

The N-terminal domains of the paralogous HycE and NuoCD govern assembly of the respective formate hydrogenlyase and NADH dehydrogenase complexes

Philipp Skorupa¹, Ute Lindenstrauß¹, Sabrina Burschel², Christian Blumenschein¹ ,
Thorsten Friedrich²  and Constanze Pinske¹ 

¹ Institute of Biology/Microbiology, Martin-Luther University Halle-Wittenberg, Germany

² Institute of Biochemistry/Molecular Bioenergetics, Albert-Ludwigs-University Freiburg, Germany

Keywords

complex I; formate hydrogenlyase; fusion proteins; HycE; NADH:ubiquinone oxidoreductase; NuoCD

Correspondence

C. Pinske, Institute of Biology/Microbiology, Martin-Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3, Halle 06120, Germany
Tel: +49 345 5526353
E-mail: constanze.pinske@mikrobiologie.uni-halle.de

Philipp Skorupa and Ute Lindenstrauß contributed equally to this work

(Received 22 October 2019, revised 17 December 2019, accepted 8 January 2020)

doi:10.1002/2211-5463.12787

Formate hydrogenlyase (FHL) is the main hydrogen-producing enzyme complex in enterobacteria. It converts formate to CO₂ and H₂ via a formate dehydrogenase and a [NiFe]-hydrogenase. FHL and complex I are evolutionarily related and share a common core architecture. However, complex I catalyses the fundamentally different electron transfer from NADH to quinone and pumps protons. The catalytic FHL subunit, HycE, resembles NuoCD of *Escherichia coli* complex I; a fusion of NuoC and NuoD present in other organisms. The C-terminal domain of HycE harbours the [NiFe]-active site and is similar to other hydrogenases, while this domain in NuoCD is involved in quinone binding. The N-terminal domains of these proteins do not bind cofactors and are not involved in electron transfer. As these N-terminal domains are separate proteins in some organisms, we removed them in *E. coli* and observed that both FHL and complex I activities were essentially absent. This was due to either a disturbed assembly or to complex instability. Replacing the N-terminal domain of HycE with a 180 amino acid *E. coli* NuoC protein fusion did not restore activity, indicating that the domains have complex-specific functions. A FHL complex in which the N- and C-terminal domains of HycE were physically separated still retained most of its FHL activity, while the separation of NuoCD abolished complex I activity completely. Only the FHL complex tolerates physical separation of the HycE domains. Together, the findings strongly suggest that the N-terminal domains of these proteins are key determinants in complex assembly.

Protein tertiary structure describes the autonomously folding regions of a protein as a domain. Domains can serve different functions, for example binding of cofactors, recognition of a motif or catalytic activity. Evolution has shuffled domains to create the huge diversity of proteins that occur today [1]. Taken on their own, a domain sometimes behaves completely different compared to the holo-protein, as observed for fibril formation by the isolated N-terminal acylphosphatase

domain of the HypF hydrogenase maturation factor [2,3]. Artificial fusion proteins have been used for many years as biochemical tools, and in many cases, these fusions have no or little impact on the function of the original protein.

It has been observed that proteins possessing coordinate functions sometimes become fused during evolution, allowing greater efficiency, based on collocation. One prominent example is the alcohol dehydrogenase

Abbreviations

BV, benzyl viologen; FdhH, formate dehydrogenase; FHL, formate hydrogenlyase; Nuo, NADH:ubiquinone oxidoreductase.

fusion with acetaldehyde dehydrogenase, which couples two sequential reactions without releasing the toxic intermediate acetaldehyde [4,5]. The most straightforward way for these fusion events to occur is when neighbouring genes of a polycistronic operon are joined through frameshift mutations. Clearly, evolutionary selective pressure is required for the proteins to remain fused, for example when the activity is enhanced or assembly is augmented. Protein complexes with several subunits are likely to assemble via ordered pathways, and natural selection tends to favour gene fusions to optimize assembly [6]. Gene fusion seems also to have happened to the *hycE* gene of the *Escherichia coli* hydrogenase three large subunit, HycE. Together with a formate dehydrogenase (FdhH), the HycE hydrogenase forms the catalytic components of the formate hydrogenlyase (FHL) complex in *E. coli* and other Enterobacteriaceae. FHL catalyses the oxidation of formate with the concomitant reduction of protons to produce hydrogen to detoxify formate during mixed-acid fermentation [7]. Along with the FdhH and HycE, three iron-sulfur proteins, including the small subunit HycG, HycF and HycB, form the soluble part of the FHL complex. This large soluble domain is anchored to the membrane by two membrane subunits [8]. The complex is phylogenetically and structurally related to the respiratory complex I, with the NuoCD protein being the homologue of HycE [9–12]. Although the relationship is less obvious on the primary structural level, where HycE and NuoCD share only 26% (145/563) identical and 43% (244/563) similar residues, it is clearly apparent when the secondary structural elements are compared between the two proteins (Fig. 1). Complex I, or NADH:ubiquinone oxidoreductase (Nuo), is the primary energy-conserving complex of many respiratory chains and couples NADH oxidation to the translocation of four protons across the membrane [13]. NuoC varies significantly in length in different organisms [14]. Unlike most organisms, which have separate NuoC and NuoD proteins, both are fused in complex I of *E. coli* and *Bacteroides fragilis*. Moreover, the NuoC subunit is homologous to the N-terminal domain of HycE and its paralogue in *E. coli*, HyfG [14]. The crystal structures of both the soluble portion and the entire complex from *Thermus thermophilus* have been determined [15,16]. NuoC is an orthologue of Nqo5 in *T. thermophilus* and of the 30-kDa protein in *Bos taurus* [13]. Unlike its counterpart in HycE, which contains the [NiFe]-active site, NuoD is not known to harbour any cofactors, but it does have a quinone-binding site [17].

The C-terminal domain of HycE extends from amino acids 175–537 and harbours the four conserved cysteinyl residues that coordinate the bimetallic [NiFe]-cofactor required for catalysis (scheme in Fig. 1 and alignment in Fig. 2). These cysteinyl residues are not present in NuoD (Figs 1 and 2). Six pleiotropic Hyp proteins are required for synthesis of the [NiFe]-cofactor including its diatomic ligands attached to the Fe atom [18,19]. Of these proteins, the HypC and HypA proteins interact directly with HycE for Fe (CN)₂CO and Ni²⁺ insertion, respectively [20,21]. In addition, HycE has a 32-amino acid C-terminal extension from amino acids 538–569 that is endoproteolytically processed after cofactor insertion and is assumed to facilitate this step by keeping the empty, cofactor-free apo-protein in an open conformation [22]. In this context, the N terminus of the unprocessed hydrogenase large subunit from *Thermococcus kodakarensis* was found to extend into the nickel-delivery protein HypA, while the C-terminal extension replaces the position of the N terminus prior to proteolytic processing [23]. A recent DFT calculation on HycE identified conserved glutamate residues in both the N- and C-terminal domains, which are proposed to govern the insertion of the Ni²⁺ ion of the cofactor, but the function of these residues has yet to be validated experimentally for HycE [24]. This strongly suggests that both domains act in concert for cofactor insertion.

The N-terminal domain of HycE, corresponding to NuoC, extends from amino acids 9–194 and thus is predicted to have a short overlap with the C-terminal domain [25]. This domain/protein bears a cofactor neither in the FHL complex nor in complex I, and its role remains enigmatic.

The N- and C-terminal domains of HycE and NuoCD are predicted to be fused via a long unstructured loop region that contacts the formate dehydrogenase or diaphorase subunits in FHL and complex I, respectively (Fig. 1). Previous work has shown that conserved residues (Glu-138, Glu-140, Asp-143) in the N-terminal domain NuoC are required for the stability of not only the NuoCD protein, but also all other subunits [26]. These residues are also conserved in the HycE protein (Figs 1 and 2). Amino acid exchanges within the N-terminal domain of HycE were better tolerated than exchanges in these conserved residues; for example, the addition of a 10-His residue stretch after Gly-83 was shown to have no effect on the stability and activity of the FHL complex because they are exposed to the surface [27]. Moreover, engineering of the HycE protein revealed that truncation of the C terminus increased H₂ production from the FHL

