SEC profiles using a Yarra 3u SEC-4000 column (B) SDS PAGE analysis under reducing conditions.

3.3.2 Redox potential of AhpC2

Redox titration was performed by a modification of the monobromobimane (mBBr) fluorescence method [129]. Fluorescence was measured using Multimode Reader Trista LB941 with excitation at 380 nm and emissiom at 460 nm. The titration give excellent fits to the Nernst equation for a single two-electron redox couple. Multiple titrations (3 titrations) of AhpC2 gave average values for $E_{m, 7.0}$ of around -310 mV (Figure 3.28).
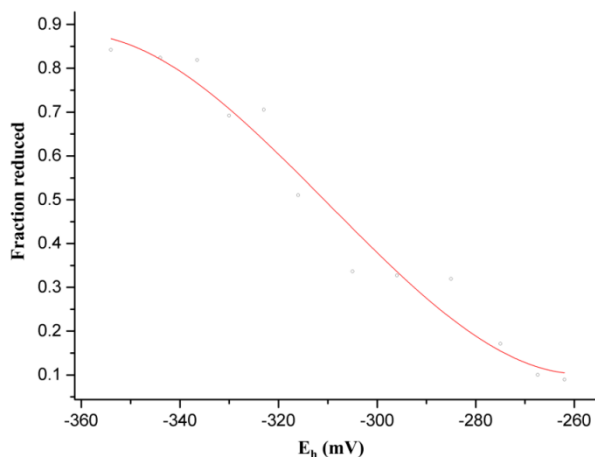
Figure 3.28. **Oxidation-reduction titrations of AhpC2 at pH 7.0.**

AhpC2, at a final concentration of 7 μM, was incubated in 500 μl of 50 mM Tris-HCl buffer (pH 7.0) containing DTT at a total concentration of 2 mM. After incubation at ambient temperature for 2 hour, excess mBBr was added and the samples were prepared for fluorescence analysis as described under Material and Method. Each point represents the average fluorescence from two replicate 50 μl samples, each of which was diluted with water into microtiter plates to a total volume of 200 μl.

3.3.3 The interaction of complex I with AhpC2

The binding affinities of the complex I to AhpC2 was evaluated using SPR. The purified *A. aeolicus* complex I from the native source was immobilized to the CM5 chip. AhpC2 was either kept as it is after purification or treated with 1 M NaCl overnight. Different concentrations of AhpC2 were flowed over the complex I immobilized CM5 surface. AhpC2 was in spots after treatment by NaCl, distinct from the ring structure of untreated AhpC2 observed by EM (Figure 3.29A, B). The binding kinetics of AhpC2 to complex I was shown in Figure 3.24 C, D. The binding constant (KD) was 0.478 nM and 4.84 nM, respectively, for the NaCl treated and untreated AhpC2.
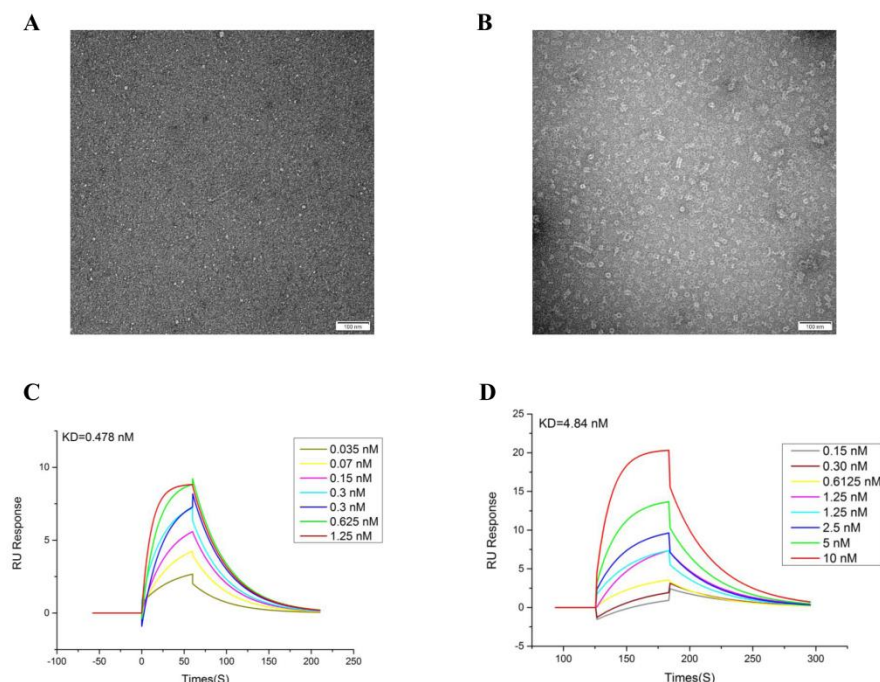
Figure 3.29. SPR analysis of the binding of AhpC2 with complex I.
(A) Negative staining of the sample treated with 1M NaCl overnight. AhpC2 was in spots. (B) Negative staining of raw sample without any treatment. AhpC2 was in bouquet. Sensograms for the AhpC2 binding to complex I immobilized on the CM5 chip surface, the concentration of AhpC2 is indicated. (C) AhpC2 treated with 1 M NaCl overnight. (D) AhpC2 without any treatment. The figure was made by Hailong Gao, the collaborator from the group of Prof. Zihe Rao in the Institute of Biophysics, Chinese Academy of Science.

3.3.4 Structural analysis of AhpC2

3.3.4.1 Electron microscopic single-particle analysis

Structural investigation of WT AhpC2 was performed by electron microscopic single-particle analysis. Electron micrographs of the negatively stained proteins showed a uniform distribution consisting of a ring-shaped top view and a dumbbell-shaped side view (Figure 3.30A). The 2D class averages of the ring-shaped view revealed a structure with 6-fold symmetry, while the dumbbell-shaped view revealed a 2-fold symmetric structure (Figure 3.30B, 1-4). The hexagonal ring has an inner diameter of about 6.6 nm and an outer diameter of 15.6 nm (Figure 3.30B, i). The height of the dumbbell was approximately 8.9 nm (Figure 3.30B, ii). It was interesting to observe that some of

AhpC2 ring stacked up to form a tube-like structure (Figure 3.30B, 5-7). These results suggest that AhpC2 is organized as a dodecamer or an even higher order assembly in solution.
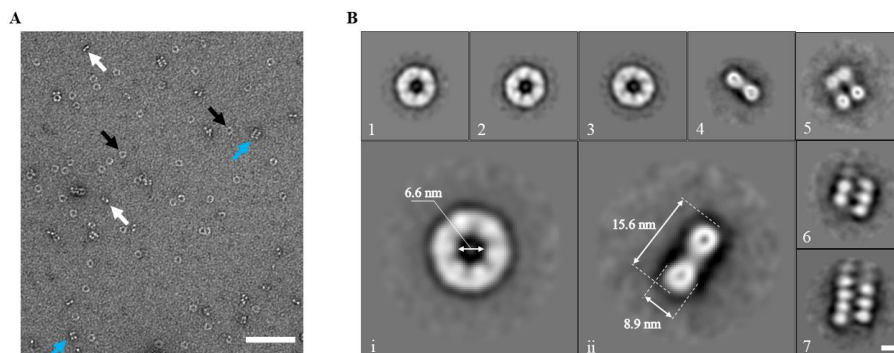

Figure 3.30. **Electron micrographs of AhpC2.**

(A) Images obtained by negative staining. The top view and side view are marked with black and white arrows, respectively. Regular stacking of AhpC2 is indicated by the blue arrow. The scale bar represents 100 nm. (B) Representative 2D class average views in different orientation. 1, 2, 3 and 4 are 2D averages derived from the translationally and rotationally aligned AhpC2 particles. Each class was averaged by about 100 particles. 1, 2 and 3 show class of top views and 4 shows class of side view. The scale bar represents 10 nm. 5, 6 and 7 show 2D images of the stacked AhpC2 particles in different forms. The images are averaged by about 50 particles respectively. The scale bar represents 10 nm. i is the 2D average of the top-view particles with inner diameter indicated. ii is the 2D average of the side-view particles with outer diameter and height indicated. The figure was made by Hailong Gao, the collaborator from the group of Prof. Zihe Rao in the Institute of Biophysics, Chinese Academy of Science.

3.3.4.2 Crystal structure of AhpC2

The crystal structure of AhpC2 as isolated was determined by molecular replacement and refined to a final R_{cryst} / R_{free} of 19% / 22% at 1.8 Å resolution. Statistics of data collection and refinement are summarized in Appendix Table S1. The final atomic model includes twelve monomers and 1091 water molecules in the asymmetric unit. The quality of the electron density allowed modeling of amino acid residues 3-222 (of 222) in two monomers and amino acid residues 5-222 in other ten monomers. The twelve monomers of AhpC2 are arranged as a toroid-shaped hexamer of homodimers or an $(\alpha_2)_6$ dodecamer with an outer diameter of ~140 Å and an inner diameter of ~ 70 Å (Appendix Figure S1).

This structure is roughly consistent with the structure of AhpC2 in solution observed by EM.

The monomer structure of AhpC2 (Figure 3.31A) is extended and can be divided into two discrete domains, the N-terminal domain (residues 1-168) and the C-terminal domain (residues 185-222) (Figure 3.31B). The two domains are connected by a long loop (residues 169-184) consisting of 16 amino acid residues.

The peroxidatic Cys^{49} (C_P^{49}) is present in the first turn of the N-terminal α2 helix which kinks at Asn^{59} and forms a β-α-β motif with the adjacent parallel β-strands (β3 and β4). During the refinement process, it was noted that C_P^{49}-SH was hyperoxidized to the sulfonic acid form (C_P^{49}-SO$_3$H), denoted as CYO (Figure 3.32A). The Cys^{212} and Cys^{218} residues are positioned on β8 and β9-strand, respectively, which construct a β-hairpin motif. Cys^{212} forms an intramolecular disulfide bond with Cys^{218} (bond length 2.1 Å. Figure 3.32B).
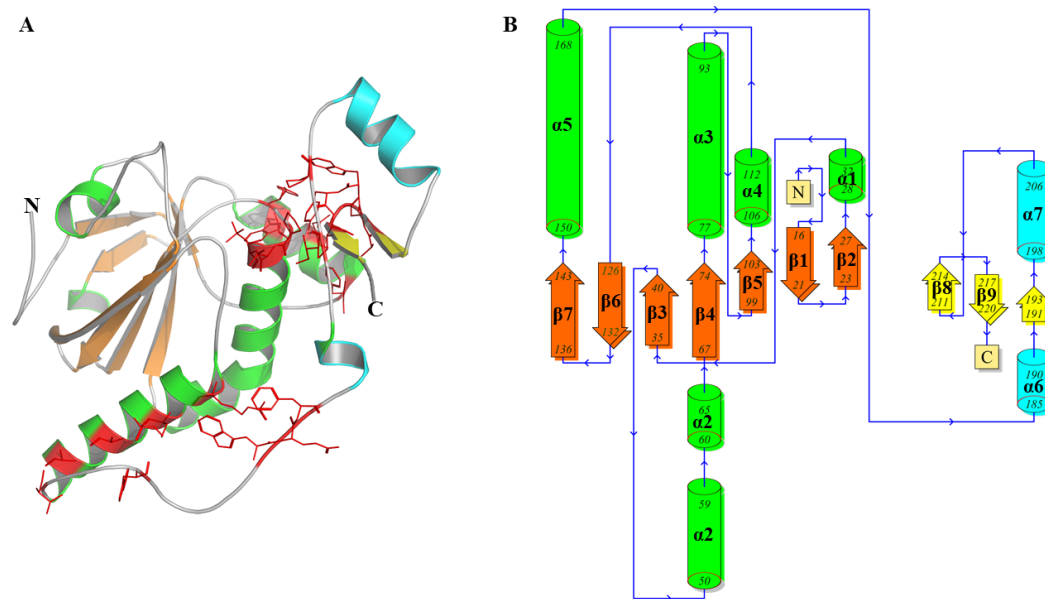
Figure 3.31. **Monomer structure and topology diagram of AhpC2.**

The α-helices and the β-strands from the N-terminal domain are depicted in green and orange, and those from the C-terminal domain in cyan and yellow, respectively. (A) Monomer structure. The residues involved in the interactions of N- and C-terminal domains are colored red and shown by sticks. (B) Topology diagram. The secondary structure was defined by using the program PDBSum [130] and polished by Inkscape (https://inkscape.org/en/). The beginnings and the ends of the secondary structural elements are labeled manually with slight modifications based on the monomer structure. Residues 50-65 belong to helix

α2 not the defined two α-helices, while residues 191-193 reside on the loop not the β-sheet. All structural figures were prepared using the program PyMOL (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC), unless specified otherwise.
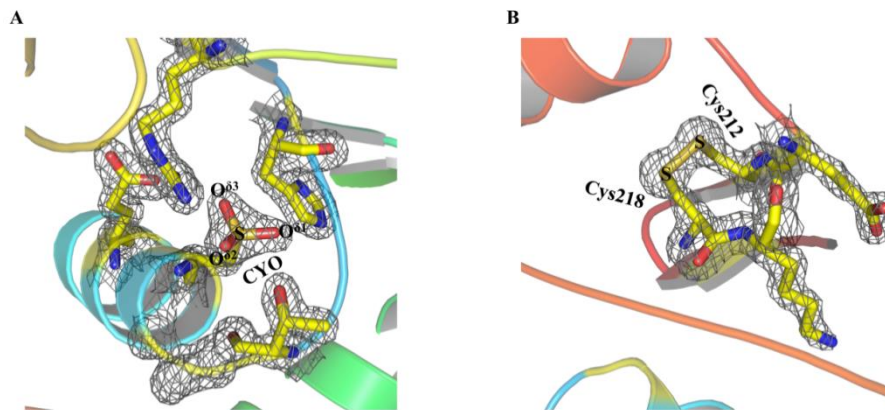
Figure 3.32. Electron-density map around the redox-active Cys residues of AhpC2.
In the stick model, the C, N, O, and S atoms are colored yellow, blue, red, and orange, respectively. (A) Electron density around the peroxidatic Cys, C_P^{49}. The C_P^{49} residue is hyperoxidized to the sulfonic form and is denoted as CYO. (B) Electron density around Cys^{212} and Cys^{218}, which form an intramonomeric disulfide linkage.

