



University of Dundee

The plant pathogen *Pectobacterium atrosepticum* contains a functional formate hydrogenlyase-2 complex

Finney, Alexander J.; Lowden, Rebecca; Fleszar, Michal; Albareda, Marta; Coulthurst, Sarah J.; Sargent, Frank

Published in:
Molecular Microbiology

DOI:
[10.1111/mmi.14370](https://doi.org/10.1111/mmi.14370)

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Finney, A. J., Lowden, R., Fleszar, M., Albareda, M., Coulthurst, S. J., & Sargent, F. (2019). The plant pathogen *Pectobacterium atrosepticum* contains a functional formate hydrogenlyase-2 complex. *Molecular Microbiology*, 112(5), 1440-1452. <https://doi.org/10.1111/mmi.14370>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.


- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



The plant pathogen *Pectobacterium atrosepticum* contains a functional formate hydrogenlyase-2 complex

Alexander J. Finney,^{1,2} Rebecca Lowden,²
Michal Fleszar,² Marta Albareda,^{2,3,†}
Sarah J. Coulthurst² and Frank Sargent^{1,2,*} 

¹School of Natural & Environmental Sciences, Newcastle University, Newcastle Upon Tyne, NE1 7RU, UK.

²School of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK.

³Centro de Biotecnología y Genómica de Plantas (C.B.G.P.) UPM-INIA, Universidad Politécnica de Madrid, Campus de Montegancedo, Pozuelo de Alarcón, 28223, Spain.

Summary

Pectobacterium atrosepticum SCRI1043 is a phytopathogenic Gram-negative enterobacterium. Genomic analysis has identified that genes required for both respiration and fermentation are expressed under anaerobic conditions. One set of anaerobically expressed genes is predicted to encode an important but poorly understood membrane-bound enzyme termed formate hydrogenlyase-2 (FHL-2), which has fascinating evolutionary links to the mitochondrial NADH dehydrogenase (Complex I). In this work, molecular genetic and biochemical approaches were taken to establish that FHL-2 is fully functional in *P. atrosepticum* and is the major source of molecular hydrogen gas generated by this bacterium. The FHL-2 complex was shown to comprise a rare example of an active [NiFe]-hydrogenase-4 (Hyd-4) isoenzyme, itself linked to an unusual selenium-free formate dehydrogenase in the final complex. In addition, further genetic dissection of the genes encoding the predicted membrane arm of FHL-2 established surprisingly that the majority of genes encoding this domain are not required for physiological hydrogen production activity. Overall, this study presents *P. atrosepticum* as a new model bacterial system for understanding anaerobic formate and hydrogen

metabolism in general, and FHL-2 function and structure in particular.

Introduction

Many members of the γ -proteobacteria are facultative anaerobes with the ability to switch their metabolisms to exploit the prevailing environmental conditions. Aerobic or anaerobic respiration is generally preferred, depending on the availability of respiratory electron acceptors. In this phylum, and specifically under anaerobic conditions, the three-carbon product of glycolysis, pyruvate, is often further metabolised by the oxygen-sensitive pyruvate formate lyase enzyme to generate acetyl CoA and the one-carbon compound formic acid (Pinske and Sawers, 2016). Studies of the model bacterium *Escherichia coli* have established that endogenously produced formate is initially excreted directly from the cell using a dedicated channel (Suppmann and Sawers, 1994; Hunger *et al.*, 2014; Mukherjee *et al.*, 2017). Under respiratory conditions this formate would be used as an electron donor through the activity of periplasmic enzymes, but under fermentative conditions the formate accumulates in the extracellular milieu until its rising concentration begins to cause a drop in extracellular pH. This is thought to trigger formate re-uptake, which in turn induces synthesis of formate hydrogenlyase (FHL) activity in the cell (Rossmann *et al.*, 1991; McDowall *et al.*, 2014; Sargent, 2016). FHL activity then proceeds to detoxify the formic acid by disproportionation to carbon dioxide and molecular hydrogen (H₂).