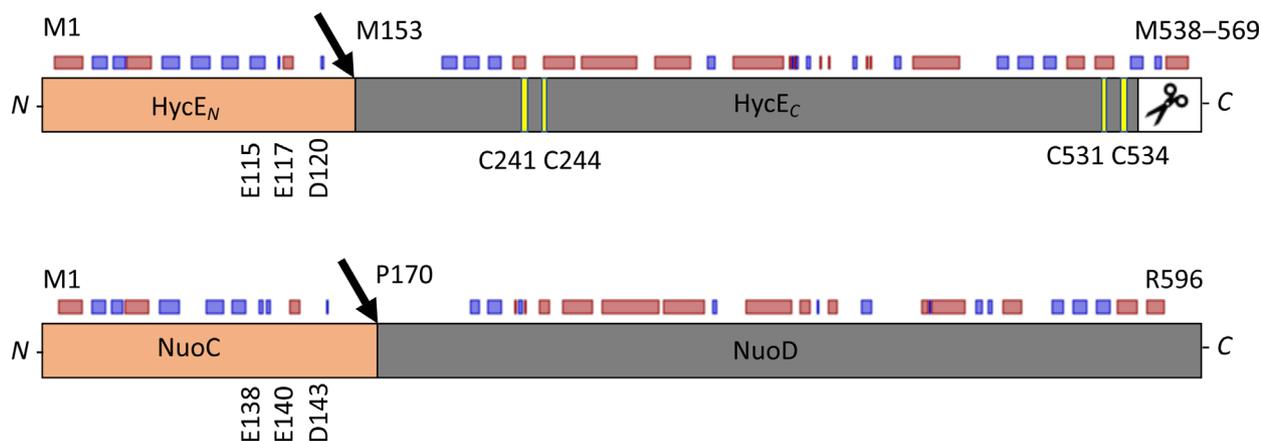


Fig. 1. Schematic representation of the HycE (top) and NuoCD (bottom) fusion proteins. The N-terminal domain of each protein is coloured in orange and the C-terminal domain in grey, and the black arrows indicate the sites where the proteins were genetically separated. The cysteinyl residues that coordinate the [NiFe]-cofactor in HycE are coloured in yellow, and the C-terminal extension that is proteolytically processed is shown in white with the scissors symbol. The locations of further relevant residues, which are mentioned in the text, are shown. The small rectangles above the respective proteins represent the secondary structure prediction (red – helix; blue – sheet) based on the Open PredictProtein server [51].

complex, despite the fact that these deletions remove the [NiFe]-binding site [28].

Therefore, in this study, we investigated the possibility of splitting the HycE and NuoCD domains into separate proteins in order to address the role of the N-terminal domain and the potential evolutionary significance of the fusion.

Methods

Growth conditions

Bacterial growth was generally carried out either aerobically on LB agar plates and in shaking liquid LB cultures or anaerobically as standing liquid cultures in TGYEP medium, pH 6.5, with 0.8% (w/v) glucose included [29] at 37 °C. Alternatively, for complex I activity, aerobic cultures were grown in glycerol medium [1% (w/v) peptone, 0.5% (w/v) yeast extract, 0.8% (w/v) NaCl, 0.4% (w/v) glycerol and separately prepared 15 mM Na₃PO₄, 10 mM Na₂HPO₄ and 10 mM KH₂PO₄, pH 8.5]. Cultures were vigorously shaken at 180 r.p.m. and 37 °C. Growth experiments were performed as biological triplicates in a 96-well microtiter plate, filled with 150 µL M9 minimal medium with 25 mM acetate as sole carbon source (47.6 mM Na₂HPO₄, 22 mM KH₂PO₄, 8.4 mM NaCl, 19 mM NH₄Cl, 2 mM MgSO₄, 0.1 mM CaCl₂, 0.3 µM thiamine dichloride, 0.1% w/v trace element solution) and agitation of 150 r.p.m. in a Tecan Infinite 200 plate reader (TECAN, Männedorf, Switzerland). Strains for maintaining the pMAK705 plasmids were grown aerobically at 30 °C. When appropriate, the antibiotics ampicillin and chloramphenicol were included at 100 and 34 µg·mL⁻¹, respectively.

DNA modifications and strain construction

All strains and plasmids are listed in Table 1 and oligonucleotides in Table 2. For constructing the N-terminal truncation of HycE in strain $\Delta hycEN$, the 514-bp DNA upstream region up to the encoded methionine start codon was amplified using the oligonucleotides HycEup_XbaI and HycEup_BamHI leaving the ribosome binding site intact, while the downstream 1254 bp was amplified from methionine codon 153 to the stop codon using the oligonucleotides HycEdown_BamHI and HycEdown_EcoRI. The two fragments were digested with XbaI/BamHI and BamHI/EcoRI, respectively, and ligated into pBSK(+) before being moved by XbaI/EcoRI restriction digestion to pMAK705 for chromosomal integration. The chromosomal integration into strain MG1655 was performed according to Ref. [30] and yielded the strain $\Delta hycEN$. Similarly, the above pMAK705 construct was modified by introducing the *nuoC* DNA sequence into the BamHI site after amplification with the oligonucleotides *nuoC*_HA_FW_BamHI and *nuoC*_RW_BamHI. The directionality of the insert was verified by DNA sequencing, and the integration onto the chromosome of MG1655 was done independently and resulted in strain *nuoC-hycE*.

Separation of the *hycE* gene into two distinct genes on plasmid pHycE was done by the NEBaseChanger method (NEB, Ipswich, MA, USA) using the oligonucleotides HycEN+C_FW and HycEN+C_RW and resulted in plasmid pHycEN+C. This construct was subcloned onto pMAK705 and moved to the chromosome as described [30].

In trans complementation was done by cloning the DNA fragment encoding the N-terminal domain including the connecting alpha helix with AA1-168 in pJET1.2 (Thermo

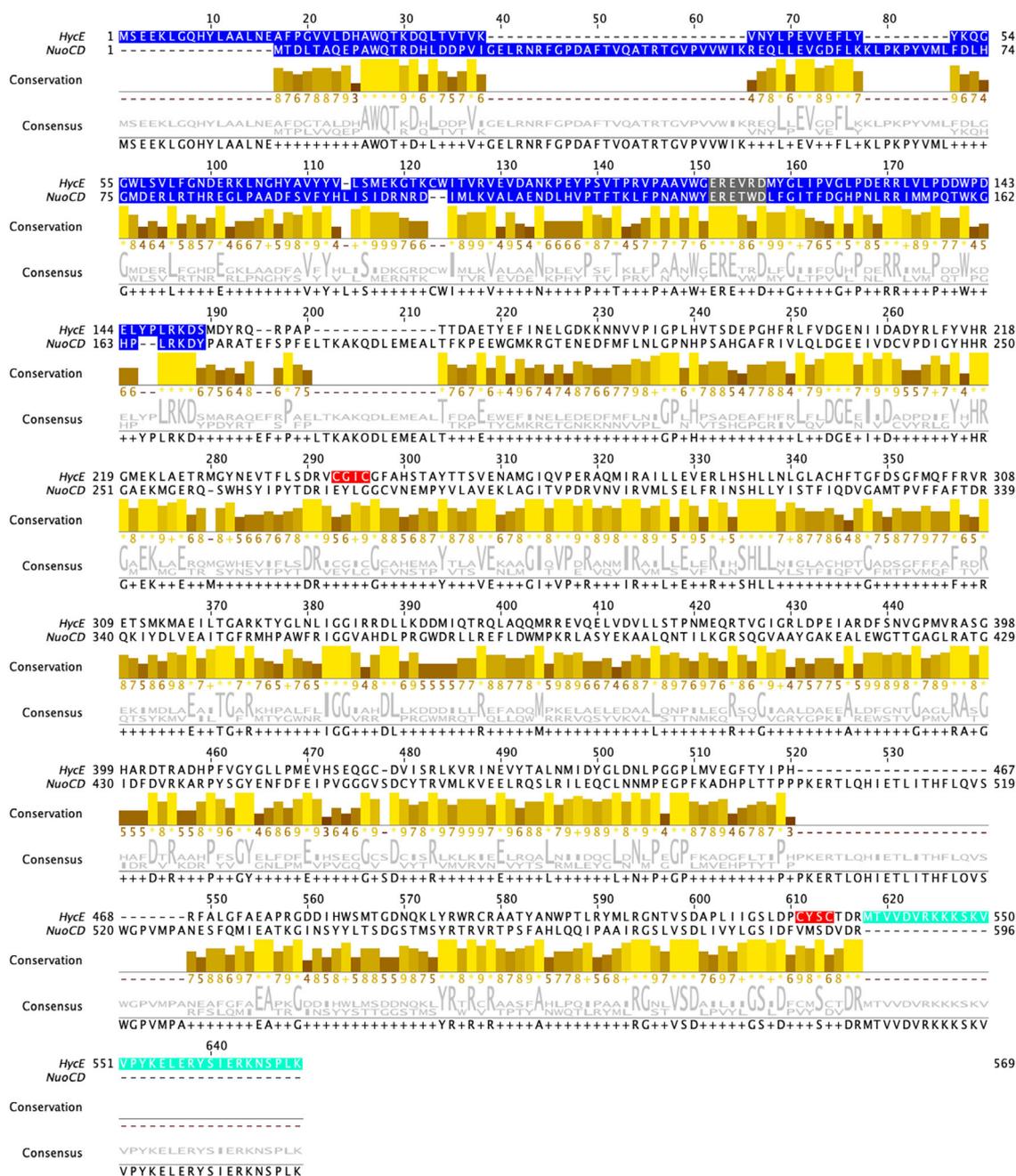


Fig. 2. Alignment of HycE and NuoCD protein sequences. The alignment was performed with CLUSTAL o (1.2.4) [55] within the UniProt database and the visualization done by JALVIEW [52]; in the conservation panel, the height of the bar represents the conservation as well as the consensus logo provided according to the size of the letter. Highlighted within the sequences are the N-terminal domains (blue highlight), conserved residues within the N-terminal domain (grey highlight), hydrogenase specific C-terminal extension (cyan highlight) and [NiFe]-cofactor binding motif (red highlight).