3D structure alignment performed by PDB search using the Dali Server [131] indicated that the basic folding of AhpC2 resembles that of the 1-Cys Prxs and shares most similarities with the 1-Cys Prxs from the archaea *Pyrococcus horikoshii* (PhPrx, PDB ID: 3W6G) [117] and *Aeropyrum pernix* K1(ApTPx, PDB ID: 1X0R) [132]. A detailed analysis was carried out using the PyMOL Molecular Graphics System (Version 2.0 Schrödinger, LLC) by superimposing the monomer structure of AhpC2 and other selected Prxs including two archaeal Prxs PhPrx, ApTPx, and human 1-Cys Prx hORF6. A significant difference was observed at the C-terminus, the orientation of AhpC2 C-terminal domain is opposite to that in other Prxs, a feature unique in *A. aeolicus* AhpC2 (Figure 3.33A). Despite the opposite orientations, a comparison of the C-termini of these Prxs was performed (Figure 3.33B). It was found that the basic topology of the C-terminal domains is similar, whereas the connecting element (including a loop and α6

helix) of the N-terminal α5 helix and the C-terminal α7 helix of AhpC2 is longer compared to other Prxs.
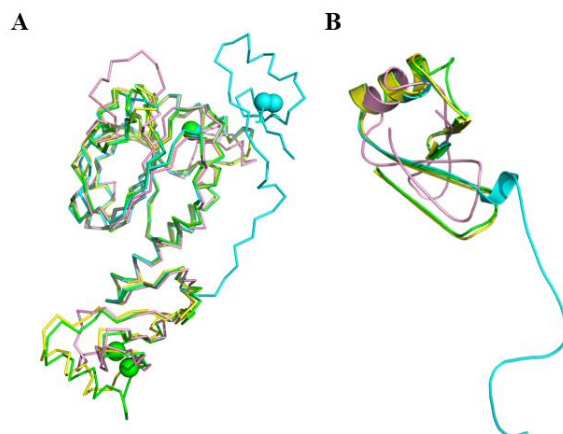
Figure 3.33. **Structure comparison of AhpC2 with selected 1-Cys Prxs.**
(A) Overlay of monomers. AhpC2 (in cyan), PhPrx (PDB: 3W6G, in green), ApTPx (PDB: 1X0R, in yellow), and hORF6 (PDB: 1PRX, in pink). (B) Superposition of the C-terminal domains of AhpC2, PhPrx, ApTPx and hORF6. For (A) and (B), the last α-helix (α7) of ApTPx has been omitted for clarity.

3.3.4.3 Co-crystallization of *A. aeolicus* complex I and AhpC2

A. aeolicus complex I and AhpC2 were co-crystallized by a molar ratio of 1:2 and 1:1. The crystals was obtained under the condition of 18 ℃, with 30% (w/v) of 2-methyl-2,4-pentanediol (MPD) as the precipitant in a buffer of 0.1 M Tri-sodium citrate, pH 5.6 after one month. The crystal could diffract but the resolution was low (8-10 Å) (Figure 3.34).
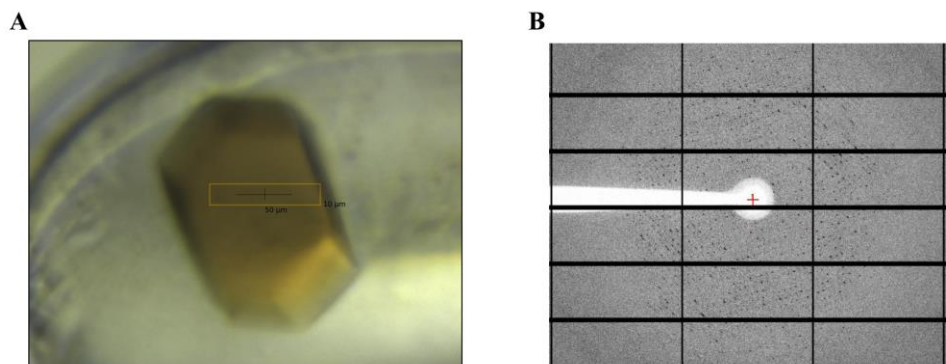
Figure 3.34. **Co-crystallization of AhpC2 with complex I.**
(A) Crystals. (B) X-ray diffraction pattern.

3.4 Outlook

In this work, we have mainly four achievements: (1) the construction of a heterologous expression system to produce *A. aeolicus* complex I; (2) the heterologous production, isolation and purification of NQOR1 from *E. coli*; (3) the identification of Fe-S clusters in the complex I both from native source and expression host; and (4) the investigation of the relationship between AhpC2 and complex I.

A fully assembled and active NQOR1 was obtained from a heterologous expression system for the first time, which constitutes a solid reference for the production of NQOR2 and allows us to perform mutagenesis and cross-linking experiments for further structural and functional analysis. The purified NQOR1 was proved to be a mixture of the intact complex I and subcomplexes lacking one or a few of membrane subunits, which provides implications of a possible assembly pathway of *A. aeolicus* complex I. However, intact NQOR1 only has a small proportion in the purified sample, which limited the quantitative analysis. In the next step, optimization will be made to eliminate the subcomplex and to obtain a homogeneous NQOR1. The EPR spectrum of the heterologously produced NQOR1 is similar to that of the native *A. aeolicus* complex I, which reveals a highly overlapped EPR signal originating from several paramagnetic centers. Mutations in the binding motif of the Fe-S cluster may be helpful to assign a specific cluster and will be performed in the future studies. In addition, although our work has suggested that AhpC2 had a very strong affinity to the complex I. However, the role of AhpC2 played in this process is still unclear. Further investigation will be carried out to understand how it works.

4. Discussion

4.1 Two isoforms of *A. aeolicus* complex I

A. aeolicus is a microaerophilic, hyperthermophilic bacterium [91], which has a complete respiratory chain and is capable of using oxygen as the electron acceptor. In comparison with complex I from other organisms, complex I from *A. aeolicus* exhibits some interesting features. One of them is that the duplicate and triplicate *nuo* genes are present in the genome sequence of *A. aeolicus* [91]. To gain more information about it, the distribution of *nuo* genes from various organisms was investigated by searching for complex I encoding proteins throughout the whole genome sequences in the DOE Joint Genome Institute databases [133]. We used the COGs of each of the complex I subunits to screen the genomic data. These are NuoA (COG:0838), NuoB (COG:0377), NuoC (COG:0852), NuoD (COG:0649), NuoE (COG:1905), NuoF (COG:1894), NuoG (COG:1034), NuoH (COG:1005), NuoI (COG:1143), NuoJ (COG:0839) NuoK (COG:0713), NuoL (COG:1009), NuoM (COG:1008), and NuoN (COG:1007). It was interesting to discover that the presence of duplicate and triplicate *nuo* genes is ubiquitous in the phylum *Aquificae*, including the *A. aeolicus*, *Hydrogenobacter thermophilus, Hydrogenobaculum* sp. Y04AAS1, *Thermocrinis albus* DSM 14484, and *Thermovibrio ammonificans* HB-1. *A. aeolicus, H. thermophilus, H.* sp. Y04AAS1, and *T. -albus* DSM 14484 are microaerophiles which belong to the family *Aquificaceae,* while *T. -ammonificans* HB-1 is a strict anaerobic bacterium, within the genus *Thermovibrio*, a member of the family *Desulfurobacteriaceae* (Figure 4.1).

To date, it is not known whether the *nuo* genes are all expressed. No complex I isoforms have been identified in any bacteria as well as the function of the remaining duplicate and triplicate *nuo* genes. Until recently, our study of *A. aeolicus* complex I revealed 20 partially homologous subunits using a combination of MALDI-TOF and LILBID mass spectrometry methods, which allowed a distinction between two different complex I isoforms, named NQOR1 and NQOR2 [123].
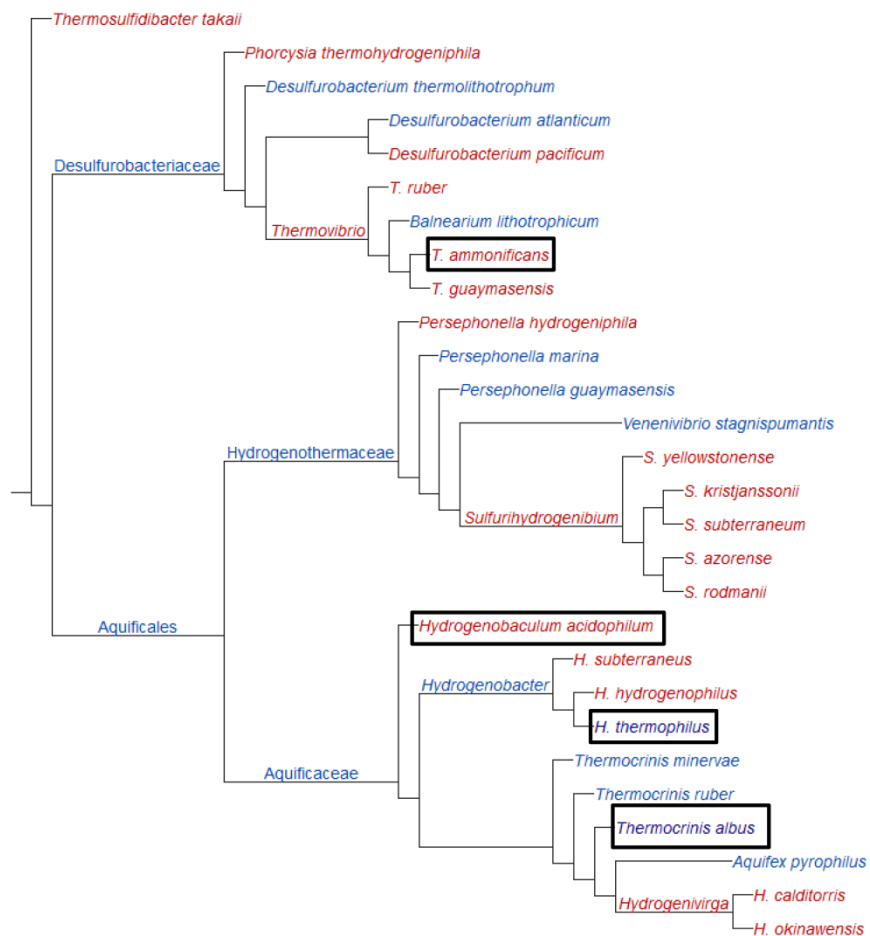
Figure 4.1. **The phylogenetic position of *Aquificae*.**
The phylogeny based on 16S rRNA-based LTP release 123 by "The All Species Living Tree Project" [134]. The figure was adapted from https://en.wikipedia.org/wiki/Aquificae.

Additionally, one isoform was also discovered in mitochondrial complex I. It was reported that subunit NDUFV3 is present in the canonical 10 kDa and in addition, in a novel 50 kDa isoform, generated through alternative splicing [135, 136]. Both isoforms assemble into complex I and their levels vary in different tissues. Either isoform alone is sufficient for assembly of mature complex I.