While FHL activity has been characterised in *E. coli* (Sargent, 2016), it is not confined to enteric bacteria and has been reported across the prokaryotic domains, including in hyperthermophilic archaea where it is not only involved in pH homeostasis but also in generating transmembrane ion gradients (Kim *et al.*, 2010; Lim *et al.*, 2014; Bae *et al.*, 2015). The ion-pumping activity stems from an evolutionary link between FHL and the respiratory NADH dehydrogenase Complex I (Bohm *et al.*, 1990; Friedrich and Scheide, 2000; Batista *et al.*, 2013; Marreiros *et al.*, 2013; Schut *et al.*, 2016). Like Complex I, FHL comprises a cytoplasm-facing catalytic domain (termed the

Accepted 13 August, 2019. *For correspondence. Email frank.sargent@newcastle.ac.uk; Tel. +44 191 20 85138; Fax +44 191 208 5651. Present address: [†]Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, Madrid, Spain.

peripheral arm in Complex I terminology) linked to an integral membrane arm. In FHL, the peripheral arm contains a [NiFe]-hydrogenase of the 'Group 4' type, which is primarily dedicated to H₂ production (Greening *et al.*, 2015), and is linked by [Fe-S]-cluster-containing proteins to a molybdenum-dependent formate dehydrogenase (Maia *et al.*, 2015). The FHL membrane arm is predicted to take two different forms allowing the enzyme to be further sub-classified as either 'FHL-1' or 'FHL-2' (Friedrich and Scheide, 2000; Marreiros *et al.*, 2013; Sargent, 2016; Finney and Sargent, 2019). The FHL-1 is the predominant archetypal FHL activity of *E. coli* (McDowall *et al.*, 2014) and comprises [NiFe]-hydrogenase-3 (Hyd-3), formate dehydrogenase-H (FdhF), and a relatively small membrane arm compared to Complex I that contains only two proteins (Fig. 1A). Genes for the much less well-understood FHL-2 enzyme are also found in *E. coli* (Andrews *et al.*, 1997). This isoenzyme is predicted to comprise a [NiFe]-hydrogenase-4 (Hyd-4), an as-yet undefined formate dehydrogenase, and a much larger membrane arm than FHL-1, containing at least five individual integral membrane subunits and more closely resembling the Complex I structure (Fig. 1B). Understanding the structure, function and physiological role of *E. coli* Hyd-4 and FHL-2 has been hindered by poor native expression levels (Skibinski *et al.*, 2002; Self *et al.*, 2004); a missing important accessory gene from the *E. coli* *hyf* cluster (Sargent, 2016); and a lack of consensus on the appropriate experimental conditions to test (Bagramyan *et al.*, 2001; Mnatsakanyan *et al.*, 2004). Thus, in order to bring fresh impetus to understanding the physiology and biochemistry of the FHL-2 complex, it was considered important that an appropriate alternative biological model system was established. In this work, *Pectobacterium atrosepticum* SCRI1043 was chosen (Bell *et al.*, 2004; Babujee *et al.*, 2012). However, very recently an operon encoding a Hyd-4 isoenzyme was cloned from *Trabulsiella guamensis*, which is a bacterium previously mistaken for a subspecies of *Salmonella* (McWhorter *et al.*, 1991), and found to be functional in *E. coli* (Lindenstrauß and Pinske, 2019).

Pectobacterium atrosepticum SCRI1043 is a phytopathogenic γ -proteobacterium that can grow under anaerobic conditions (Babujee *et al.*, 2012). A global transcriptomic study identified a chromosomal locus (Fig. 1C) that was transcribed under anaerobic conditions in this organism (Bell *et al.*, 2004; Babujee *et al.*, 2012). This locus neatly collects together almost all of the known genes for hydrogen metabolism (Fig. 1C), including genes for a bidirectional Hyd-2-type [NiFe]-hydrogenase; genes for specialist metallo-cofactor biosynthesis; a putative formate-responsive transcriptional regulator; a predicted formate dehydrogenase

gene; and an 11-cistron operon apparently encoding a Hyd-4 isoenzyme and its associated accessory proteins (Babujee *et al.*, 2012).