Table 1. Strains and plasmids used in this study.

	Genotype	Reference
Strain		
MG1655	F ⁻ λ ⁻ <i>ilvG</i> ⁻ <i>rfb-50</i> <i>rph-1</i>	[53]
BW25113	<i>lacI</i> ⁺ <i>rrnB</i> _{T14} Δ <i>lacZ</i> _{WJ16} <i>hsdR514</i> Δ <i>araBAD</i> _{AH33} Δ <i>rhaBAD</i> _{LD78} <i>rph-1</i> Δ(<i>araB-D</i>)567 Δ(<i>rhaD-B</i>)568 Δ <i>lacZ4787</i> :: <i>rrnB-3</i> <i>hsdR514</i> <i>rph-1</i>	[33,54]
JW2691	Like BW25113, but Δ <i>hycE</i>	[54]
<i>hycEN</i> +C	Like MG1655, but <i>hycE</i> separated as AA1-152 and AA153-569	This study
Δ <i>hycEN</i>	Like MG1655, but Δ <i>hycE</i> N-terminal domain corresponding to AA1-152	This study
<i>nuoC-hycE</i>	Like MG1655, but <i>hycE</i> N-terminal domain AA1-152 replaced with <i>nuoC</i> coding for AA1-169	This study
JW4040	Like BW25113, but Δ <i>fdhF</i>	[54]
Δ <i>ndh</i>	JW1095: like BW25113, but Δ <i>ndh</i>	[31,54]
Δ <i>nuoC</i>	BW25113 Δ <i>ndh</i> , Δ <i>nuoC</i> AA1-169	This study
<i>nuoC</i> +D	BW25113 Δ <i>ndh</i> , <i>nuoC</i> separated as AA1-169 and <i>nuoD</i> as AA170-596	This study
Δ <i>nuoC</i> -L	BW25113 Δ <i>ndh</i> , Δ <i>nuoC</i> -L	This study
Plasmids		
pMAK705	Temperature-sensitive replicon, Cm ^R	[30]
pHycE	pACYCDuet-1, <i>hycE</i> ⁺ , internal His-tag, Cm ^R	[34]
pHycEN+C	pACYCDuet-1, <i>hycE</i> N-terminal domain AA1-152, <i>hycE</i> C-terminal domain AA153-569 Cm ^R	This study
pHycEN	pJET1.2, <i>hycE</i> N-terminal domain AA1-168, Amp ^R	This study
pJET- <i>nuoCD</i>	pJET1.2, <i>nuoCD</i> including RBS, Amp ^R	This study
pKD3	Cm ^R	[33]
pCP20	Contains flippase gene for λ Red mutagenesis; Cm ^R , Amp ^R	[32]

Table 2. Oligonucleotides used in this study.

Oligonucleotide	Sequence 5' → 3'
HycEup_XbaI	GCGTCTAGATGCTGCTTGGTCTGTGGGTT
HycEup_BamHI	GCGGGATCCACTCTCTTTAATCACGCCGC
HycEdown_BamHI	GCGGGATCCATGGATTATCGTCAGCGTCC
HycEdown_EcoRI	GCGGAATTCCTTATTTACGCGCGAGTTTTTAC
<i>nuoC</i> _HA_FW_BamHI	GCGGGATCCATGTATCCGTATGATGTGCCGGATTATGCGACCGACTTAACCGCGCAAG
<i>nuoC</i> _RW_BamHI	GCGGGATCCATAATCTTTACGCAGCGGGTG
HycEN+C_FW	ATATACCATGGATTATCGTCAGCGTCC
HycEN+C_RW	CTCCTTAGCTGTCTTTACGCAGCGG
HycEN_RW_EcoRI	GCGGAATTCCTTAGTAGGTTTCAGCATCGGTGG
DnuoC_5'_FW	CGCCGACAGTCACCACGGACCATTTGCAATGGTGAACAAT ccatgggtccatgatgaatc
DnuoC_3'_RW_KpnI	TGGTCAGCTCAAACGGCGAGAATTCGGTAGCGCGCCGGCATggtaccctcct gcgattgtgtaggctggagc
DnuoLM_3'_RW	ACATGTTCGATGGCAAGGAACACGCCGATAACGCCGCCAG gcgattgtgtaggctggagc
<i>nuoC</i> _up_control	CGCATTGCCGTAAC TAACC
<i>nuoCD</i> _FW_BamHI	ggatccCACGGACCAATTTGCAATGGTG
<i>nuoCD</i> _RW_HindIII	aagcttTTAGCGGTCCACATCTGACATAAC
<i>hycB</i> _FW	GCGAAGCTTGATGAATCGTTTTGTAATTGCTG
<i>hycB</i> _RW	GCGGGATCCTCATTTAGCCTCTCCACTTT
<i>hycF</i> _FW	GCGCTGCAGGATGTTTACCTTTATCAAAAAAG
<i>hycF</i> _RW	GCGGGATCCTCAGATGGCCTCTTTCATATG
<i>tatB</i> _FW	GTGTTTGATATCGGTTTTAG
<i>tatB</i> _RW	TTACGGTTTTATCACTCGACG

Scientific, Waltham, MA, USA) using the primers HycEup_XbaI and HycEN_RW_EcoRI.

Strain BW25113 Δ*ndh* has been described before [31]. This strain was used for further modification of complex I. Deletion of the DNA sequence encoding the NuoC domain

was performed by initially replacing the sequence encoding the domain with the *cat* gene from vector pKD3 (oligonucleotides DnuoC_5'_FW and DnuoC_3'_RW_KpnI) and subsequently removing the resistance cassette using vector pCP20 as described [32,33]. The DNA sequence encoding

the downstream NuoD domain received a novel ribosome binding site and start codon. The deletion of the entire region from *nuoC-L* was performed similarly, but using the oligonucleotides DnuoC_5'_FW and DnuoLM_3'_RW [33]. The *nuoC+D* strain was constructed by ordering the corresponding region as gBlocks fragment (DNA sequence in Fig. 3), which contained a stop codon after the codon encoding Y169 in *nuoC*, a ribosome binding site and a start codon for *nuoD* (IDT DNA, Coralville, IA, USA), cloning it as BamHI/BamHI fragment into pMAK705 and introducing it into BW25113 Δndh .

The pJET-*nuoCD* plasmid was constructed by amplification of the chromosomal *nuoCD* gene, including its own

ribosome binding site, using the oligonucleotides *nuoCD_FW_BamHI* and *nuoCD_RW_HindIII* and Q5 DNA polymerase (NEB) and cloning it blunt-end into pJET1.2 vector according to the manufacturer's instructions (Thermo Scientific). Expression of the gene is under the control of the *lacUV5* promoter of the vector.

All authenticity of all DNA sequences in the plasmid inserts and strains was verified by sequencing.