Isoforms generated by gene duplication and alternative splicing are two fundamentally different mechanisms. In the former, multiple copies of a gene are generated in the genome. The generation of isoforms appears to have played a fundamental role in the evolution of biological diversity [137]. As a general rule, isoforms do not precisely co-

localize in time and space in the life of an organism. This observation suggests on one hand that the generation of isoforms has helped to evolve multicellularity, developmental processes and the specialization of intracellular compartments. It is therefore not surprising that eukaryotic cells contain many protein isoforms, including the respiratory chain complexes. There are multiple homologous subunits of cytochrome c oxidase [138, 139], V-type ATP synthase with two different homologous peripheral stalk subunits [140] and most recently there are reports that complex I in murine and bovine tissues possesses different homologous subunits, and also in rat tissues and hematoma cell lines [141]. The common feature in eukaryotes is that expression of the isoforms is tissue or cell type-specific.

Protein isoforms themselves usually differ in their amino acid sequence and are therefore expected to differ in their functional properties, which can have an impact on a range of functional properties that include intra- and extracellular location, regulatory properties and catalytic and/or structural function. One of the goal of our work is to separate the two isoforms of *A. aeolicus* complex I to investigate the variations in its structure and function.

4.2 Isolation of the individual isoform of complex I by immunoprecipitation

Separation of the two isoforms in native preparations is a great challenge due to their high homology. They are co-eluted as a mixture in the same chromatographic fractions and appeared as one band on IEF and BN-PAGE gel. One approach was, therefore, to perform a small scale purification of the individual isoform by immunoprecipitation. It is known that complex I is a multi-subunit, enormous membrane protein. Additionally, the subunits composed two isoforms share a high sequence similarity. Therefore, a lot of factors need to be taken into consideration to obtain the optimal result.

- **Selection of the antibody:** Polyclonal antibodies bind to multiple epitopes on the target protein and thus form tighter antibody-protein complexes in comparison with monoclonal antibodies. This high retention rates between the polyclonal antibodies with the target protein also implicates that the target protein is less likely to be washed away during the washing steps. Thus, polyclonal antibodies were used as

capture antibodies here. Additionally, subunit NuoD and NuoI, located in the hydrophilic arm of complex I, are better accessible for antibodies than subunits of the membrane domain. Therefore, we selected the polyclonal antibody of $NuoD_2/NuoI_1$ and $NuoD_1/NuoI_2$ to target NQOR1 and NQOR2, respectively.

- **Binding efficiency and specificity:** To improve the binding affinity of Ab/Ag, affinity-purified antibody was used instead of serum, in which the antibody that is specific for the antigen of interest may account for only 2–5% of the total IgG and results in low antigen yields. In addition, S3 sample containing two isoforms was chosen as antigen for purification to eliminate the non-specific binding of contaminants to the antibody. Furthermore, different molar ratios of Ab/Ag were tested to ensure sufficient antibodies needed to capture of the target protein.

Optimal result was achieved using Anti-NuoI$_2$ with molar ratio of Ab/Ag at 10:1 (see chapter 3.2). However, the binding of Ab/Ag was still too weak and non-specific binding could be detected. It is promising to capture individual isoform of complex I by immunoprecipitation, but more optimization is required. It is a significant challenge to separate the two isoforms directly from the native source. This approach was abandoned because the second strategy was more successful at a certain stage.

4.3 Heterologous production of the individual isoform in *E. coli*

4.3.1 The *nuo* gene encoding complex I

To date, little is known about the assembly pathway of bacterial complex I. There are some reports on the regulation of the expression of *nuo* genes and the assembly of complex I in *E. coli* cells. The *nuo* genes in *E. coli* are organized in one operon ordered as *nuo*A-N [11], which is conserved in several other bacteria, including *S. typhimurium* [12]*, P. denitrificans* [13], *R. capsulatus* [14], and *T. thermophilus* [15]. The 5' half of the locus contains a previously identified promoter (*nuo*P) [142, 143], which is located upstream of *nuo*A. The 3' end of nuoG encodes a C-Terminal region (CTR) of the NuoG subunit. Defects caused by deletion or duplication in this region prevent a correct functionality of complex I [144]. The assembly of *A. aeolicus* complex I in native cell is,

in particular, mysterious, which must be rather complex in respect to other organisms because the *nuo* genes are dispersed in different loci throughout the genomic DNA.

4.3.2 Heterologous production of protein complexes from hyperthermophilic organisms in mesophilic hosts

Proteins from hyperthermophilic organisms are much more stable than their counterparts from mesophiles or moderate thermophiles. Therefore, structural comparison of hyperthermophilic and mesophilic enzymes is expected to provide us with a better understanding of the mechanisms by which highly thermostable proteins are stabilized. However, studies of hyperthermophiles are often hampered by the lack of genetic manipulation system and obstacles in cell cultivation. Therefore, construction of a heterologous expression system is particularly important for expressing genes from hyperthermophiles. Although there are some difficulties by producing hyperthermophilic proteins in the mesophilic hosts, the genes encoding hyperthermophilic proteins and enzymes have been extensively expressed in heterologous organisms such as *E. coli* and their good productions have been achieved in the last decades [122, 145, 146], including membrane proteins [147, 148] and large multimeric complexes [149-151]. In this work, *E. -coli* was chosen as the host organism. It is a well-established host for a long time that accounts for short culturing time, easy genetic manipulation and low cost media. Furthermore, *E. coli* possesses complete ISC and SUF system for Fe-S cluster assembly (Figure 4.2).

Organism	Variant Code	Sulfur transfer: Cysteine desulfurase (CDS)	Iron-binding: IscA-2, SufA-3, SufA2-4, HesB-5	Scaffold: IscU-6, NifU-7	SufE	SufD	SufB	SufC	Chaperones: HscB-12, HscA-13	Fdx
Escherichia coli K12	1	1664, 2500, 2766	2498-2, 1668-3, 157-5	2499-6	1663, 2767	1665	1667	1666	2497-12, 2496-13	2495
Chlamydophila pneumoniae AR39	10	354, 56			693	55	53	54		
Pyrococcus furiosus	100	1103, 167				1329	1330	1331		
Methanocaldococcus jannaschii	1000						34	35		
Thermoplasma volcanium	2000		1393-4			1389	1390	1391		
Mycobacterium tuberculosis CDC1551	11	1552, 3216	1554-4, 2323-5	1553-6	3503	1550	1549	1551		
Erwinia carotovora	111	2606, 3010, 3844, 3965	2604-2, 3969-3, 4492-5	2605-6, 3845-7	3011, 3964	3966	3968	3967	2603-12, 2602-13	2601
Prochlorococcus marinus str. MIT 9313	2	1597, 2037, 716	1961-5		1140	1598	1600	1599		
Burkholderia pseudomallei K96243	3	4159, 4250, 6551, 782	4252-2, 4158-4, 3382-5	4251-6, 6621-6		4160	4162	4161	4253-12, 4254-13	4255
Pirellula sp.	33	1940, 4145, 5102, 5876	2543-5	1939-6		5520	5521	5522		
Haemophilus influenzae Rd KW20	4	1234, 1278, 352	350-2, 1635-5	351-6	1233				349-12, 347-13	346
Azotobacter vinelandii	40	1435, 2228, 3056, 4190, 548	1433-2, 2230-2, 313-5	1434-6, 2229-7	1320				1432-12, 1431-13	1430
Anopheles gambiae [E]	44	11207, 13036	12371-5, 4948-5	1256-6	13037					2254
Neisseria meningitidis Z2491	5	1453	1456-2, 681-5	1455-6					1457-12, 1230-13	1233
Aquifex aeolicus VF5	55	523, 730	1292-5	622-6						
Leuconostoc mesenteroides	6	1009		1010-6		1008	1011	1007		
Chlorobium tepidum TLS	66	1965		1964-6						
Campylobacter jejuni RM1221	7	284		283-7						
Geobacter sulfurreducens PCA	77	1855, 1998, 2554, 2768		1399-6, 1999-7						
Bacillus subtilis	8	267, 2754, 2791, 2962, 3274	3222-5	3273-6		3275	3272	3276		
Nostoc sp. PCC 7120	9	1766, 2802, 2812, 3395, 3517, 4174	1741-5, 2692-5, 4648-5	1765-7	3820	2801	2799	2800		

Functional variants:
#1: complete ISC and SUF systems;
#111: complete ISC, SUF, NIF systems;
#10: SUF: SufBCD+SufE+SufS;
#100: SUF: SufBCD+SufS;
#1000: SUF: only SufBC;
#2000: SUF: only SufABC;

#2: ISC: only IscA+IscS, SUF - complete;
#3: ISC: complete, SUF lacks SufE;
#33: ISC: IscA+IscU+IscS, SUF lacks SufE;
#4: ISC: complete, SUF: only SufE;
#40: ISC: complete, SUF: only SufE, NIF
#44: ISC: IscA+IscU+IscS, SUF: only SufE;
#8: ISC: IscA+IscU+IscS, SUF: SufBCD;

#5: ISC - complete;
#55: ISC: IscA+IscU+IscS;
#6: ISC: only IscS+IscU, SUF: SufBCD;
#66: ISC: only IscS+IscU;
#7: NIF: NifU+NifS;
#77: NIF: NifU+NifS; ISC: IscS+IscU
#9: NIF: IscA+NifU+NifS, SUF: SufBCD+SufE.

Figure 4.2. System for Fe-S cluster assembly.
The figure was adapted from http://www.theseed.org/SubsystemStories/Fe-S_cluster_assembly/story.pdf.

Homologous overproduction of recombinant single subunits, sub-complexes and the entire complex I were carried out in *E. coli* for structural and functional analysis [24, 54, 55, 57, 58, 95, 152-155]. It was reported that the fully assembled NADH dehydrogenase fragment consisting of NuoE, F, and G could be obtained when the genes of *nuo*E, F, and G from *E. coli* were simultaneously overexpressed with the genes *nuo*B, C, and D [24]. Furthermore, it was proved that engineering the Strep-tag II sequence to the C-termini of NuoE, F or G led to the overproduction of NuoB, CD, E, F and G and the assembly of NADH dehydrogenase fragment in the cytoplasm, however, fusion of strep-tag to the N-terminus of either NuoE or NuoF disturbed the assembly of the NADH dehydrogenase fragment [152]. The entire *E. coli* complex I was homologously overproduced in an *E.coli nuo* deletion strain [57, 95, 154].

To date, there is no successful example of heterologous production of intact complex I reported, although the production of *A. aeolicus* complex I was previously performed by *Marta Macedo Vranas* in Thorsten Friedrich's lab [122, 156]. In their study, several expression plasmids were constructed, with His6-tag inserted into the C-terminus/N-

terminus of NuoF, His10-tag into the C-terminus of NuoM, and the N-terminus of NuoB, respectively. Cell growth and protein production were optimized and purification strategies were manifold varied. Although a membrane-bound subunit was detected in one preparation, and hydrophilic subunits NuoB, NuoE, NuoF, NuoG and $NuoI_1$ in the preparation could be identified by Mass Spectrometry when the purification was performed at 25-30 ℃, only the soluble NuoEF subcomplex was purified to homogeneity. There was no indication of a fully assembled complex I.

Based on information from the previous work [24, 122, 152, 156], we succeeded for the first time to produce fully-assembled complex I from a hyperthermophilic organism in *E. -coli*.

4.3.3 Constructs designed for recombinant NQOR1

As described in chapter 1.4, the genes encoding *A. aeolicus* complex I NQOR1 are spread in three loci , with genes of *nuo*E, *nuo*F overlapped and organized in one operon; the gene of nuoG alone in the second operon; and the genes of $nuoA_2$, $nuoB$, $nuoD_2$, $nuoH_1$, $nuoI_2$, $nuoJ_1$, $nuoK_1$, $nuoL_1$, $nuoM_1$ and $nuoN_1$ in the third operon. Given the fact that the total size of *nuo* gene is enormous and reorganization of 13 *nuo* genes into one artificial operon will interrupt and change the intergenic regions and regulatory elements, which may further affect gene transcription and protein translation, etc., we tended to keep the *nuo* gene as the same as it is in *A. aeolicus*. In consequence, we adopted a co-expression strategy using two plasmids pBADCM1 and pBAD33, both possessing P_{BAD} (systematically araBp) promoter, compatible origins of replication and independent antibiotic selection for maintenance (see chapter 3.3.2.1). The gene of *nuo*EF and *nuo*G were rearranged into one artificial operon and inserted into ORF2 of pBADCM1, the rest *nuo* genes in the original operon of *A. aeolicus* genomic DNA was inserted into pBAD33. The intergenic region between nuoF and nuoG was created by 58 bp upstream of *nuo*G in *A. aeolicus* genome. The translation initiation region (TIR) of *nuo*A is 38 bp upstream of *nuo*A in *A. aeolicus* genome.