In this work, a molecular genetic approach was taken to characterise the hydrogen metabolism locus of *P. atrosepticum*. A bank of un-marked and in-frame gene deletion mutants was constructed and used to demonstrate unequivocally that the unusual FHL-2 identified in the genome is functional in *P. atrosepticum* and responsible for the majority of H₂ production under anaerobic conditions. The complex was shown to contain an active Hyd-4 and, unusually, a version of formate dehydrogenase that does not rely on selenocysteine. Surprisingly, it was shown that many of the genes encoding the large membrane arm of FHL-2 can be removed without adversely affecting H₂ production activity. This has potential implications for the molecular architecture of the membrane arm. Overall, this work introduces *P. atrosepticum* as a tractable model system and presents important genetic, biochemical and physiological characterisation of FHL-2 and [NiFe]-hydrogenase-4.

Results

P. atrosepticum produces molecular hydrogen under anaerobic conditions

P. atrosepticum SCRI1043 (Bell *et al.*, 2004) contains the genes for potentially H₂-evolving enzymes (Babujee *et al.*, 2012). Therefore, the initial goal of this study was to establish the growth conditions under which molecular hydrogen could be evolved. First, the SCRI1043 wild-type strain was grown under anaerobic fermentative conditions in a minimal medium supplemented with 0.8% (w/v) glucose. The culture headspace was sampled at periodic intervals and the amount of H₂ present quantified by gas chromatography (GC). Under these conditions, H₂ evolution activity was found to be temperature dependent, with H₂ accumulation in the headspace observed to be maximal when the phytopathogen was incubated at 20 or 24°C (Fig. 2A). Taking forward 24°C as standard incubation temperature, H₂ evolution was observed and found to level off after 40 h incubation (Fig. 2B).

When anaerobic respiratory conditions were tested, comprising 0.5% (v/v) glycerol and 0.4% (w/v) nitrate, H₂ production was found to cease with no H₂ detectable after 48 h growth (Fig. 2C). However, replacement of nitrate with 0.4% (w/v) fumarate as a terminal electron acceptor allowed the generation of low, but detectable, levels of H₂ (Fig. 2C). Maximal H₂ production is observable under fermentative conditions (Fig. 2C).

Hyd-4 is the predominant hydrogen-producing enzyme in *P. atrosepticum*

To determine the molecular basis of the observed H₂ production activity, a molecular genetic approach was taken. Initially, the genes encoding the catalytic subunits of the [NiFe]-hydrogenases were targeted. First, a strain PH001 (Table 1) was constructed carrying an unmarked in-frame deletion of the *hyfG* gene, predicted to encode the catalytic subunit of a Hyd-4 isoenzyme. When cultured

fermentatively in the presence of glucose, the PH001 (Δ *hyfG*) strain produced less than 5% of the total H₂ accumulated by the wild-type control under the same conditions (Fig. 3A). Next, the gene encoding the catalytic subunit of Hyd-2 (*hybC*) was tested. Mutant strain PH002 (Table 1) was prepared carrying only a Δ *hybC* allele and, in this case, H₂ evolution under fermentative conditions was essentially indistinguishable from the wild-type strain (Fig. 3A). Finally, a Δ *hybC* Δ *hyfG* double mutant (PH003,