RT-PCR

The RT-PCR was performed according to previously published methods [34]. RNA was isolated from strains in the

```

AGGAACCTGCTTTACCAAAATGGCACC GGTTATTCAGCGTCTGTATGACCAGATGCTGGAACCAAAATGGGTTATCTCAA
TGGGTGCTGTGCCAACTCTGGTGGTATGTACGATATTTATTCCGTTGTGCAGGGCGTCGATAAATTCATCCCGTTGAT
GTGTATATCCCGGGCTGCCCGCCGCGTCTTGAAGCGTACATGCAGGCACTGATGCTGTTGCAGGAATCTATCGGC AAAAGA
ACGTCGTCCGCTCTCTGGGTGGTTGGCGATCAGGGCGTTTATCGCGCCAATATGCAATCAGAGCGCGAACGCAAGCGCG
GTGAACGCATTGCCGTAACCTGCGTACACCTGACGAGATTTAATTTGCGCCTGTCCGCAAAGGGATTTTCTTCGCG
TTATTCCCTAAATCTATTTTCGCGAAGCTTACTGCGCCGACAGTCACCACGGACCATTTGCAATGGTGAACAATATGATATCC
GTATGATGTGCCGGATTATGCGgttaccACCGACTTAACCGCGAAGAACC CGCTGGCAGACCCGCGATCATCTTGATG
ATCCGGTGATTGGCGAACTGCGCAACCGTTTTGGGCCGATGCCTTTACTGTT CAGGCGACTCGCACCCGGGTTCCCGTT
GTGTGGATCAAGCGTGAACAATTACTGGAAGTTGGCGATTTCTTAAAGAAACTGCCGAAACCTTACGTCATGCTGTTTGA
CTTACACGGCATGGACGAACGCTCTGCGCACACCCGCGAAGGGTTACCTGCCGCGGATTTTTCCGTTTTCTACCATCTGA
TTTCTATCGATCGTAACCGCGACATCATGCTGAAGGTGGCGCTGGCAGAAAACGACCTGCACGTACCGACCTTCACAAA
CTGTTCCCGAACGCTAACTGGTATGAGCGTGAACCTGGGATCTGTTTGGCATTACTTTTCGACGGTCACCCGAACCTGCG
ACGCATCATGATGCCGCAAACCTGGAAAGGTCACCCGCTGCGTAAAGATTATtaaggagggtaccATGCCGCGCGCGCT
ACCGAATTCTCGCCGTTTGTAGCTGACCAAAGCCAAAACAGGATCTGGAGATGGAAGCCCTGACCTTCAAACCGGAAGAGTG
GGGGATGAAGCGCGGCACCGAAAACGAGGACTTCATGTTCTCAACCTCGGTCCGAACCACCCGTCGGCGCACGGGGCTT
TCCGTATCGTTTTTGCAACTCGATGGCGAAGAGATTGTCGACTGCGTACCAGACATCGGTTACCACCACCGTGGTGGCGAG
AAAATGGGCGAACGCCAGTCTTGGCACAGCTACATCCGTATACTGACCGTATCGAATACCTCGGCGGCTGCGTTAACGA
AATGCCTTACGTGCTGGCGGTAGAGAACTGGCCGGGATCACCGTGCCGGATCGCGTTAACGTCATTCGCGTTATGCTCT
CCGAACTGTTCCGCATCAACAGTCACCTGCTGTATATCTCGACCTTTATTCAGGACGTCG

```

Legend:

XX KpnI site
XX HA-epitope tag
ATG start/stop codons
XX *nuoC*

Fig. 3. DNA sequence of *nuoC+D* ordered as gBlocks fragment. The 1500-bp fragment above was obtained from IDT DNA and designed to contain the *nuoC* upstream region, *nuoC* as independent gene (start and stop codons are shown in red; gene is underlined), DNA sequence coding for an included HA-epitope tag (yellow highlight) and a ribosome binding site for the downstream *nuoD* gene linked to the start codon by a KpnI restriction site (italics).

exponential growth phase using the SV Total RNA Isolation System according to the manufacturer's instructions (Promega, Fitchburg, WI, USA). An extra DNase I digestion step was performed using RQ1 RNase-Free DNase (Promega). The RT-PCR to generate cDNA was performed using M-MLV Reverse Transcriptase, RNase H Minus and random hexamer primers according to the manufacturer's instructions (Promega). The presence of transcripts was tested with gene-specific oligonucleotides, which are listed in Table 2, by using 17 PCR cycles to remain within the semi-quantitative part of the amplification reaction.

Western blotting and signal quantification

Aliquots of 50 µg total proteins of anaerobically grown cells were separated on a 10% (w/v acrylamide) SDS/PAGE, transferred to a nitrocellulose membrane and challenged with antibodies raised against HycG (1 : 3000, a kind gift from A. Böck/R. G. Sawers) [35,36]. The secondary antibody used was conjugated to HRP enzyme (Bio-Rad, Hercules, CA, USA), and the light signal generated by the ECL reaction (Thermo Scientific) was recorded with a film. The signal intensities were analysed using IMAGEJ software [37].

Enzymatic assays

Hydrogenase activity was determined as H₂-dependent benzyl viologen (BV) reduction using 50 mM MOPS buffer, pH 7.0, with 4 mM BV in a stoppered 1.6-mL cuvette with 0.8 mL H₂ gas headspace as described [38]. For kinetic determination of FHL activity, cell suspensions were applied to a modified Clark-type electrode (Oxytherm, Hansatech Instruments, Norfolk, UK), which was equipped with an OXY/EUCU module to reverse the polarizing voltage to -0.7 V, where the electrode disc is only sensitive to H₂ (Hansatech Instruments). A volume of 2 mL degassed MOPS buffer, pH 7.0, was used, and the reaction was started with 15 mM formate. The H₂ content of the gas headspace of a 15-mL Hungate tube filled with 7 mL culture was determined by gas chromatography using a GC2010 Plus Gas Chromatograph (Shimadzu, Kyōto, Japan) as described [39].

To measure complex I activity, the NADH-dependent oxidase activity was determined on a Clark-type electrode sensitive to O₂ (Hansatech Instruments) and was performed essentially identical to the H₂ measurements, with the exception that 1.25 mM NADH was used as substrate [31]. Similarly, the succinate oxidase activity was monitored in the presence of 10 mM succinate. Protein concentrations were determined by the method of Ref. [40].

Sucrose gradient centrifugation

The sucrose gradients were prepared and NADH/ferricyanide oxidoreductase activity was determined basically as described

[31]. Briefly, membrane extracts from cells grown under the appropriate growth conditions were prepared after cell rupture with an Emulsiflex (EF-C5; Avestin Europe, Mannheim, Germany) in the presence of PMSF (0.1 mM) and DNase at 160 000 g. The sediment was resuspended in buffer A (50 mM MES/NaOH, 50 mM NaCl) with 1% (w/v) DDM and 5 mM MgCl₂. After solubilization, the membrane fraction was briefly centrifuged and subsequently applied to a freshly poured 24 mL linear sucrose gradient ranging from 5% to 30% (w/v) sucrose in buffer A supplemented with 5 mM MgCl₂. After centrifugation for 16 h at 140 000 g, the gradient was fractionated into 24 × 1 mL fractions. The NADH/ferricyanide oxidoreductase activity of 50 µL from each fraction was assayed in 1 mL buffer A containing 1 mM ferricyanide (AppliChem, Darmstadt, Germany). The reaction was started by an addition of 0.2 mM NADH, and the reaction was followed by monitoring the slope at 410 nm in an Ultrospec photometer (Amersham Pharmacia Biotech, Amersham, UK). The rate was calculated using an ε₄₁₀ of 1 mM⁻¹·cm⁻¹. Protein concentration was determined by biuret assay [41].

Results and Discussion

Formate hydrogenlyase complex retains activity after separation of the N- and C-terminal domains of HycE

In order to investigate the effect of dividing the HycE protein into two separate domains on the activity and stability of the FHL complex, the efficiency of conversion of formate to H₂ and CO₂ was quantified by measuring the headspace H₂ content of the culture after fermentative growth. This activity represents the total yield of H₂ after growth on glucose medium, but also provides an accurate indication of the activity of the FHL complex, and is particularly useful for assessing low FHL activities. The parental strain produced 7.03 µmol H₂·mL culture⁻¹·OD_{600 nm}⁻¹, while H₂ production was basically absent in the negative control strain JW2691, which lacks the *hycE* gene encoding the catalytic subunit HycE. This strain can be partially (71%) complemented for H₂ production by introducing plasmid pHycE carrying the *hycE* gene. We have previously observed that *hycE* cannot fully complement a Δ*hycE* strain when provided *in trans*, and we assume this to be due to the altered chromosomal context [42]. Surprisingly, the Δ*hycE* strain can also be complemented by the plasmid pHycEN+C, which consists of the same insert sequence as on pHycE with the exception that it includes a stop codon within *hycE* (Table 3). This 'split', or separated, *hycE* gene in the chromosomal context in strain *hycEN+C* also resulted

Table 3. Hydrogenase activities and their percentage of the parental activities. The activities were determined from three biological replicates.

Strain (+Plasmid)	H ₂ -headspace quantification (μmol·OD _{600 nm} ⁻¹ ·mL ⁻¹)		Formate-dependent evolution (nmol H ₂ ·mg ⁻¹ ·min ⁻¹)	H ₂	Total hydrogenase activity H ₂ :BV (U·mg ⁻¹)	
MG1655 (parental)	7.03 ± 0.28	100%	636 ± 28	100%	1.12 ± 0.37	100%
JW2691 ($\Delta hycE$)	0.04 ± 0.07	< 1%	1 ± 2	< 1%	0.32 ± 0.10	29%
JW2691 + pHycE	4.99 ± 0.71	71%	190 ± 62	30%	0.56 ± 0.32	50%
JW2691 + pHycEN+C	7.46 ± 1.94	106%	337 ± 173	53%	0.64 ± 0.53	57%
<i>hycEN</i> +C	7.23 ± 0.39	103%	350 ± 19	55%	0.67 ± 0.34	60%
$\Delta hycEN$	< 0.01	< 1%	1 ± 2	< 1%	0.19 ± 0.05	17%
$\Delta hycEN$ + pHycEN	3.23 ± 0.19	46%	73 ± 15	11%	0.39 ± 0.18	35%
<i>nuc-hycE</i>	< 0.01	< 1%	1 ± 2	< 1%	0.19 ± 0.01	17%

in H₂ accumulation to a level similar to that of the parental strain. Strain $\Delta hycEN$, which carried a truncation of the N-terminal domain of HycE, but retained a protein fragment of the C-terminal domain, was unable to evolve H₂, a phenotype like that of the *hycE* deletion strain. The $\Delta hycEN$ strain could be complemented *in trans*, however, by introducing a plasmid encoding the missing N-terminal portion of HycE, which restored almost 50% of the H₂ production (Table 3). Complementation could not be achieved by using a fusion of the related NuoC to the C-terminal domain of HycE. Hence, despite sharing structural similarity, they fulfil distinct roles in both complexes. These assays also showed that the HycE protein, when physically separated into N- and C-terminal domains, was as efficient as the native HycE protein in converting formate to H₂ and CO₂ during mixed-acid fermentation.