In order to simplify the isolation, purification and detection of a recombinant protein, affinity tags are usually incorporated during vector construction. His-tag and Strep-tag II are widely used because they are small and therefore, often do not interfere the

downstream applications [157]. Here we select Strep-tag II. The pros and cons of the Strep-tag and His-tag system are listed in Table 4.1.

Table 4.1 Pros and Cons of the Strep-tag® and His-tag system

		Strep-Tag	His-Tag
Pros (+)		• High purity, even under denaturing conditions and in batch purification • Physiological buffer conditions • Mild reversibility without changing the overall buffer conditions • Purification of metallo-proteins • Beneficial for membrane proteins	• Resin can be recharged • Denaturing conditions with Gua-HCl
Cons (-)		• Lower yield compared to His-tag	• Not suitable for metal proteins • Non-specific interaction with complex forming amino acids • Does not tolerate chelating agents • Time consuming elution gradient for higher purity

We fused the StrepII-tag to the C-terminus of NuoG on the basis of two observations. First, the C-terminus is located one uppermost tip of the molecule hydrophilic domain, exposed to solvent in the known 3-D structures [10]. Second the *nuo*G gene is located at the 3' end of artificial operon *nuo*EFG, which will not disrupt the intergenic region. Consequently, a tag in that position should not interfere with structural assembly or with substrate binding. In order to avoid interferences and chimera formation from chromosomally encoded *E. coli nuo* subunits, a chromosomal *nuo* deletion strain named BA14 (Δ *nuo*) [95] was used for heterologous production of *A. aeolicus* complex I.

4.3.4 Production and purification of NQOR1 from *E. coli*

Ferric ammonium citrate ($(NH4)_5[Fe(C_6H_4O_7)_2]$), ferrous sulfate ($FeSO4$), sodium sulfide (Na_2S), L-Cysteine and Riboflavin (Vitamin B2) were supplemented in the LB medium to supply ions necessary for the synthesis of Fe-S clusters. Purification of NQOR1 was

first attempt using Strep-Tactin purification; however, a protein contaminant with a molecular weight approximately 720 kDa cannot be eliminated. Afterwards, the method for purification of the native *A. aeolicus* complex I was tried but other protein contaminants appeared. Finally, we obtained pure complex I with an anion exchange and subsequent Strep-Tactin affinity chromatography (See chapter 3.3.2.2). Protease inhibitor, EDTA and reducing agent (DTT) were added in the buffer used for protein purification to reduce oxidation damage.

4.4 Structural and functional analysis of NQOR1 produced in *E. coli*

4.4.1 Identification of the purified NQOR1

After the two-step purification, the resulting protein sample was evaluated by molecular mass and NADH dehydrogenase activity staining, respectively. Multiple bands appeared in the gel of native page, which all exhibited NADH oxidase activity (see chapter 3.3.3.2). The composition of the purified protein solution and individual band on native page were further identified by MS. All subunits of NQOR1 could be detected, with trace protein contaminants from *E. coli*. It was demonstrated that the purified protein was inhomogeneous, not only containing the fully assembled NQOR1, but also subcomplexes lacking one or a few of membrane subunits (see chapter 3.3.3.3). From the native page gel, it could be estimated that the hydrophilic arm occupied highest proportion in the purified complex I. One possibility would be that the entire membrane arm dissociated in solution or that the membrane subunits degraded despite the presence of protease inhibitors, reducing agents and metal chelators in buffer. The composition of one band on the native page gel with a higher molecular weight and NADH hydrogenase activity (gel band 6, see chapter 3.3.3.3) aroused our interest and raised another possibility. This band indicated a large protein assembly, which contains all components of *A. aeolicus* complex -I hydrophilic arm and a subunit from the membrane arm $NuoH_1$, transcription termination factor, ribosomal protein and polypeptide elongation factor for protein biosynthesis, chaperonin, proteins involving anabolism and catabolism, etc. (see Appendix Table S2). It seems that an active ribosome with protein under synthesis and assembly was captured during protein purification. Based on this observation, the

partially assembled complex I could also be the intermediates generated during complex I biosynthesis. The successful assembly of the intact complex I was further confirmed by EM, which indicates the characteristic L-shape structure. The 2-D classification was attempted, but there were no enough particles of the fully assembled complex I to reconstitute an entire model. Only a low-quality model could be obtained with a truncated membrane arm. In addition, our enzymatic studies indicated that the purified protein can catalyze electron transfer from NADH to ubiquinone.

4.4.2 The putative assembly pathway of bacterial complex I

Little is known about the assembly of bacterial complex I. This process has been studied in *E. coli* [80] (see chapter 1.2.6) and a possible assembly pathway was proposed. Accordingly, the initial production of the membrane-spanning NuoA anchors the ribosome and the *nuo*-mRNA to the bacterial membrane to circumvent time-consuming diffusion processes and the use of additional chaperones for the translation of *nuo*B to *nuo*I. The ribosome stays at the membrane during the entire synthesis process of the globular subunits [158]. Membrane-spanning NuoH is translated from the polycistronic mRNA in contact with the already synthesized peripheral arm. Next, NuoJ and K are acquired and, finally, an assembly intermediate consisting of NuoL, M and N was assembled.

I agreed that the assembly of complex I already proceeds during the synthesis of the Nuo subunits, but the performed heterologous production of the *A. aeolicus* complex I (NQOR1) suggested differences in the assembly sequence of the subunits. I proposed a model mainly based on a found protein assembly on the blue native PAGE gel which is mainly composed of the complete hydrophilic arm and ribosomal proteins. Due to the fact that the genes encoding subunits of the Q-module (NuoB, D_2 and I_1) and the N-module (NuoE, F and G) were designed into two operons, the transcription, translation and assembly of the Q-module and the N-module of NQOR1 should be independent. The assembly of the N- and Q-modules to the hydrophilic arm proceeds afterwards. The N-terminal helix of $NuoI_1$ might serve as an anchor for further assembly of the complex at the membrane surface. $NuoH_1$ was first added to the assembled peripheral arm, followed

by NuoA$_2$, NuoJ$_1$, NuoK$_1$, and NuoN$_1$. Finally, NuoL$_1$ and NuoM$_1$ were attached sequentially.

4.4.3 Fe-S clusters in *A. aeolicus* complex I

A. aeolicus complex I is equipped with FMN and eight to ten Fe-S clusters which are required for electron transfer from NADH to ubiquinone. *A. aeolicus* complex I contains ten Fe-S clusters. The clusters, N1a, N1b, N2, N3, N4, N5, N6a, and N6b, are conserved, N7 cluster is discovered in some bacteria, such as *E. coli* and *T. thermophilus*. N8 cluster is a novel tetranuclear Fe-S cluster, which is only observed in *A. aeolicus*, both in amino acid sequence (Figure 4.3) and 3D structure (see chapter 1.4).
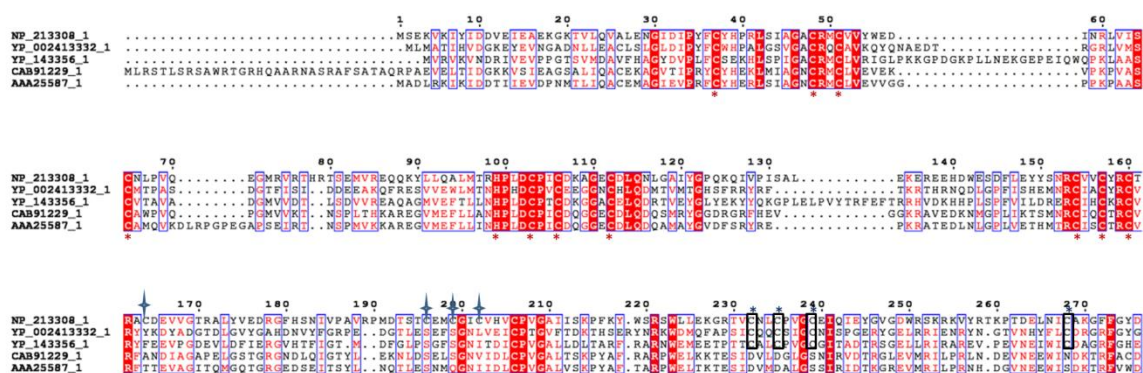

Figure 4.3. Sequence alignment of the N-terminal segment of NuoG from *A. aeolicus* VF5 with its homologues from various organisms.

NP_213308_1 *A. aeolicus* VF5*,* YP_002413332_1 *Escherichia coli* UMN026, YP_143356.1 *T. thermophilus* HB8, CAB91229.1 *N. crassa*, and AAA25587.1 *P. denitrificans.* The motif for N7 cluster was highlighted in black box, for N8 cluster by blue star.

Two Fe-S clusters (N1b and N2) could be assigned unambiguously in native *A. aeolicus* complex I by EPR measurement (see chapter 3.1.3). The major amount of NQOR1 produced in *E. coli* consisting of the hydrophilic arm, which harbors all the redox centers and is less complicated compared to the preparations from the native source, might be a better material for EPR measurement. Overall, the EPR spectra obtained is similar to that from native *A. aeolicus* complex I (see chapter 3.3.3.6). The EPR spectra of the individual Fe-S clusters cannot be distinguished so that the production of the single subunits of complex I such as the subunit containing N8 is still necessary.

4.5 Relationship between complex I and AhpC2

Previous MS data indicated that AhpC2 is a potential substrate or binding partner of *A. - aeolicus* complex I which might play a role as terminal electron donor/acceptor in an alternative electron transfer pathway via the novel Fe-S cluster N8. To test this hypothesis, we first performed a heterologous production and characterization of AhpC2. Amino acid sequence analysis classified AhpC2 as a member of the 1-Cys peroxiredoxin family, which has the capability to reduce a broad range of hydroperoxide substrates including phospholipid hydroperoxides (PLOOH) [159]. This protein could be produced in *E. coli* and purified to homogeneity. The structures of AhpC2 were determined by X-ray crystallography and EM, which both reveal a ring-shaped dodecamer. The redox potential of AhpC2 was determined at around -310 mV, which was comparable to that of NADH (around -320 mV) and lower than that of Fe-S clusters in complex I with an expectation of N1a (around -340 mV). Therefore, electron transfer between AhpC2 and complex I is possible. The interaction of AhpC2 with complex I was evaluated by SPR. The dissociation constant (KD) suggested that the AhpC2 had a very high affinity to complex I (see chapter 3.4.2). The K_D value is 0.478 nM and 4.84 nM for an AhpC2 sample treated with and without NaCl, respectively. Negative stain electron microscopic data revealed a dodecameric structure of AhpC2 that dissociates into monomers after the treatment with 1 M NaCl. Obviously, the oligomeric state of AhpC2 influences its affinity to complex I and perhaps also the function AhpC2. Co-crystallization experiments between AhpC2 and complex I were also attempted, but the diffraction power of the resulting crystals was too low for further structural analysis.

5. References

[1] Y. Hatefi, The mitochondrial electron transport and oxidative phosphorylation system, Annu Rev Biochem, 54 (1985) 1015-1069.

[2] P. Mitchell, Coupling of Phosphorylation to Electron and Hydrogen Transfer by a Chemi-Osmotic Type of Mechanism, Nature, 191 (1961) 144-&.

[3] R. Baradaran, J.M. Berrisford, G.S. Minhas, L.A. Sazanov, Crystal structure of the entire respiratory complex I, Nature, 494 (2013) 443-448.

[4] F. Sun, X. Huo, Y. Zhai, A. Wang, J. Xu, D. Su, M. Bartlam, Z. Rao, Crystal structure of mitochondrial respiratory membrane protein complex II, Cell, 121 (2005) 1043-1057.

[5] S. Iwata, J.W. Lee, K. Okada, J.K. Lee, M. Iwata, B. Rasmussen, T.A. Link, S. Ramaswamy, B.K. Jap, Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex, Science, 281 (1998) 64-71.

[6] T. Tsukihara, H. Aoyama, E. Yamashita, T. Tomizaki, H. Yamaguchi, K. ShinzawaItoh, R. Nakashima, R. Yaono, S. Yoshikawa, The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 angstrom, Science, 272 (1996) 1136-1144.

[7] L.A. Baker, I.N. Watt, M.J. Runswick, J.E. Walker, J.L. Rubinstein, Arrangement of subunits in intact mammalian mitochondrial ATP synthase determined by cryo-EM, Proc Natl Acad Sci U S A, 109 (2012) 11675-11680.

[8] L.A. Sazanov, A giant molecular proton pump: structure and mechanism of respiratory complex I, Nat Rev Mol Cell Biol, 16 (2015) 375-388.