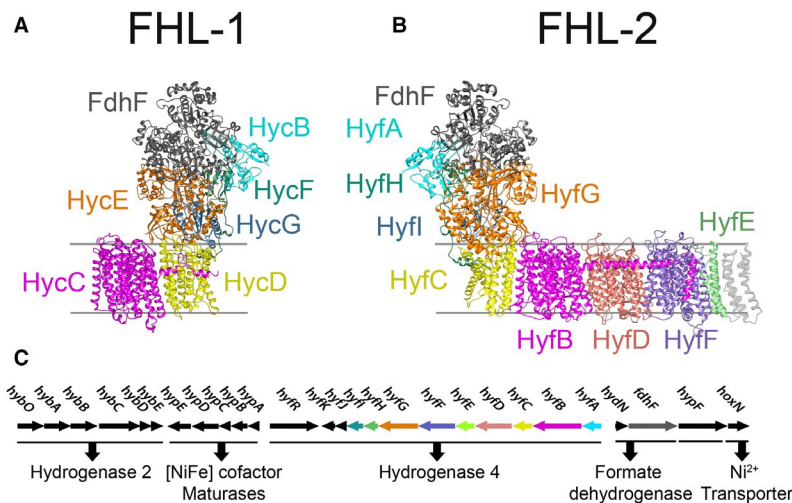


Fig. 1. Biochemistry and genetics of formate hydrogenlyase. Structural models of (A) formate hydrogenlyase-1 (FHL-1) from *Escherichia coli* and (B) formate hydrogenlyase-2 (FHL-2) from *Pectobacterium atrosepticum*. Subunits related at the primary and tertiary levels are coloured similarly. Structural modelling of the formate hydrogenlyases complexes was performed using Phyre² predictions of respective subunits (Kelley and Sternberg, 2009). Using Chimera (Pettersen et al., 2004) and the Cryo-EM structure of the *Pyrococcus furiosus* Membrane Bound Hydrogenase, MBH (PDB: 6CFW), individual FHL-2 subunits were manually assembled. FdhF, which is not present in *Pyrococcus furiosus* MBH, was positioned principally to align its [4Fe–4S] cluster with that of the surface-exposed [4Fe–4S] cluster from HyfA. C. The genetic organisation of the hydrogen metabolism gene cluster of *P. atrosepticum* (ECA1225–ECA1252). Predicted gene product functions are indicated and the operon for Hyd-4 is colour coded to match the structure model in panel (B).

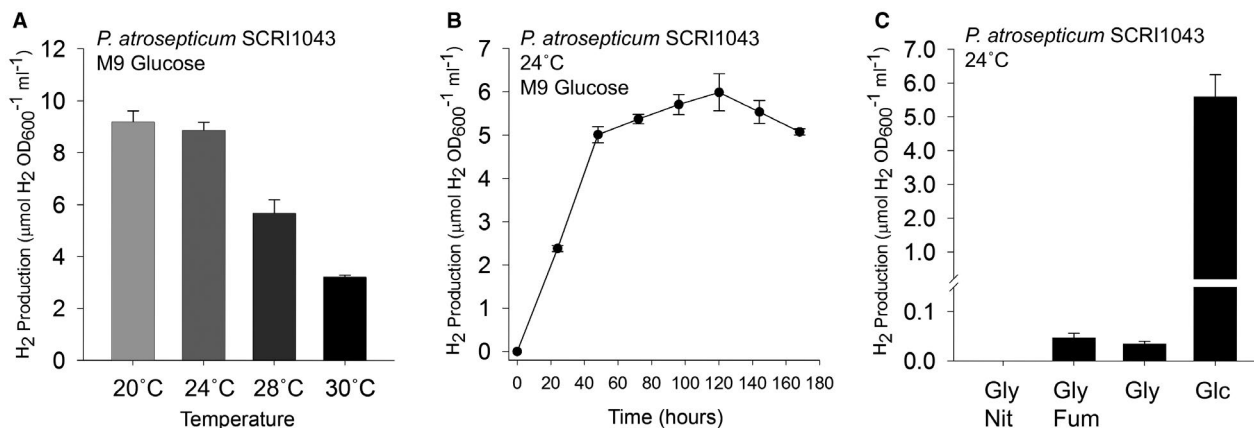


Fig. 2. *P. atrosepticum* produces molecular hydrogen gas.

A. Anaerobic hydrogen production is optimal at lower temperatures. The *P. atrosepticum* SCRI1043 parent strain was incubated in M9 medium supplemented with 0.8% (w/v) glucose for 168 h at the temperatures indicated before gaseous H₂ accumulation was quantified.

B. A time course of H₂ accumulation. *P. atrosepticum* SCRI1043 was incubated in M9 medium supplemented with 0.8% (w/v) glucose at 24°C and gaseous H₂ accumulation was measured every 24 h.

C. *P. atrosepticum* SCRI1043 was incubated in M9 medium supplemented with either 0.5% (v/v) glycerol and 0.4% (w/v) nitrate ('Gly Nit'); 0.5% (v/v) glycerol and 0.4% (w/v) fumarate ('Gly Fum'); 0.5% (v/v) glycerol only (Gly); or 0.8% (w/v) glucose only ('Glc') at 24°C for 48 h. In all cases, the levels of molecular H₂ in the culture headspace were quantified by GC and normalised to OD₆₀₀ and culture volume. Error bars represent SD ($n = 3$).