In addition to the headspace analysis, which measures accumulation of H₂ over 16 h, cell suspensions were assayed on a modified Clark-type electrode with formate as substrate allowing quantitative measurement of the specific activity of the FHL complex. These measurements verified the absence of activity in strains JW2691 ($\Delta hycE$), $\Delta hycEN$ and *nuc-hycE*. As could be shown for the H₂-headspace measurement, the pHycE and pHycEN plasmids were functional, but were only able to complement the corresponding deletion strains JW2691 ($\Delta hycE$) and $\Delta hycEN$, to 30% and 11% of the level of the wild-type strain, respectively (Table 3). While determination of the accumulated H₂ revealed 100% complementation of H₂ production, the direct FHL assay led to approximately 50% of the FHL activity rate of the parental strain when the domain-split HycE protein was introduced either *in trans* on plasmid pHycEN+C in strain JW2691 ($\Delta hycE$) or after chromosomal integration in strain *hycEN*+C (Table 2). This result revealed that the two physically separated domains of HycE retain

half of the activity of the fused, native protein, despite being fully functional in conversion of glucose during growth. Moreover, the findings demonstrate that providing each component in multiple copies did not result in higher activity, indicating that the amount of the two domains in the complex did not limit the reaction rate.

The samples used in the electrode experiments were subsequently used to generate crude extracts, allowing spectrophotometric quantification of total hydrogenase activity, which is a composite of Hyd-1, Hyd-2 and the Hyd-3 component of the FHL complex. Hyd-1 and Hyd-2 are both H₂-oxidizing hydrogenases, which allow the cell to scavenge some of the FHL-produced H₂. Under these conditions, Hyd-3 is generally the major contributor to the total hydrogenase activity [42]. A fourth hydrogenase, predicted to form a second FHL complex, is usually not expressed under the conditions used in this study and does not contribute to the activity [43,44]. Here, the residual total hydrogenase activity in the $\Delta hycE$ strain JW2691, representing the combined activities of Hyd-1 and Hyd-2, was approximately 30% of that of the wild-type strain MG1655 (Table 3). Strains $\Delta hycEN$ and *nuc-hycE* showed levels of total hydrogenase activity of 17% of the parental, which, surprisingly, is lower than when only Hyd-1 and Hyd-2 are present. Using this assay, the HycE protein separated into two domains complemented to a level of 60% of the wild-type activity, which was slightly higher than the activity recovered by complementation with a plasmid, pHycE13, encoding the native HycE polypeptide. The formate dehydrogenase H activity, which is the other catalytic component of the FHL complex, followed these trends, but with slightly greater variation (Table 4). This indicates that the formate dehydrogenase polypeptide was able to stably interact with the complex carrying the separated domains of HycE.

Table 4. Formate dehydrogenase H activity measured in crude extracts and the percentage of the parental activity. The activities were determined from three biological replicates.

Strain/plasmid	FDH-H formate: BV (U mg ⁻¹)	% of parental activity
MG1655	2.61 ± 2.71	100
JW2691 ($\Delta hycE$)	0.15 ± 0.24	6
JW2691 + pHycE13	0.03 ± 0.03	1
JW2691 + pHycEN+C	0.43 ± 0.73	16
<i>hycEN</i> +C	0.46 ± 0.51	18
$\Delta hycEN$	0.14 ± 0.72	5
$\Delta hycEN$ + pHycEN	0.34 ± 0.54	13
<i>nuoC-hycE</i>	0.14 ± 0.24	5

Modifications of the HycE N-terminal domain do not influence transcription of the *hyc* operon

Bacterial operons often contain internal promoters and regulatory elements [45]. In the case of the *hyc* operon, computational evidence predicts a further sigma 70-dependent promoter in front of *hycI*, the last gene of the operon [46]. Therefore, it was investigated whether the genomic modifications in strains $\Delta hycEN$ and *nuoC-hycE*, which completely lacked FHL activity, influenced the transcription of the downstream genes. Reverse transcriptase PCR on RNA isolated from those strains in comparison with cDNA obtained from the parental strain MG1655 showed similarly intensive bands for the upstream *hycB* gene, as well as for the downstream *hycF* gene (Fig. 4). Despite this analysis not being quantitative, the overall levels of transcripts

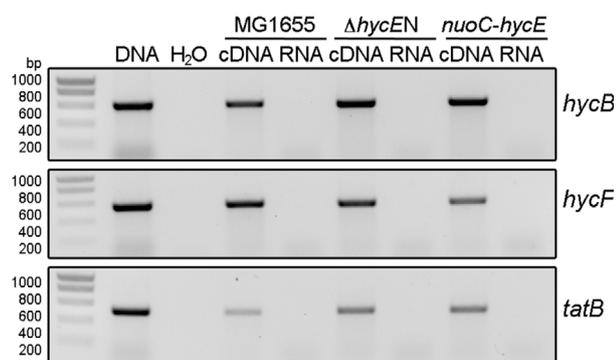


Fig. 4. Transcriptional analysis of the *hyc* operon after genomic modifications. RNA was isolated from strains MG1655, $\Delta hycEN$ and *nuoC-hycE*, and reverse transcription was performed and subsequently PCR on the cDNA and, for control of DNA contamination, also from the RNA and water. The transcripts that were analysed are *hycB*, *hycF* and *tatB* as labelled on the right side. The first lane shows the size of the correct PCR fragment when obtained from genomic DNA, and the size of the ladder (SmartLadder; Eurogentec, Seraing, Belgium) is shown on the very left.

were similar, as shown for the transcripts of the constitutive mRNA of the *tatB* gene (Fig. 4). It is concluded, therefore, that the introduced genetic modifications do not alter the transcription of up- and downstream genes within the operon.

Presence of the small subunit protein HycG correlates with the FHL activities

Analysis of the hydrogenase small subunit HycG of the FHL complex by quantitative western blotting showed that the protein amount is subject to change [8]. The absence of HycE in the $\Delta hycE$ strain JW2691 caused an approximate 70% reduction in HycG levels compared to the wild-type (Fig. 5). A similar observation for a $\Delta hycE$ strains was made previously [8]. Generally, for hydrogenase assembly to occur, the large subunit must receive its [NiFe]-cofactor and therefore interact with various delivery proteins and undergo endoproteolytic processing [19,47]. Only then is the protein primed for interaction with its small subunit, and this allows the subsequent initiation of the assembly of the entire FHL complex. However, if processing of the large subunit is delayed or prevented, the small subunit is rapidly degraded [47]. This reduced amount of HycG was also observed in the $\Delta hycEN$ strain, both with and without *in trans* complementation, and the *nuoC-hycE* fusion strain (Fig. 5). The strains *hycEN*+C and, to a lesser extent, the complementation of $\Delta hycE$ with pHycE and pHycEN+C show high amounts of HycG, which correlates with their FHL activities. This indicates that the absence of hydrogenase activity, which was observed for some of the HycE variants, correlates with loss of complex integrity and suggests that the N-terminal domain could be the driving force for assembly of the core FeS-carrying proteins.

It is notable in this context that the $\Delta fdhF$ strain JW4040, despite lacking FHL activity due to the absence of the electron-input subunit FdhH, showed high amounts of HycG (Fig. 5). As FdhH is only loosely associated with the FHL complex, this result indicates that this subunit is not required for the stable assembly of the core FHL complex [27] and that the core FHL complex remains stable even without electron flow through the complex.