[9] H. Leif, U. Weidner, A. Berger, V. Spehr, M. Braun, P. Vanheek, T. Friedrich, T. Ohnishi, H. Weiss, *Escherichia coli* Nadh Dehydrogenase-I, a Minimal Form of the Mitochondrial Complex-I, Biochem Soc T, 21 (1993) 998-1001.

[10] L.A. Sazanov, P. Hinchliffe, Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*, Science, 311 (2006) 1430-1436.

[11] U. Weidner, S. Geier, A. Ptock, T. Friedrich, H. Leif, H. Weiss, The gene locus of the proton-translocating NADH: ubiquinone oxidoreductase in *Escherichia coli*. Organization of the 14 genes and relationship between the derived proteins and subunits of mitochondrial complex I, J Mol Biol, 233 (1993) 109-122.

[12] C.D. Archer, T. Elliott, Transcriptional Control of the Nuo Operon Which Encodes the Energy-Conserving Nadh Dehydrogenase of *Salmonella typhimurium*, Journal of Bacteriology, 177 (1995) 2335-2342.

[13] X. Xu, A. Matsuno-Yagi, T. Yagi, DNA sequencing of the seven remaining structural genes of the gene cluster encoding the energy-transducing NADH-quinone oxidoreductase of *Paracoccus denitrificans*, Biochemistry-Us, 32 (1993) 968-981.

[14] A. Dupuis, A. Peinnequin, M. Chevallet, J. Lunardi, E. Darrouzet, B. Pierrard, V. Procaccio, J.P. Issartel, Identification of five *Rhodobacter capsulatus* genes encoding the equivalent of ND subunits of the mitochondrial NADH-ubiquinone oxidoreductase, Gene, 167 (1995) 99-104.

[15] T. Yano, S.S. Chu, V.D. Sled, T. Ohnishi, T. Yagi, 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 (1997) 4201-4211.

[16] U. Brandt, Energy converting NADH:quinone oxidoreductase (complex I), Annu Rev Biochem, 75 (2006) 69-92.

[17] J. Hirst, Why does mitochondrial complex I have so many subunits?, Biochemical Journal, 437 (2011) e1-e3.

[18] S. Papa, S. Scacco, A.M. Sardanelli, V. Petruzzella, R. Vergari, A. Signorile, Z. Technikova-Dobrova, Complex I and the cAMP cascade in human physiopathology, Biosci Rep, 22 (2002) 3-16.

[19] H.C. Au, B.B. Seo, A. Matsuno-Yagi, T. Yagi, I.E. Scheffler, The NDUFA1 gene product (MWFE protein) is essential for activity of complex I in mammalian mitochondria, P Natl Acad Sci USA, 96 (1999) 4354-4359.

[20] N. Yadava, P. Potluri, E.N. Smith, A. Bisevac, I.E. Scheffler, Species-specific and mutant MWFE proteins. Their effect on the assembly of a functional mammalian mitochondrial complex I, J Biol Chem, 277 (2002) 21221-21230.

[21] U. Schulte, V. Haupt, A. Abelmann, W. Fecke, B. Brors, T. Rasmussen, T. Friedrich, H. Weiss, A reductase/isomerase subunit of mitochondrial NADH : ubiquinone oxidoreductase (complex I) carries an NADPH and is involved in the biogenesis of the complex, Journal of Molecular Biology, 292 (1999) 569-580.

[22] J.E. Cronan, I.M. Fearnley, J.E. Walker, Mammalian mitochondria contain a soluble acyl carrier protein, Febs Letters, 579 (2005) 4892-4896.

[23] Y. Hatefi, A.G. Haavik, D.E. Griffiths, Reconstitution of the electron transport system. I. Preparation and properties of the interacting enzyme complexes, Biochem Biophys Res Commun, 4 (1961) 441-446.

[24] M. Braun, S. Bungert, T. Friedrich, Characterization of the overproduced NADH dehydrogenase fragment of the NADH:ubiquinone oxidoreductase (complex I) from *Escherichia coli*, Biochemistry-Us, 37 (1998) 1861-1867.

[25] T. Friedrich, The NADH:ubiquinone oxidoreductase (complex I) from *Escherichia coli*, Biochim Biophys Acta, 1364 (1998) 134-146.

[26] T. Yano, V.D. Sled, T. Ohnishi, T. Yagi, Expression of the 25-kilodalton iron-sulfur subunit of the energy-transducing NADH-ubiquinone oxidoreductase of *Paracoccus denitrificans*, Biochemistry-Us, 33 (1994) 494-499.

[27] T. Yano, T. Yagi, V.D. Sled, T. Ohnishi, Expression and Characterization of the 66-Kilodalton (Nqo3) Iron-Sulfur Subunit of the Proton-Translocating Nadh-Quinone Oxidoreductase of *Paracoccus denitrificans*, Journal of Biological Chemistry, 270 (1995) 18264-18270.

[28] T. Yano, S. Magnitsky, V.D. Sled', T. Ohnishi, T. Yagi, 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, Journal of Biological Chemistry, 274 (1999) 28598-28605.

[29] T. Yano, J. Sklar, E. Nakamaru-Ogiso, Y. Takahashi, T. Yagi, T. Ohnishi, 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*, Journal of Biological Chemistry, 278 (2003) 15514-15522.

[30] Y. Zu, S. Di Bernardo, T. Yagi, J. Hirst, Redox properties of the [2Fe-2S] center in the 24 kDa (NQO2) subunit of NADH:ubiquinone oxidoreductase (complex I), Biochemistry-Us, 41 (2002) 10056-10069.

[31] T. Yano, W.R. Dunham, T. Ohnishi, Characterization of the Delta mu(H)+-sensitive ubisemiquinone species (SQ(Nf)) and the interaction with cluster N2: New insight into the energy-coupled electron transfer in complex I, Biochemistry-Us, 44 (2005) 1744-1754.

[32] G. Hofhaus, H. Weiss, K. Leonard, Electron microscopic analysis of the peripheral and membrane parts of mitochondrial NADH dehydrogenase (complex I), J Mol Biol, 221 (1991) 1027-1043.

[33] V. Guenebaut, R. Vincentelli, D. Mills, H. Weiss, K.R. Leonard, Three-dimensional structure of NADH-dehydrogenase from *Neurospora crassa* by electron microscopy and conical tilt reconstruction, J Mol Biol, 265 (1997) 409-418.

[34] V. Guenebaut, A. Schlitt, H. Weiss, K. Leonard, T. Friedrich, Consistent structure between bacterial and mitochondrial NADH : ubiquinone oxidoreductase (complex I), Journal of Molecular Biology, 276 (1998) 105-112.

[35] N. Grigorieff, Structure of the respiratory NADH:ubiquinone oxidoreductase (complex I), Curr Opin Struct Biol, 9 (1999) 476-483.

[36] M. Radermacher, T. Ruiz, T. Clason, S. Benjamin, U. Brandt, V. Zickermann, The three-dimensional structure of complex I from *Yarrowia lipolytica*: A highly dynamic enzyme, Journal of Structural Biology, 154 (2006) 269-279.

[37] R.G. Efremov, L.A. Sazanov, Structure of the membrane domain of respiratory complex I, Nature, 476 (2011) 414-420.

[38] R.G. Efremov, R. Baradaran, L.A. Sazanov, The architecture of respiratory complex I, Nature, 465 (2010) 441-445.

[39] C. Hunte, V. Zickermann, U. Brandt, Functional Modules and Structural Basis of Conformational Coupling in Mitochondrial Complex I, Science, 329 (2010) 448-451.

[40] A.A. Agip, J.N. Blaza, H.R. Bridges, C. Viscomi, S. Rawson, S.P. Muench, J. Hirst, Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states, Nat Struct Mol Biol, 25 (2018) 548-556.

[41] V. Zickermann, C. Wirth, H. Nasiri, K. Siegmund, H. Schwalbe, C. Hunte, U. Brandt, Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I, Science, 347 (2015) 44-49.

[42] K. Fiedorczuk, J.A. Letts, G. Degliesposti, K. Kaszuba, M. Skehel, L.A. Sazanov, Atomic structure of the entire mammalian mitochondrial complex I, Nature, 538 (2016) 406-410.

[43] J.M. Berrisford, L.A. Sazanov, Structural basis for the mechanism of respiratory complex I, J Biol Chem, 284 (2009) 29773-29783.

[44] K.R. Vinothkumar, J. Zhu, J. Hirst, Architecture of mammalian respiratory complex I, Nature, 515 (2014) 80-84.

[45] J. Zhu, K.R. Vinothkumar, J. Hirst, Structure of mammalian respiratory complex I, Nature, 536 (2016) 354-358.

[46] J.N. Blaza, K.R. Vinothkumar, J. Hirst, Structure of the Deactive State of Mammalian Respiratory Complex I, Structure, 26 (2018) 312-319 e313.

[47] R. Guo, S. Zong, M. Wu, J. Gu, M. Yang, Architecture of Human Mitochondrial Respiratory Megacomplex I2III2IV2, Cell, 170 (2017) 1247-1257 e1212.

[48] T. Yagi, A. Matsuno-Yagi, The proton-translocating NADH-quinone oxidoreductase in the respiratory chain: the secret unlocked, Biochemistry-Us, 42 (2003) 2266-2274.

[49] T. Ohnishi, Iron-sulfur clusters semiquinones in Complex I, Bba-Bioenergetics, 1364 (1998) 186-206.

[50] T. Reda, C.D. Barker, J. Hirst, Reduction of the iron-sulfur clusters in mitochondrial NADH:ubiquinone oxidoreductase (complex I) by EuII-DTPA, a very low potential reductant, Biochemistry-Us, 47 (2008) 8885-8893.

[51] J. Hirst, Towards the molecular mechanism of respiratory complex I, Biochem J, 425 (2009) 327-339.

[52] H. Leif, V.D. Sled, T. Ohnishi, H. Weiss, T. Friedrich, Isolation and characterization of the proton-translocating NADH: ubiquinone oxidoreductase from *Escherichia coli*, Eur J Biochem, 230 (1995) 538-548.

[53] G. Yakovlev, T. Reda, J. Hirst, Reevaluating the relationship between EPR spectra and enzyme structure for the iron–sulfur clusters in NADH:quinone oxidoreductase, Proceedings of the National Academy of Sciences, 104 (2007) 12720-12725.

[54] T. Pohl, T. Bauer, K. Dorner, S. Stolpe, P. Sell, G. Zocher, T. Friedrich, Iron-sulfur cluster N7 of the NADH : Ubiquinone oxidoreductase (Complex I) is essential for stability but not involved in electron transfer, Biochemistry-Us, 46 (2007) 6588-6596.

[55] P.K. Sinha, E. Nakamaru-Ogiso, J. Torres-Bacete, M. Sato, N. Castro-Guerrero, T. Ohnishi, A. Matsuno-Yagi, T. Yagi, Electron transfer in subunit NuoI (TYKY) of *Escherichia coli* NADH:quinone oxidoreductase (NDH-1), J Biol Chem, 287 (2012) 17363-17373.

[56] M.L. Verkhovskaya, N. Belevich, L. Euro, M. Wikstrom, M.I. Verkhovsky, Real-time electron transfer in respiratory complex I, Proc Natl Acad Sci U S A, 105 (2008) 3763-3767.

[57] M. Narayanan, D.J. Gabrieli, S.A. Leung, M.M. Elguindy, C.A. Glaser, N. Saju, S.C. Sinha, E. Nakamaru-Ogiso, Semiquinone and cluster N6 signals in His-tagged proton-translocating NADH:ubiquinone oxidoreductase (complex I) from *Escherichia coli*, J Biol Chem, 288 (2013) 14310-14319.

[58] E. Nakamaru-Ogiso, T. Yano, T. Yagi, T. Ohnishi, Characterization of the iron-sulfur cluster N7 (N1c) in the subunit NuoG of the proton-translocating NADH-quinone oxidoreductase from *Escherichia coli*, J Biol Chem, 280 (2005) 301-307.

[59] L.A. Sazanov, Respiratory complex I: mechanistic and structural insights provided by the crystal structure of the hydrophilic domain, Biochemistry-Us, 46 (2007) 2275-2288.

[60] J.A. Birrell, K. Morina, H.R. Bridges, T. Friedrich, J. Hirst, Investigating the function of [2Fe-2S] cluster N1a, the off-pathway cluster in complex I, by manipulating its reduction potential, Biochemical Journal, 456 (2013) 139-146.

[61] E. Gnandt, K. Dorner, M.F.J. Strampraad, S. de Vries, T. Friedrich, The multitude of iron-sulfur clusters in respiratory complex I, Bba-Bioenergetics, 1857 (2016) 1068-1072.

[62] V.K. Moparthi, B. Kumar, Y. Al-Eryani, E. Sperling, K. Gorecki, T. Drakenberg, C. Hagerhall, Functional role of the MrpA- and MrpD-homologous protein subunits in enzyme complexes evolutionary related to respiratory chain complex I, Biochim Biophys Acta, 1837 (2014) 178-185.

[63] N.J. Hu, S. Iwata, A.D. Cameron, D. Drew, Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT, Nature, 478 (2011) 408-+.

[64] E. Screpanti, C. Hunte, Discontinuous membrane helices in transport proteins and their correlation with function, Journal of Structural Biology, 159 (2007) 261-267.

[65] C. Hunte, E. Screpanti, M. Venturi, A. Rimon, E. Padan, H. Michel, Structure of a Na+/H+ antiporter and insights into mechanism of action and regulation by pH, Nature, 435 (2005) 1197-1202.

[66] C.C. Page, C.C. Moser, X. Chen, P.L. Dutton, Natural engineering principles of electron tunnelling in biological oxidation-reduction, Nature, 402 (1999) 47-52.

[67] K. Huoponen, J. Vilkki, P. Aula, E.K. Nikoskelainen, M.L. Savontaus, A new mtDNA mutation associated with Leber hereditary optic neuroretinopathy, Am J Hum Genet, 48 (1991) 1147-1153.

[68] A.D. Vinogradov, Kinetics, Control, and Mechanism of Ubiquinone Reduction by the Mammalian Respiratory Chain-Linked Nadh-Ubiquinone Reductase, J Bioenerg Biomembr, 25 (1993) 367-375.

[69] M.D. Esposti, A. Ghelli, The Mechanism of Proton and Electron-Transport in Mitochondrial Complex-I, Bba-Bioenergetics, 1187 (1994) 116-120.

[70] U. Brandt, S. Kerscher, S. Drose, K. Zwicker, V. Zickermann, Proton pumping by NADH:ubiquinone oxidoreductase. A redox driven conformational change mechanism?, FEBS Lett, 545 (2003) 9-17.

[71] T. Ohnishi, J.C. Salerno, Conformation-driven and semiquinone-gated proton-pump mechanism in the NADH-ubiquinone oxidoreductase (complex I), Febs Letters, 579 (2005) 4555-4561.

[72] R.G. Efremov, L.A. Sazanov, The coupling mechanism of respiratory complex I - A structural and evolutionary perspective, Bba-Bioenergetics, 1817 (2012) 1785-1795.

[73] J. Nouws, L.G.J. Nijtmans, J.A. Smeitink, R.O. Vogel, Assembly factors as a new class of disease genes for mitochondrial complex I deficiency: cause, pathology and treatment options, Brain, 135 (2012) 12-22.

[74] S. Guerrero-Castillo, F. Baertling, D. Kownatzki, H.J. Wessels, S. Arnold, U. Brandt, L. Nijtmans, The Assembly Pathway of Mitochondrial Respiratory Chain Complex I, Cell Metab, 25 (2017) 128-139.

[75] L. Sanchez-Caballero, S. Guerrero-Castillo, L. Nijtmans, Unraveling the complexity of mitochondrial complex I assembly: A dynamic process, Bba-Bioenergetics, 1857 (2016) 980-990.

[76] L.E. Formosa, M.G. Dibley, D.A. Stroud, M.T. Ryan, Building a complex complex: Assembly of mitochondrial respiratory chain complex I, Semin Cell Dev Biol, 76 (2018) 154-162.

[77] A. Signes, E. Fernandez-Vizarra, Assembly of mammalian oxidative phosphorylation complexes I-V and supercomplexes, Essays Biochem, 62 (2018) 255-270.

[78] H. Erhardt, S. Steimle, V. Muders, T. Pohl, J. Walter, T. Friedrich, Disruption of individual nuo-genes leads to the formation of partially assembled NADH:ubiquinone oxidoreductase (complex I) in *Escherichia coli*, Biochim Biophys Acta, 1817 (2012) 863-871.

[79] M.A. Spero, F.O. Aylward, C.R. Currie, T.J. Donohue, Phylogenomic analysis and predicted physiological role of the proton-translocating NADH:quinone oxidoreductase (complex I) across bacteria, MBio, 6 (2015).

[80] T. Friedrich, D.K. Dekovic, S. Burschel, Assembly of the *Escherichia coli* NADH:ubiquinone oxidoreductase (respiratory complex I), Biochim Biophys Acta, 1857 (2016) 214-223.

[81] B. Py, C. Gerez, S. Angelini, R. Planel, D. Vinella, L. Loiseau, E. Talla, C. Brochier-Armanet, R. Garcia Serres, J.M. Latour, S. Ollagnier-de Choudens, M. Fontecave, F. Barras, Molecular organization, biochemical function, cellular role and evolution of NfuA, an atypical Fe-S carrier, Mol Microbiol, 86 (2012) 155-171.

[82] T. Pohl, J. Walter, S. Stolpe, J.H. Soufo, P.L. Grauman, T. Friedrich, Effects of the deletion of the *Escherichia coli* frataxin homologue CyaY on the respiratory NADH:ubiquinone oxidoreductase, BMC Biochem, 8 (2007) 13.

[83] H.T. Le, V. Gautier, F. Kthiri, A. Malki, N. Messaoudi, M. Mihoub, A. Landoulsi, Y.J. An, S.S. Cha, G. Richarme, YajL, Prokaryotic Homolog of Parkinsonism-associated Protein DJ-1, Functions as a Covalent Chaperone for Thiol Proteome, Journal of Biological Chemistry, 287 (2012) 5861-5870.

[84] L.M. Iyer, D.D. Leipe, E.V. Koonin, L. Aravind, Evolutionary history and higher order classification of AAA+ ATPases, J Struct Biol, 146 (2004) 11-31.

[85] T. Frickey, A.N. Lupas, Phylogenetic analysis of AAA proteins, J Struct Biol, 146 (2004) 2-10.

[86] J. Snider, I. Gutsche, M. Lin, S. Baby, B. Cox, G. Butland, J. Greenblatt, A. Emili, W.A. Houry, Formation of a distinctive complex between the inducible bacterial lysine decarboxylase and a novel AAA plus ATPase, Journal of Biological Chemistry, 281 (2006) 1532-1546.

[87] R. Huber, T. Wilharm, D. Huber, A. Trincone, S. Burggraf, H. Konig, R. Rachel, I. Rockinger, H. Fricke, K.O. Stetter, Aquifex-Pyrophilus Gen-Nov Sp-Nov Represents a Novel Group of Marine Hyperthermophilic Hydrogen-Oxidizing Bacteria, Syst Appl Microbiol, 15 (1992) 340-351.

[88] S. Burggraf, G.J. Olsen, K.O. Stetter, C.R. Woese, A Phylogenetic Analysis of *Aquifex pyrophilus*, Syst Appl Microbiol, 15 (1992) 352-356.

[89] C. Pitulle, Y.Q. Yang, M. Marchiani, E.R.B. Moore, J.L. Siefert, M. Aragno, P. Jurtshuk, G.E. Fox, Phylogenetic Position of the Genus Hydrogenobacter, Int J Syst Bacteriol, 44 (1994) 620-626.

[90] Y. Sekiguchi, Y. Kamagata, K. Syutsubo, A. Ohashi, H. Harada, K. Nakamura, Phylogenetic diversity of mesophilic and thermophilic granular sludges determined by 16S rRNA gene analysis, Microbiology, 144 (Pt 9) (1998) 2655-2665.

[91] G. Deckert, P.V. Warren, T. Gaasterland, W.G. Young, A.L. Lenox, D.E. Graham, R. Overbeek, M.A. Snead, M. Keller, M. Aujay, R. Huber, R.A. Feldman, J.M. Short, G.J. Olsen, R.V. Swanson, The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*, Nature, 392 (1998) 353-358.

[92] M. Neuberger, DNA Cloning - a Practical Approach, Vol 1-2 - Glover,Dm, Nature, 317 (1985) 679-679.

[93] B. Miroux, J.E. Walker, Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels, J Mol Biol, 260 (1996) 289-298.

[94] E.A. Raleigh, N.E. Murray, H. Revel, R.M. Blumenthal, D. Westaway, A.D. Reith, P.W. Rigby, J. Elhai, D. Hanahan, McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning, Nucleic Acids Res, 16 (1988) 1563-1575.

[95] B. Amarneh, J. De Leon-Rangel, S.B. Vik, Construction of a deletion strain and expression vector for the *Escherichia coli* NADH:ubiquinone oxidoreductase (Complex I), Biochim Biophys Acta, 1757 (2006) 1557-1560.

[96] S. Surade, M. Klein, P.C. Stolt-Bergner, C. Muenke, A. Roy, H. Michel, Comparative analysis and "expression space" coverage of the production of prokaryotic membrane proteins for structural genomics, Protein Sci, 15 (2006) 2178-2189.

[97] L.M. Guzman, D. Belin, M.J. Carson, J. Beckwith, Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter, J Bacteriol, 177 (1995) 4121-4130.

[98] C. Li, A. Wen, B. Shen, J. Lu, Y. Huang, Y. Chang, FastCloning: a highly simplified, purification-free, sequence- and ligation-independent PCR cloning method, BMC Biotechnol, 11 (2011) 92.

[99] R.S. Haun, I.M. Serventi, J. Moss, Rapid, reliable ligation-independent cloning of PCR products using modified plasmid vectors, Biotechniques, 13 (1992) 515-518.

[100] C. Aslanidis, P.J. de Jong, Ligation-independent cloning of PCR products (LIC-PCR), Nucleic Acids Res, 18 (1990) 6069-6074.

[101] D. Hanahan, Studies on transformation of *Escherichia coli* with plasmids, J Mol Biol, 166 (1983) 557-580.

[102] P.K. Smith, R.I. Krohn, G.T. Hermanson, A.K. Mallia, F.H. Gartner, M.D. Provenzano, E.K. Fujimoto, N.M. Goeke, B.J. Olson, D.C. Klenk, Measurement of Protein Using Bicinchoninic Acid, Anal Biochem, 150 (1985) 76-85.

[103] G. Peng, G. Fritzsch, V. Zickermann, H. Schagger, R. Mentele, F. Lottspeich, M. Bostina, M. Radermacher, R. Huber, K.O. Stetter, H. Michel, Isolation, characterization and electron microscopic single particle analysis of the NADH:ubiquinone oxidoreductase (complex I) from the hyperthermophilic eubacterium *Aquifex aeolicus*, Biochemistry-Us, 42 (2003) 3032-3039.

[104] M. Sabar, J. Balk, C.J. Leaver, Histochemical staining and quantification of plant mitochondrial respiratory chain complexes using blue-native polyacrylamide gel electrophoresis, The Plant Journal, 44 (2005) 893-901.

[105] I. Wittig, M. Karas, H. Schagger, High resolution clear native electrophoresis for in-gel functional assays and fluorescence studies of membrane protein complexes, Mol Cell Proteomics, 6 (2007) 1215-1225.

[106] E. Estornell, R. Fato, F. Pallotti, G. Lenaz, Assay conditions for the mitochondrial NADH:coenzyme Q oxidoreductase, FEBS Lett, 332 (1993) 127-131.

[107] P. Infossi, E. Lojou, J.P. Chauvin, G. Herbette, M. Brugna, M.T. Giudici-Orticoni, *Aquifex aeolicus* membrane hydrogenase for hydrogen biooxidation: Role of lipids and physiological partners in enzyme stability and activity, Int J Hydrogen Energ, 35 (2010) 10778-10789.

[108] A. Garofano, K. Zwicker, S. Kerscher, P. Okun, U. Brandt, Two aspartic acid residues in the PSST-homologous NUKM subunit of complex I from *Yarrowia lipolytica* are essential for catalytic activity, J Biol Chem, 278 (2003) 42435-42440.

[109] L.G. Fägerstam, Biospecific Interaction Analysis in Real Time Using a Biosensor System with Surface Plasmon Resonance Detection, in: E. Kurstak, R.G. Marusyk, F.A. Murphy, M.H.V. Van Regenmortel (Eds.) Applied Virology Research: New Diagnostic Procedures, Springer US, Place Published, 1994, pp. 113-129.

[110] S. Bazan, E. Mileykovskaya, V.K. Mallampalli, P. Heacock, G.C. Sparagna, W. Dowhan, Cardiolipin-dependent reconstitution of respiratory supercomplexes from purified *Saccharomyces cerevisiae* complexes III and IV, J Biol Chem, 288 (2013) 401-411.