Removal or separation of NuoC causes complete loss of complex I activity

Based on the findings made with HycE, it was decided to test whether the homologous protein of complex I, NuoCD, also revealed similar properties when the two-domain fusion protein was separated into two

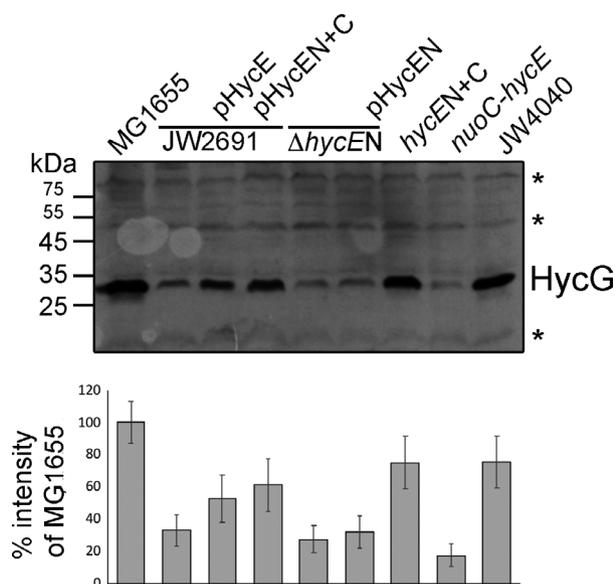


Fig. 5. Analysis of HycG levels in HycE variant strains. The amount of HycG protein from anaerobically grown samples was analysed by western blotting using antibodies raised against HycG. A protein amount of 50 μ g of strains MG1655, JW2691 ($\Delta hycE$), $\Delta hycEN$, *hycEN+C*, *nuoC-hycE* and JW4040 ($\Delta fdhF$) was separated by SDS/PAGE either directly or after complementation with the plasmids as indicated. Unspecific and nonhydrogenase related cross-reactions of the antibody served as an internal loading control and are indicated by an asterisk on the right side of the panel. The histograms in the lower panel show the quantification of the HycG band intensity (\pm SD) in comparison with the HycG protein amount in MG1655 using IMAGEJ.

distinct proteins. Strain *nuoC+D*, in which the gene encoding NuoCD was physically separated into two genes, was generated. Additionally, the DNA encoding the NuoC domain was deleted in strain $\Delta nuoC$ to examine the consequences on complex I assembly and activity. All changes were introduced into a strain lacking the *ndh* gene, which codes for the alternative NADH dehydrogenase, and whose presence would interfere with the NADH oxidase activity assay. As a control, a double-deletion strain $\Delta nuoC-L$ was constructed, which, in addition to the Δndh deletion, was also devoid of most of the genes of the *nuo* operon, including *nuoEFG* encoding the diaphorase domain and part of the membrane domain (*nuoHIJKL*).

The functionality of complex I was initially tested by monitoring the NADH oxidase activity, which is catalysed by complex I and the cytochrome *bo₃* and cytochrome *bd* oxidases. The strain Δndh , lacking the alternative NADH dehydrogenase, showed an activity of 0.2 U·mg⁻¹ (Table 5), which also has been observed in a previous study [31]. The negative control strain lacking complex I (strain $\Delta nuoC-L$) had only a residual

Table 5. NADH and succinate oxidase activity of various *Escherichia coli* complex I deletion strains. The activities were determined from at least two biological replicates.

Strain	NADH oxidase activity (mU·mg protein ⁻¹)		Succinate oxidase activity (mU·mg protein ⁻¹)
	activity (mU·mg protein ⁻¹)	% Activity of parental	activity (mU·mg protein ⁻¹)
Δndh	213 \pm 12	100	63 \pm 14
Δndh , $\Delta nuoC$	3 \pm 1	1	50 \pm 7
Δndh , <i>nuoC+D</i>	8 \pm 2	4	40 \pm 15
Δndh , $\Delta nuoC-L$	9 \pm 8	4	48 \pm 7

oxidase activity of 4%. The two mutant strains $\Delta nuoC$ and *nuoC+D* had almost no NADH oxidase activity, which indicated a loss of complex I activity (Table 5). A previous study established that deletion of any of the *nuo*-genes encoding complex I abolished its ability to transfer electrons into the respiratory chain [48]. In order to rule out a general effect of the mutations on the respiratory chain, succinate oxidase activity was assessed, which uses succinate dehydrogenase, complex II, instead of complex I to channel electrons into the quinone pool. Succinate oxidase activity was very similar in these strains, ranging from 40 to around 60 mU per mg protein (Table 5), demonstrating that the lack of NADH oxidase activity is specifically due to the absence of complex I.

The detergent-solubilized membrane fractions were subjected to a sucrose gradient centrifugation in order to determine the stability and the assembly of the complex in the mutant strains. The fractions were separately assayed for NADH/ferricyanide oxidoreductase activity, which indicates the presence of the FMN-containing subunit NuoF, which is part of the diaphorase domain of complex I. The activity peaks in fraction 14, corresponding to intact complex I as observed in the Δndh strain (shaded area in Fig. 6). Complex I mutant strain $\Delta nuoC-L$, which lacks the diaphorase proteins NuoEFG, did not show activity in these fractions. Fractions 5–9 contained a broader, more disperse activity peak, which was originally thought to include the diaphorase activity that had dissociated from the complex. However, because it is also present in the membrane preparations of the $\Delta nuoC-L$ strain, this seems unlikely and rather represents an unspecific reaction of the ferricyanide. Furthermore, the activity profile of the extract from strains $\Delta nuoC$ and *nuoC+D* resembles the negative control, indicating that NuoF is absent.

Assuming that the mutations could destabilize complex I after extraction from the membrane, but nevertheless allowed retention of a residual activity in whole cells, the strains were subjected to growth experiments

Fig. 6. Stability of complex I in various *Escherichia coli* deletion mutants. NADH/ferricyanide oxidoreductase activity of sucrose gradient fractions loaded with detergent extracts. The gradients were normalized to a total load of 20 mg protein. The shaded area represents the distribution of holo-complex I (scale right). Strains are Δndh – grey circles with grey connecting line; $\Delta nuoC$ – red triangles with black border; $nuoC+D$ – black triangles; and $\Delta nuoC-L$ – grey squares.

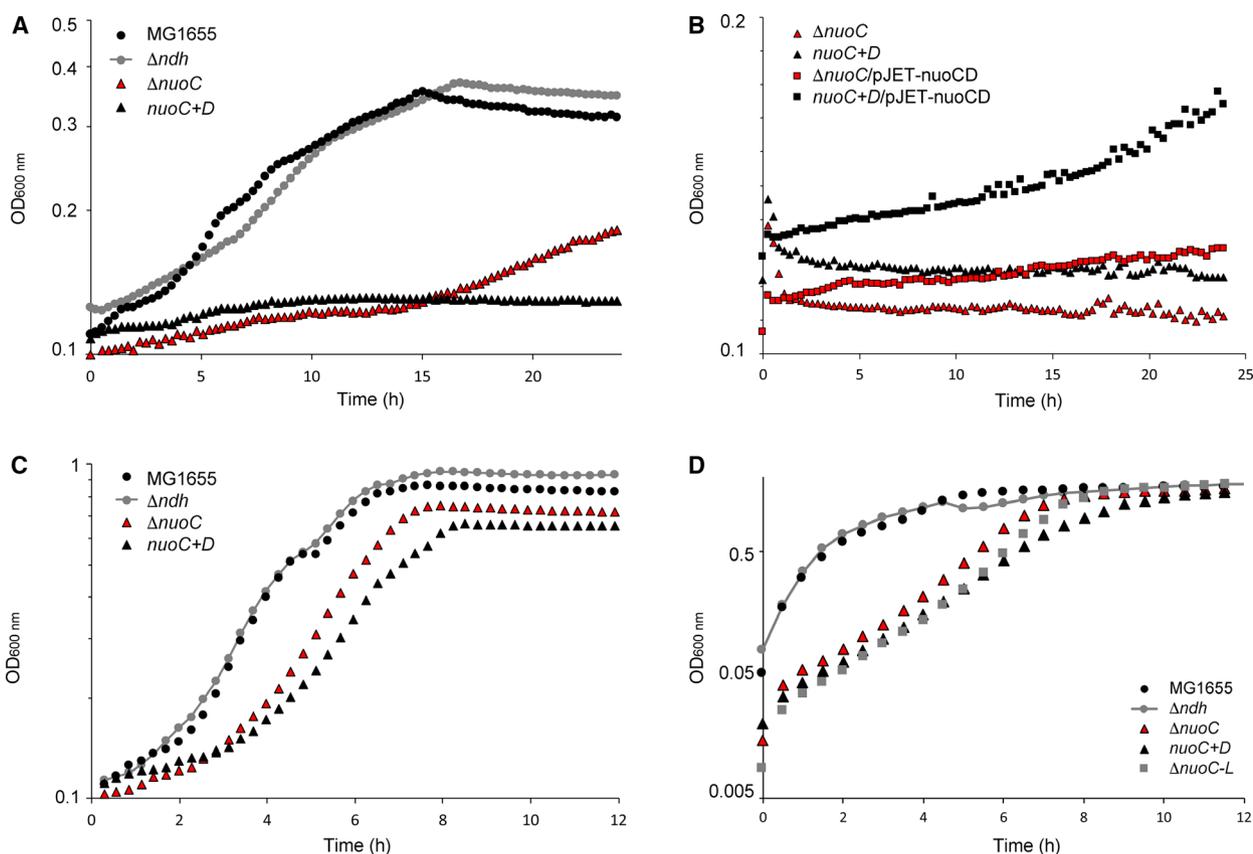
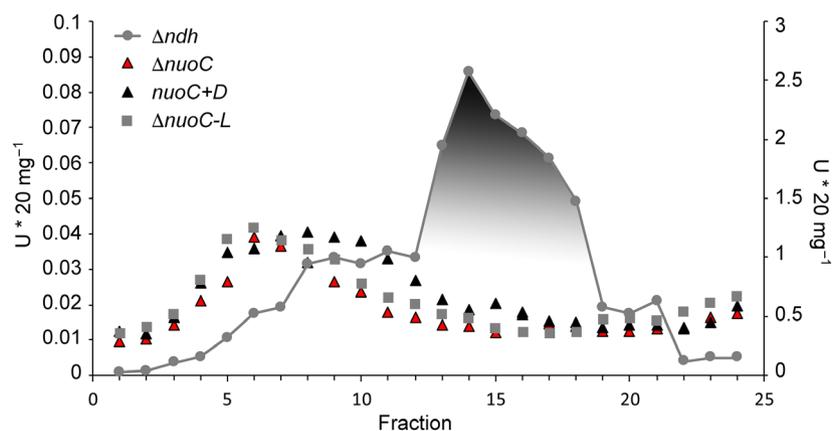