[111] G. Tang, L. Peng, P.R. Baldwin, D.S. Mann, W. Jiang, I. Rees, S.J. Ludtke, EMAN2: An extensible image processing suite for electron microscopy, J Struct Biol, 157 (2007) 38-46.

[112] J.A. Mindell, N. Grigorieff, Accurate determination of local defocus and specimen tilt in electron microscopy, J Struct Biol, 142 (2003) 334-347.

[113] S.H.W. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination, J Struct Biol, 180 (2012) 519-530.

[114] Y. Thielmann, J. Koepke, H. Michel, The ESFRI Instruct Core Centre Frankfurt: automated high-throughput crystallization suited for membrane proteins and more, Journal of structural and functional genomics, 13 (2012) 63-69.

[115] W. Kabsch, Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants, Journal of Applied Crystallography, 26 (1993) 795-800.

[116] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, Journal of Applied Crystallography, 40 (2007) 658-674.

[117] T. Nakamura, A. Mori, M. Niiyama, H. Matsumura, C. Tokuyama, J. Morita, K. Uegaki, T. Inoue, Structure of peroxiredoxin from the anaerobic hyperthermophilic archaeon *Pyrococcus horikoshii*, Acta Crystallographica Section F, 69 (2013) 719-722.

[118] G.N. Murshudov, P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, Acta Crystallographica Section D, 67 (2011) 355-367.

[119] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta crystallographica. Section D, Biological crystallography, 66 (2010) 213-221.

[120] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallographica Section D, 66 (2010) 486-501.

[121] M.J. Stark, Multicopy expression vectors carrying the lac repressor gene for regulated high-level expression of genes in *Escherichia coli*, Gene, 51 (1987) 255-267.

[122] M. Kohlstadt, K. Dorner, R. Labatzke, C. Koc, R. Heilscher, E. Schiltz, O. Einsle, P. Hellwig, T. Friedrich, Heterologous production, isolation, characterization and crystallization of a soluble fragment of the NADH:ubiquinone oxidoreductase (complex I) from *Aquifex aeolicus*, Biochemistry-Us, 47 (2008) 13036-13045.

[123] G. Peng, B. Meyer, L. Sokolova, W. Liu, S. Bornemann, J. Juli, K. Zwicker, M. Karas, B. Brutschy, H. Michel, Identification and characterization two isoforms of NADH:ubiquinone oxidoreductase from the hyperthermophilic eubacterium *Aquifex aeolicus*, Biochim Biophys Acta, 1859 (2018) 366-373.

[124] R. Djafarzadeh, S. Kerscher, K. Zwicker, M. Radermacher, M. Lindahl, H. Schagger, U. Brandt, Biophysical and structural characterization of proton-translocating NADH-dehydrogenase (complex I) from the strictly aerobic yeast *Yarrowia lipolytica*, Bba-Bioenergetics, 1459 (2000) 230-238.

[125] L. Euro, D.A. Bloch, M. Wikstrom, M.I. Verkhovsky, M. Verkbovskaya, Electrostatic interactions between FeS clusters in NADH : Ubiquinone oxidoreductase (complex I) from *Escherichia coli*, Biochemistry-Us, 47 (2008) 3185-3193.

[126] L. Soito, C. Williamson, S.T. Knutson, J.S. Fetrow, L.B. Poole, K.J. Nelson, PREX: PeroxiRedoxin classification indEX, a database of subfamily assignments across the diverse peroxiredoxin family, Nucleic Acids Research, 39 (2011) D332-D337.

[127] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Molecular Systems Biology, 7 (2011).

[128] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, Nucleic Acids Research, 42 (2014) W320-W324.

[129] M. Hirasawa, P. Schurmann, J.P. Jacquot, W. Manieri, P. Jacquot, E. Keryer, F.C. Hartman, D.B. Knaff, Oxidation-reduction properties of chloroplast thioredoxins, Ferredoxin : Thioredoxin reductase, and thioredoxin f-regulated enzymes, Biochemistry-Us, 38 (1999) 5200-5205.

[130] R.A. Laskowski, E.G. Hutchinson, A.D. Michie, A.C. Wallace, M.L. Jones, J.M. Thornton, PDBsum: a Web-based database of summaries and analyses of all PDB structures, Trends Biochem Sci, 22 (1997) 488-490.

[131] L. Holm, L.M. Laakso, Dali server update, Nucleic Acids Res, 44 (2016) W351-355.

[132] T. Nakamura, T. Yamamoto, T. Inoue, H. Matsumura, A. Kobayashi, Y. Hagihara, K. Uegaki, M. Ataka, Y. Kai, K. Ishikawa, Crystal structure of thioredoxin peroxidase from aerobic hyperthermophilic archaeon *Aeropyrum pernix* K1, Proteins, 62 (2006) 822-826.

[133] V.M. Markowitz, I.M. Chen, K. Palaniappan, K. Chu, E. Szeto, Y. Grechkin, A. Ratner, I. Anderson, A. Lykidis, K. Mavromatis, N.N. Ivanova, N.C. Kyrpides, The integrated microbial genomes system: an expanding comparative analysis resource, Nucleic Acids Res, 38 (2010) D382-390.

[134] R. Munoz, P. Yarza, W. Ludwig, J. Euzeby, R. Amann, K.H. Schleifer, F.O. Glockner, R. Rossello-Mora, Release LTPs104 of the All-Species Living Tree, Syst Appl Microbiol, 34 (2011) 169-170.

[135] M.G. Dibley, M.T. Ryan, D.A. Stroud, A novel isoform of the human mitochondrial complex I subunit NDUFV3, FEBS Lett, 591 (2017) 109-117.

[136] S. Guerrero-Castillo, A. Cabrera-Orefice, M.A. Huynen, S. Arnold, Identification and evolutionary analysis of tissue-specific isoforms of mitochondrial complex I subunit NDUFV3, Biochim Biophys Acta Bioenerg, 1858 (2017) 208-217.

[137] P.W. Gunning, Protein Isoforms and Isozymes, eLS, Place Published, 2006.

[138] S. Arnold, The power of life--cytochrome c oxidase takes center stage in metabolic control, cell signalling and survival, Mitochondrion, 12 (2012) 46-56.

[139] S. Arnold, Cytochrome c oxidase and its role in neurodegeneration and neuroprotection, Adv Exp Med Biol, 748 (2012) 305-339.

[140] S. Rahman, Y. Ishizuka-Katsura, S. Arai, S. Saijo, I. Yamato, M. Toyama, N. Ohsawa, M. Inoue, K. Honda, T. Terada, M. Shirouzu, S. Yokoyama, S. Iwata, T. Murata, Expression, purification and characterization of isoforms of peripheral stalk subunits of human V-ATPase, Protein Expr Purif, 78 (2011) 181-188.

[141] H.R. Bridges, K. Mohammed, M.E. Harbour, J. Hirst, Subunit NDUFV3 is present in two distinct isoforms in mammalian complex I, Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1858 (2017) 197-207.

[142] J. Bongaerts, S. Zoske, U. Weidner, G. Unden, Transcriptional Regulation of the Proton-Translocating Nadh Dehydrogenase Genes (Nuoa-N) of *Escherichia coli* by Electron-Acceptors, Electron-Donors and Gene Regulators, Molecular Microbiology, 16 (1995) 521-534.

[143] M.M. Zambrano, R. Kolter, *Escherichia coli* mutants lacking NADH dehydrogenase I have a competitive disadvantage in stationary phase, Journal of Bacteriology, 175 (1993) 5642-5647.

[144] H.J. Falk-Krzesinski, A.J. Wolfe, Genetic analysis of the nuo locus, which encodes the proton-translocating NADH dehydrogenase in *Escherichia coli*, J Bacteriol, 180 (1998) 1174-1184.

[145] C. Klinger, M. Rossbach, R. Howe, M. Kaufmann, Thermophile-specific proteins: the gene product of aq_1292 from *Aquifex aeolicus* is an NTPase, BMC Biochem, 4 (2003) 12.

[146] M. Obuchi, K. Kawahara, D. Motooka, S. Nakamura, M. Yamanaka, T. Takeda, S. Uchiyama, Y. Kobayashi, T. Ohkubo, Y. Sambongi, Hyperstability and crystal structure of cytochrome c555 from hyperthermophilic *Aquifex aeolicus*, Acta Crystallographica Section D, 65 (2009) 804-813.

[147] M. Schutz, B. Schoepp-Cothenet, E. Lojou, M. Woodstra, D. Lexa, P. Tron, A. Dolla, M.C. Durand, K.O. Stetter, F. Baymann, The naphthoquinol oxidizing cytochrome bc1 complex of the hyperthermophilic knallgasbacterium *Aquifex aeolicus*: properties and phylogenetic relationships, Biochemistry-Us, 42 (2003) 10800-10808.

[148] S.E. Rollauer, M.J. Tarry, J.E. Graham, M. Jaaskelainen, F. Jager, S. Johnson, M. Krehenbrink, S.M. Liu, M.J. Lukey, J. Marcoux, M.A. McDowell, F. Rodriguez, P. Roversi, P.J. Stansfeld, C.V. Robinson, M.S. Sansom, T. Palmer, M. Hogbom, B.C. Berks, S.M. Lea, Structure of the TatC core of the twin-arginine protein transport system, Nature, 492 (2012) 210-214.

[149] T. Matsui, M. Yoshida, Expression of the wild-type and the Cys-/Trp-less alpha 3 beta 3 gamma complex of thermophilic F1-ATPase in *Escherichia coli*, Biochim Biophys Acta, 1231 (1995) 139-146.

[150] D.G. McMillan, S. Keis, P. Dimroth, G.M. Cook, A specific adaptation in the a subunit of thermoalkaliphilic F1FO-ATP synthase enables ATP synthesis at high pH but not at neutral pH values, J Biol Chem, 282 (2007) 17395-17404.

[151] C. Zhang, M. Allegretti, J. Vonck, J.D. Langer, M. Marcia, G. Peng, H. Michel, Production of fully assembled and active Aquifex aeolicus F1FO ATP synthase in *Escherichia coli*, Biochimica et Biophysica Acta (BBA) - General Subjects, 1840 (2014) 34-40.

[152] S. Bungert, B. Krafft, R. Schlesinger, T. Friedrich, One-step purification of the NADH dehydrogenase fragment of the *Escherichia coli* complex I by means of Strep-tag affinity chromatography, FEBS Lett, 460 (1999) 207-211.

[153] V. Spehr, A. Schlitt, D. Scheide, V. Guenebaut, T. Friedrich, Overexpression of the *Escherichia coli* nuo-operon and isolation of the overproduced NADH:ubiquinone oxidoreductase (complex I), Biochemistry-Us, 38 (1999) 16261-16267.

[154] T. Pohl, M. Uhlmann, M. Kaufenstein, T. Friedrich, Lambda Red-mediated mutagenesis and efficient large scale affinity purification of the *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I), Biochemistry-Us, 46 (2007) 10694-10702.

[155] E. Nakamaru-Ogiso, A. Matsuno-Yagi, S. Yoshikawa, T. Yagi, T. Ohnishi, Iron-sulfur cluster N5 is coordinated by an HXXXCXXCXXXXXC motif in the NuoG subunit of *Escherichia coli* NADH : quinone oxidoreductase (complex I), Journal of Biological Chemistry, 283 (2008) 25979-25987.

[156] M. Vranas, Heterologous production, purification and characterization of respiratory proteins from *Aquifex aeolicus* in *Escherichia coli* strains, Place Published, 2013.

[157] K. Terpe, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, Appl Microbiol Biotechnol, 60 (2003) 523-533.

[158] T. Bornemann, J. Jöckel, M.V. Rodnina, W. Wintermeyer, Signal sequence–independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel, Nature Structural &Amp; Molecular Biology, 15 (2008) 494.

[159] A.B. Fisher, C. Dodia, Y. Manevich, J.W. Chen, S.I. Feinstein, Phospholipid hydroperoxides are substrates for non-selenium glutathione peroxidase, J Biol Chem, 274 (1999) 21326-21334.