Fig. 7. Aerobic growth phenotype of complex I mutants in (A) in acetate minimal medium (B) after complementation (C) glucose minimal medium and (D) glycerol/LB medium. All growth curves were recorded in a plate reader with 150 r.p.m. shaking in 150 μ L volume. The graphs represent the average of three biological triplicates. For all graphs: BW25113 – black circles; Δndh – grey circles with grey connecting line; $\Delta nuoC$ – red triangles with black border; $nuoC+D$ – black triangles; and $\Delta nuoC-L$ – grey squares. For panel 7B only: $\Delta nuoC/pJET-nuoCD$ – red squares with black border; and $nuoC+D/pJET-nuoCD$ – black squares.

in an *in vivo* assay (Fig. 7A). The cells were grown aerobically with acetate as sole nonfermentable carbon source, which requires activity of complex I in order to maintain the NADH/NAD⁺ ratio within the cells. This in turn is necessary for the function of the TCA

cycle and the glyoxylate shunt to produce amino acids [49]. The growth monitoring during 24 h showed that the Δndh strain had no growth defect compared to the parental strain, as observed before [31]. In contrast, both the $\Delta nuoC$ strain and the $nuoC+D$ strain did not

grow within the first 16 h of the experiment, while eventually the $\Delta nuoC$ strain showed a slight increase in cell density (Fig. 7A). The absence of growth with acetate, as in the $\Delta nuoC$ and $nuoC+D$ strains, typically indicates lack of complex I activity. A complementation of the strains $\Delta nuoC$ and $nuoC+D$ with a plasmid encoding the $nuoCD$ gene, including its ribosome binding site, was attempted (Fig. 7B). However, the growth of the $\Delta nuoC$ or $nuoC+D$ strain on acetate-containing minimal medium could only be restored to a small degree with the $nuoCD$ plasmid, attaining optical densities of 0.18 ($nuoC+D$), while the parental strain grew to an OD of 0.33. This could indicate a dominant phenotype of the chromosomal $nuoCD$ modification; however, it also does not entirely exclude effects caused by disruption of the nuo operon. All strains were in parallel subjected to growth in glucose medium, where they exhibited no difference in growth behaviour (Fig. 7C). Furthermore, glycerol medium was also tested, where the strains still attained the same final optical density in the stationary phase (Fig. 7D). The Δndh strain and its BW25113 parental derivative grew very similarly, with a growth rate μ of 1.45 and 1.74 h^{-1} , respectively. In contrast, the strains $\Delta nuoC$, $nuoC+D$ and $\Delta nuoC-L$ had significantly, but similarly, diminished growth rates of 0.54, 0.46 and 0.52 h^{-1} , respectively.

Therefore, all of the growth experiments indicate that the $nuoC$ deletion and separation of $nuoCD$ into two genes abolish complex I activity in whole cells in *in vivo* experiments. The absence of activity in the $\Delta nuoC$ strain was not surprising, because even single amino acid exchanges in NuoC had a similar effect [26]. What was surprising, however, was the different effects of separating the paralogous genes into two parts in the case of FHL and complex I. The FHL complex tolerated the introduced changes, while complex I presumably did not assemble anymore. The region within the proteins where separation of the domains occurred was based on protein alignments and the presence of the Met153 residue in HycE, which was directly hijacked as initiation codon (see alignment in Fig. 2). However, the amino acid sequence in this region is not as conserved as other regions of the protein, and without further structural information of HycE and NuoCD, it is very difficult to predict the location of the loop region and its function in connecting the modules of the complexes.

Conclusions

The research conducted here on the function of the N-terminal domains of the HycE and NuoCD subunits

of the *E. coli* FHL complex and complex I, respectively, reveals that these domains have been evolutionarily conserved and appear to have an important role in complex assembly. It is still not understood what the regulatory steps in the assembly of these large complexes are, but it is reasonable to assume that assembly must involve a highly ordered sequence of events, where accumulated subcomplexes represent intermediate steps [12,50]. The FeS proteins tend to be unstable in the absence of other components, and thus, the nucleation of the core complex (HycB, E, F, G or NuoB, CD, H, I, L) probably starts with the catalytic subunit HycE and its paralogue NuoCD, respectively. These subunits reside next to the membrane interface, while the formate dehydrogenase or NuoEFG proteins represent the last complex-specific steps in the separate evolution of the two complexes and are located further away from the membrane [12]. We identified the N-terminal domains of the HycE and NuoCD proteins as key determinants in initiating the assembly of this core complex. The different effects of splitting the proteins in two parts show that despite the existing phylogenetic relationship between the FHL and complex I, the latter tolerates splitting of the $nuoCD$ gene less well.

Acknowledgements

The authors wish to thank Frank Sargent (Newcastle University, UK) for providing resources during the early stages of this study and R. Gary Sawers for providing language help and critical discussions. The authors acknowledge the financial support within the funding programme Open Access Publishing by the German Research Foundation (DFG) to CP and further grants to CP (Grant Number PI1252/2) and TF (Grant Number FR 1140/11-1) within the SPP1927.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

PS and SB investigated complex I activity; UL, CB and CP investigated FHL activity; TF and CP supervised the study, provided resources and validated the results; and CP wrote the original draft.

References

- 1 Heringa J and Taylor WR (1997) Three-dimensional domain duplication, swapping and stealing. *Curr Opin Struct Biol* 7, 416–421.