6. Appendix

Table S1. Data collection and refinement statistics

Dataset	
Space group	P1
Cell dimensions	a= 51.1 b=141.5 c=143.0 α=61.00 β=79.00 γ=80.90
Wavelength (Å)	1.00067
Resolution range (Å)	20-1.8 (1.9-1.8)
Observed reflections	1050439 (144788)
Unique reflections	306783 (45076)
Completeness (%)	96.3 (94.2)
Redundancy	3.4 (3.2)
$<I/\sigma I>$	12.2 (2.48)
Refinement	
Resolution range (Å)	20-1.8 (1.82-1.80)
Rcryst (%)	19.0 (32.8)
Rfree (%)	22.0 (35.7)
R.m.s. deviation bonds (Å)	0.013
R.m.s. deviation angles (°)	1.265
No. of amino acids in AU	2604
Solvent	1091

Table S2. Identification of BN-PAGE gel band 6

Application	Description	Coverage [%]	PSMs	Unique Peptides
Components of *A. aeolicus* complex I hydrophilic arm	NADH dehydrogenase I subunit D [OS=*Aquifex aeolicus* VF5]	78	121	47
	NADH dehydrogenase I subunit G [OS=*Aquifex aeolicus* VF5]	72	123	50
	NADH dehydrogenase I subunit F [OS=*Aquifex aeolicus* VF5]	58	60	30
	NADH dehydrogenase I subunit I [OS=*Aquifex aeolicus* VF5]	54	23	13
	NADH dehydrogenase I subunit B [OS=*Aquifex aeolicus* VF5]	47	19	11
	NADH dehydrogenase I subunit E [OS=*Aquifex aeolicus* VF5]	36	16	8
	NADH dehydrogenase I chain H [OS=Aquifex aeolicus VF5]	2	1	1
Transcription	Transcription termination factor Rho [OS=*Escherichia coli* K-12]	7	2	2
Ribosome protein for protein synthesis	50S ribosomal protein L5 [OS=*Escherichia coli* K-12]	67	13	11
	30S ribosomal protein S5 [OS=*Escherichia coli* K-12]	62	9	7
	30S ribosomal protein S3 [OS=*Escherichia coli* K-12]	33	7	6
	30S ribosomal protein S10 [OS=*Escherichia coli* K-12]	35	4	4
	50S ribosomal protein L13 [OS=*Escherichia coli* K-12]	33	4	4

	50S ribosomal protein L3 [OS=*Escherichia coli* K-12]	23	6	4
	50S ribosomal protein L4 [OS=*Escherichia coli* K-12]	15	3	3
	chaperone protein DnaJ [OS=*Escherichia coli* K-12]	11	3	3
	30S ribosomal protein S1 [OS=*Escherichia coli* K-12]	8	6	3
	formate acetyltransferase 1 [OS=*Escherichia coli* K-12]	4	3	3
	30S ribosomal protein S13 [OS=*Escherichia coli* K-12]	26	2	2
	50S ribosomal protein L14 [OS=*Escherichia coli* K-12]	23	2	2
	50S ribosomal protein L28 [OS=*Escherichia coli* K-12]	23	2	2
	50S ribosomal protein L1 [OS=*Escherichia coli* K-12]	11	2	2
	30S ribosomal protein S7 [OS=*Escherichia coli* K-12]	9	2	2
	50S ribosomal protein L27 [OS=*Escherichia coli* K-12]	16	1	1
	30S ribosomal protein S9 [OS=*Escherichia coli* K-12]	9	1	1
	50S ribosomal protein L22 [OS=*Escherichia coli* K-12]	9	1	1
	50S ribosomal protein L18 [OS=*Escherichia coli* K-12]	8	1	1

	50S ribosomal protein L11 [OS=*Escherichia coli* K-12]	7	1	1
	50S ribosomal protein L23 [OS=*Escherichia coli* K-12]	7	1	1
	30S ribosomal protein S8 [OS=*Escherichia coli* K-12]	6	1	1
	50S ribosomal protein L17 [OS=*Escherichia coli* K-12]	6	1	1
	50S ribosomal protein L6 [OS=*Escherichia coli* K-12]	4	1	1
	30S ribosomal protein S2 [OS=*Escherichia coli* K-12]	3	1	1
GTPase that associates with the 50S ribosomal subunit and may have a role during protein synthesis or ribosome biogenesis	GTPase HflX [OS=*Escherichia coli* K-12]	4	2	2
Polypeptide elongation during protein synthesis	Elongation factor Tu 1 [OS=*Escherichia coli* K-12]	50	27	16
Molecular chaperones	60 kDa chaperonin [OS=*Escherichia coli* K-12]	4	2	2
Chaperone activity	ATP-dependent protease ATPase subunit HslU [OS=*Escherichia coli* K-12]	2	1	1
Controls the flux of glucose into the hexosamine pathway	Glutamine--fructose-6-phosphate aminotransferase [isomerizing] [OS=*Escherichia coli* K-12]	2	1	1
Regulation of glycerol uptake and metabolism	glycerol kinase [OS=*Escherichia coli* K-12]	2	1	1
Galactose metabolism	D-tagatose-1,6-bisphosphate aldolase subunit gatZ	15	4	4

	[OS=*Escherichia coli* K-12]			
Glycolytic process	Pyruvate dehydrogenase E1 component [OS=*Escherichia coli* K-12]	52	57	36
Converts pyruvate, the final product of glycolysis, to lactate when oxygen is absent or in short supply	L-lactate dehydrogenase [OS=*Escherichia coli* K-12]	3	1	1
Tricarboxylic acid cycle	Succinate dehydrogenase flavoprotein subunit [OS=*Escherichia coli* K-12]	2	1	1
Tricarboxylic acid cycle	2-oxoglutarate dehydrogenase E1 component [OS=*Escherichia coli* K-12]	16	11	10
Component of terminal oxidase, in the aerobic respiratory chain	Cytochrome bd-I ubiquinol oxidase subunit 1 [OS=*Escherichia coli* K-12]	4	3	2
Electron transfer subunit of the terminal reductase	Anaerobic dimethyl sulfoxide reductase chain B [OS=*Escherichia coli* K-12]	5	1	1
ATP synthesis and/or hydrolysis	ATP synthase subunit beta [OS=*Escherichia coli* K-12]	56	29	18
ATP synthesis and/or hydrolysis	ATP synthase subunit alpha [OS=*Escherichia coli* K-12]	35	16	14
In aerobic conditions it acts as a hydrogen peroxide scavenger	Aldehyde-alcohol dehydrogenase [OS=*Escherichia coli* K-12]	3	2	2
Biosynthesis of nucleotides (purines and pyrimidines), cofactors NAD and NADP, and amino acids histidine and tryptophan	Ribose-phosphate pyrophosphokinase [OS=*Escherichia coli* K-12]	8	2	2
Aspartate metabolic process amino acid biosynthesis and	Aspartate ammonia-lyase [OS=*Escherichia coli* K-12]	7	1	1

tricarboxylic acid cycle				
Catalyse the oxidation of D-amino acids into their corresponding oxoacids	D-amino acid dehydrogenase [OS=*Escherichia coli* K-12]	5	2	2
Receptor for the attractant L-aspartate and related amino and dicarboxylic acids.	methyl-accepting chemotaxis protein II [OS=*Escherichia coli* K-12]	3	1	1
Threonine biosynthesis	Threonine--tRNA ligase [OS=*Escherichia coli* K-12]	2	1	1
Lysine biosynthesis	Lysine--tRNA ligase [OS=*Escherichia coli* K-12]	4	2	2
acetyl-CoA biosynthesis	Phosphate acetyltransferase [OS=*Escherichia coli* K-12]	3	2	2
Hydrogen cycling	Hydrogenase-1 small chain [OS=*Escherichia coli* K-12]	5	1	1
Maintain a near-neutral intracellular pH when cells are exposed to extremely acidic conditions	Glutamate decarboxylase alpha [OS=*Escherichia coli* K-12]	4	1	1
Involved in formation of the rod shape of the cell	Rod shape-determining protein MreB [OS=*Escherichia coli* K-12]	4	1	1
Essential cell division protein	Cell division protein FtsZ [OS=*Escherichia coli* K-12]	7	2	2
Bacterial cell wall biosynthesis	D-alanyl-D-alanine carboxypeptidase dacB [OS=*Escherichia coli* K-12]	3	1	1
	Uncharacterized protein yubM [OS=*Escherichia coli* K-12]	4	1	1
Fatty acid biosynthesis	3-oxoacyl-[acyl-carrier-protein] synthase 1 [OS=*Escherichia coli* K-12]	3	1	1

Biosynthesis of lipid A	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase [OS=*Escherichia coli* K-12]	8	2	1
ATP-dependent breakage, passage and rejoining of double-stranded DNA.	DNA topoisomerase 4 subunit A [OS=*Escherichia coli* K-12]	1	1	1
DNA recombinational repair	Protein RecA [OS=*Escherichia coli* K-12]	3	1	1
Protein translocase	Protein translocase subunit SecF [OS=*Escherichia coli* K-12]	4	1	1
ATP-dependent zinc metallopeptidase for both cytoplasmic and membrane proteins	ATP-dependent zinc metalloprotease FtsH [OS=*Escherichia coli* K-12]	1	1	1
Metallopeptidase activity	UPF0758 protein ykfG [OS=*Escherichia coli* K-12]	10	1	1
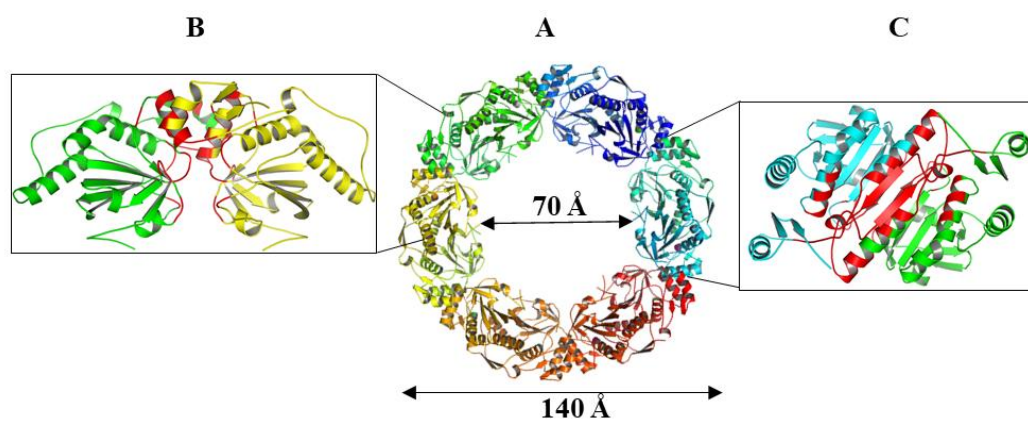
Figure S1. Overall structure of AhpC2. (A) Top view of the overall structure. Twelve subunits are arranged as a ring-shaped dodecamer through the interactions of six B-type dimers via A-type dimeric interface. Each monomer is shown in different colors. (B) A-type dimer. (C) B-type dimer. The elements involved in the dimer interactions are highlighted in red.

performing all the electron microscopic experiments, the surface plasmon resonance assays. **Prof. Song Gao's group** at Department of Chemistry, Beijing University For the initial investigations on electron paramagnetic resonance measurements for the recombinant NQOR1 samples.

Prof. Dr. Thomas Prisner and **Dr. Alberto collauto** at the Johann Wolfgang Goethe University, **Dr. Klaus Zwicker** at Universitätsklinikum Frankfurt for the collaboration on the electron paramagnetic resonance measurements.

The **INSTRUCT Core Centre** crystallization (**Barbara Rathman**, **Dr. Yvonne Thielmann**) and mass spectrometry facilities (**Dr. Julian Langer**, **Imke Wüllenweber**, **Martin Eisinger**, and **Fiona Rupprecht**) for support during this thesis work.

My friends, **Sandra Zakrzewska**, **Kanwal Kayastha**, **Tejaswi Kalavacherla**, **Hui Zeng**, **Radhika Khera**, and **Aditya Patra**, for their precious friendship, for all the good time together, and for their support during the disappointed period.

My parents for their unconditional loves throughout my life. They are open-minded and always support me, trust me. I am able to overcome all the difficulties with their strongest backup.

Curriculum vitae

Personal Information

Name	Wenxia Liu	
Date of birth	7th August 1987	
Place of birth	Laiwu, Shandong, China	

Educations

9/2006 - 7/2010	• Bachelor at Ludong University, Yantai, Shandong, China • Major: Biological Engineering • GPA: 3.57/4
10/2010 - 7/2013	• Master at Institute of Microbiology Chinese Academy of Science, Beijing, China • Major: Microbiology • Supervisor: Prof. Yanfen Xue and Prof. Yanhe Ma
11/2013 - present	• Ph.D. student at Max Planck Institute of Biophysics, Frankfurt am Main, Germany • Department: Molecular Membrane Biology • Group leader: Dr. Guohong Peng • Internal supervisor: Prof. Dr. Klaas Martinus Pos • External supervisor: Prof. Dr. Hartmut Michel

Conferences

- CEF-Symposium 2015: Structures, Mechanisms, and Dynamics of Membrane Complexes
- International Symposium, New Horizons in Membrane Transport and Communication, October 4-6, 2017
- 20th European Bioenergetics Conference Budapest, August 25-30, 2018