- 2 Rosano C, Zuccotti S, Stefani M, Bucciantini M, Ramponi G and Bolognesi M (2002) Crystallization and preliminary X-ray characterization of the acylphosphatase-like domain from the *Escherichia coli* hydrogenase maturation factor HypF. *Acta Crystallogr D Biol Crystallogr* **58**, 524–525.
- 3 Chiti F, Bucciantini M, Capanni C, Taddei N, Dobson CM and Stefani M (2001) Solution conditions can promote formation of either amyloid protofilaments or mature fibrils from the HypF N-terminal domain. *Protein Sci* **10**, 2541–2547.
- 4 Koo OK, Jeong D-W, Lee JM, Kim MJ, Lee J-H, Chang HC, Kim JH and Lee HJ (2005) Cloning and characterization of the bifunctional alcohol/acetaldehyde dehydrogenase gene (*adhE*) in *Leuconostoc mesenteroides* isolated from kimchi. *Biotechnol Lett* **27**, 505–510.
- 5 Sawers RG and Clark D (2004) Fermentative pyruvate and acetyl-coenzyme A metabolism. In *EcoSal – Escherichia coli and Salmonella: Cellular and Molecular Biology* (Stewart V ed). Chapter 3.5.4. American Society for Microbiology, Washington, DC. <https://doi.org/10.1128/ecosalplus.3.5>.
- 6 Marsh JA, Hernández H, Hall Z, Ahnert SE, Perica T, Robinson CV and Teichmann SA (2013) Protein complexes are under evolutionary selection to assemble via ordered pathways. *Cell* **153**, 461–470.
- 7 Pinske C and Sawers RG (2016) Anaerobic formate and hydrogen metabolism. In *EcoSal Plus2016* (Stewart V ed). American Society for Microbiology, Washington, DC, ESP-0011–2016.
- 8 Sauter M, Böhm R and Böck A (1992) Mutational analysis of the operon (*hyc*) determining hydrogenase 3 formation in *Escherichia coli*. *Mol Microbiol* **6**, 1523–1532.
- 9 Marreiros BC, Batista AP, Duarte AMS and Pereira MM (2012) A missing link between complex I and group 4 membrane-bound [NiFe] hydrogenases. *Biochim Biophys Acta* **1827**, 198–209.
- 10 Böhm R, Sauter M and Böck A (1990) Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogenlyase components. *Mol Microbiol* **4**, 231–243.
- 11 Albracht SP (1993) Intimate relationships of the large and the small subunits of all nickel hydrogenases with two nuclear-encoded subunits of mitochondrial NADH: ubiquinone oxidoreductase. *Biochim Biophys Acta* **1144**, 221–224.
- 12 Friedrich T and Scheide D (2000) The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Lett* **479**, 1–5.
- 13 Hirst J (2013) Mitochondrial complex I. *Annu Rev Biochem* **82**, 551–575.
- 14 Moparthy V and Hägerhäll C (2011) The evolution of respiratory chain complex I from a smaller last common ancestor consisting of 11 protein subunits. *J Mol Evol* **72**, 484–497.
- 15 Baradaran R, Berrisford JM, Minhas GS and Sazanov LA (2013) Crystal structure of the entire respiratory complex I. *Nature* **494**, 443–448.
- 16 Sazanov LA and Hinchliffe P (2006) Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*. *Science* **311**, 1430–1436.
- 17 Yu H, Wu C-H, Schut GJ, Haja DK, Zhao G, Peters JW, Adams MWW and Li H (2018) Structure of an ancient respiratory system. *Cell* **173**, 1636–1649.e16.
- 18 Forzi L and Sawers RG (2007) Maturation of [NiFe]-hydrogenases in *Escherichia coli*. *Biometals* **20**, 565–578.
- 19 Sawers RG and Pinske C (2017) NiFe-hydrogenase assembly. In *Encyclopedia of Inorganic and Bioinorganic Chemistry*. John Wiley & Sons, Ltd, Hoboken, NJ. <https://doi.org/10.1002/9781119951438.eibc2484>.
- 20 Drapal N and Böck A (1998) Interaction of the hydrogenase accessory protein HypC with HycE, the large subunit of *Escherichia coli* hydrogenase 3 during enzyme maturation. *Biochemistry* **37**, 2941–2948.
- 21 Chan Chung KC and Zamble DB (2011) Protein interactions and localization of the *Escherichia coli* accessory protein HypA during nickel insertion to [NiFe] hydrogenase. *J Biol Chem* **286**, 43081–43090.
- 22 Rossmann R, Sauter M, Lottspeich F and Böck A (1994) Maturation of the large subunit (HYCE) of *Escherichia coli* hydrogenase 3 requires nickel incorporation followed by C-terminal processing at Arg537. *Eur J Biochem* **220**, 377–384.
- 23 Kwon S, Watanabe S, Nishitani Y, Kawashima T, Kanai T, Atomi H and Miki K (2018) Crystal structures of a [NiFe] hydrogenase large subunit HyhL in an immature state in complex with a Ni chaperone HypA. *Proc Natl Acad Sci USA* **115**, 7045–7050.
- 24 Albareda M, Pacios LF and Palacios JM (2019) Computational analyses, molecular dynamics, and mutagenesis studies of unprocessed form of [NiFe] hydrogenase reveal the role of disorder for efficient enzyme maturation. *Biochim Biophys Acta Bioenerg* **1860**, 325–340.
- 25 Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztányi Z, El-Gebali S, Fraser M *et al.* (2017) InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res* **45**, D190–D199.
- 26 Castro-Guerrero N, Sinha PK, Torres-Bacete J, Matsuno-Yagi A and Yagi T (2010) Pivotal roles of three conserved carboxyl residues of the NuoC (30k) segment in the structural integrity of proton-translocating NADH-quinone oxidoreductase from *Escherichia coli*. *Biochemistry* **49**, 10072–10080.
- 27 McDowall JS, Murphy BJ, Haumann M, Palmer T, Armstrong FA and Sargent F (2014) Bacterial formate

- hydrogenlyase complex. *Proc Natl Acad Sci USA* **111**, E3948–E3956.
- 28 Maeda T, Sanchez-Torres V and Wood TK (2008) Protein engineering of hydrogenase 3 to enhance hydrogen production. *Appl Microbiol Biotechnol* **79**, 77–86.
- 29 Begg YA, Whyte JN and Haddock BA (1977) The identification of mutants of *Escherichia coli* deficient in formate dehydrogenase and nitrate reductase activities using dye indicator plates. *FEMS Microbiol Lett* **2**, 47–50.
- 30 Hamilton CM, Aldea M, Washburn BK, Babitzke P and Kushner SR (1989) New method for generating deletions and gene replacements in *Escherichia coli*. *J Bacteriol* **171**, 4617–4622.
- 31 Burschel S, Kreuzer Decovic D, Nuber F, Stiller M, Hofmann M, Zupok A, Siemiatkowska B, Gorka M, Leimkühler S and Friedrich T (2019) Iron-sulfur cluster carrier proteins involved in the assembly of *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I). *Mol Microbiol* **111**, 31–45.
- 32 Cherepanov P and Wackernagel W (1995) Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9–14.
- 33 Datsenko K and Wanner B (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97**, 6640–6645.
- 34 Lindenstrauß U, Skorupa P, McDowall JS, Sargent F and Pinske C (2017) The dual-function chaperone HycH improves assembly of the formate hydrogenlyase complex. *Biochem J* **474**, 2937–2950.
- 35 Laemmli U (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- 36 Towbin H, Staehelin T and Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci USA* **76**, 4350–4354.
- 37 Abràmoff M, Magalhaes P and Ram S (2004) Image processing with ImageJ. *Biophoton Int* **11**, 36–42.
- 38 Ballantine S and Boxer D (1985) Nickel-containing hydrogenase isoenzymes from anaerobically grown *Escherichia coli* K-12. *J Bacteriol* **163**, 454–459.
- 39 Pinske C (2018) The ferredoxin-like proteins HydN and YsaA enhance redox dye-linked activity of the formate dehydrogenase H component of the formate hydrogenlyase complex. *Front Microbiol* **9**, 1238.
- 40 Lowry O, Rosebrough N, Farr A and Randall R (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* **193**, 265–275.
- 41 Lubran MM (1978) The measurement of total serum proteins by the Biuret method. *Ann Clin Lab Sci* **8**, 106–110.
- 42 Pinske C, Krüger S, Soboh B, Ihling C, Kuhns M, Braussemann M, Jaroschinsky M, Sauer C, Sargent F, Sinz A *et al.* (2011) Efficient electron transfer from hydrogen to benzyl viologen by the [NiFe]-hydrogenases of *Escherichia coli* is dependent on the coexpression of the iron-sulfur cluster-containing small subunit. *Arch Microbiol* **193**, 893–903.
- 43 Andrews SC, Berks BC, McClay J, Ambler A, Quail MA, Golby P and Guest JR (1997) A 12-cistron *Escherichia coli* operon (*hyf*) encoding a putative proton-translocating formate hydrogenlyase system. *Microbiology* **143** (Pt 11), 3633–3647.
- 44 Redwood MD, Mikheenko IP, Sargent F and Macaskie LE (2008) Dissecting the roles of *Escherichia coli* hydrogenases in biohydrogen production. *FEMS Microbiol Lett* **278**, 48–55.
- 45 Sawers RG (2006) Differential turnover of the multiple processed transcripts of the *Escherichia coli* *focA-pflB* operon. *Microbiology* **152**, 2197–2205.
- 46 Huerta AM and Collado-Vides J (2003) Sigma70 promoters in *Escherichia coli*: specific transcription in dense regions of overlapping promoter-like signals. *J Mol Biol* **333**, 261–278.
- 47 Pinske C and Sawers RG (2014) The importance of iron in the biosynthesis and assembly of [NiFe]-hydrogenases. *Biomol Concepts* **5**, 55–70.
- 48 Erhardt H, Steimle S, Muders V, Pohl T, Walter J and Friedrich T (2012) Disruption of individual *nuo*-genes leads to the formation of partially assembled NADH:ubiquinone oxidoreductase (complex I) in *Escherichia coli*. *Biochim Biophys Acta* **1817**, 863–871.
- 49 Prüß BM, Nelms JM, Park C and Wolfe AJ (1994) Mutations in NADH:ubiquinone oxidoreductase of *Escherichia coli* affect growth on mixed amino acids. *J Bacteriol* **176**, 2143–2150.
- 50 McDowall JS, Hjersing MC, Palmer T and Sargent F (2015) Dissection and engineering of the *Escherichia coli* formate hydrogenlyase complex. *FEBS Lett* **589**, 3141–3147.
- 51 Yachdav G, Koppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, Hönigschmid P, Schaffner A, Roos M, Bernhofer M *et al.* (2014) PredictProtein – an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res* **42**, W337–W343.
- 52 Waterhouse AM, Procter JB, Martin DMA, Clamp M and Barton GJ (2009) Jalview version 2 – a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191.
- 53 Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF *et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462.

- 54 Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko K, Tomita M, Wanner B and Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006 0008.
- 55 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M and Söding J *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539.