Table 1. *P. atrosepticum* strains and plasmids used in this study.

Strain	Relevant genotype	Genomic identifier	Source
SCRI1043	–		Bell <i>et al.</i> (2004)
PH001	$\Delta hyfG$	ECA1241	This work
PH002	$\Delta hybC$	ECA1228	This work
PH003	$\Delta hybC \Delta hyfG$	ECA1228, ECA1241	This work
PH004	$\Delta fdhF$	ECA1250	This work
PH005	$\Delta hybC \Delta fdhF$	ECA1228, ECA1250	This work
PH007	$\Delta hybC \Delta hyfB-F$	ECA1228, ECA1246-2	This work
PH008	$\Delta hybC \Delta hyfD-F$	ECA1228, ECA1244-2	This work
PH009	$\Delta hybC hyfG^{His}$	ECA1228, ECA1241	This work
PH010	$\Delta hypF$	ECA1251	This work
PH011	$\Delta hoxN$	ECA1252	This work
PH013	$\Delta hybC \Delta fdhD$	ECA1228, ECA0093	This work
PH015	$\Delta hybC \Delta hyfR$	ECA1228, ECA1236	This work
PH018	$\Delta hybC \Delta ECA1964$	ECA1228, ECA1964	This work
PH019	$\Delta hybC \Delta fdhF \Delta ECA1964$	ECA1228, ECA1250, ECA1964	This work
PH020	$\Delta hybC hyfG^{His} \Delta hyfB-F$	ECA1228, ECA1241, ECA1246-2	This work
PH021	$\Delta hybC hyfG^{His} \Delta hyfD-F$	ECA1228, ECA1241, ECA1244-2	This work
PH027	$\Delta hybC \Delta ECA1507$	ECA1228, ECA1507	This work
PH028	$\Delta hybC \Delta fdhF \Delta ECA1507$	ECA1228, ECA1250, ECA1507	This work

Plasmid	Description	Source
pSUPROM	Vector for expression of genes under control of the <i>E. coli</i> <i>tatA</i> promoter (Kan ^R)	Jack <i>et al.</i> (2004)
pSUPROM- <i>hyfG</i>	as pSUPROM with <i>hyfG</i> (ECA1241)	This work
pSUPROM- <i>hybC</i>	as pSUPROM with <i>hybC</i> (ECA1228)	This work
pSUPROM- <i>fdhF</i>	as pSUPROM with <i>fdhF</i> (ECA1250)	This work
pSUPROM- <i>ECA1507</i>	as pSUPROM with <i>ECA1507</i>	This work
pSUPROM- <i>ECA1964</i>	as pSUPROM with <i>ECA1964</i>	This work
pSUPROM- <i>fdhD</i>	as pSUPROM with <i>fdhD</i> (ECA0093)	This work
pSUPROM- <i>hypF</i>	as pSUPROM with <i>hypF</i> (ECA1251)	This work
pSUPROM- <i>hoxN</i>	as pSUPROM with <i>hoxN</i> (ECA1252)	This work
pSUPROM- <i>hyfR</i>	as pSUPROM with <i>hyfR</i> (ECA1236)	This work

Table 1) was constructed and was found to be completely devoid of the ability to produce gaseous H₂ (Fig. 3A).

P. atrosepticum can be stably transformed and plasmids encoding either *hyfG* or *hybC* were constructed. In the case of PH001 ($\Delta hyfG$) and PH003 ($\Delta hybC \Delta hyfG$), H₂ evolution could be rescued in the mutant strains by supplying extra copies of *hyfG* on a plasmid (Fig. 3B).

Taken altogether, the data presented in Figs 2 and 3 demonstrate that Hyd-4 is responsible for the majority of physiological H₂ production by *P. atrosepticum*, and that this activity is present under fermentative conditions at temperate growth temperatures $\leq 24^{\circ}\text{C}$.

P. atrosepticum contains an active FHL-2 with a selenium-free formate dehydrogenase

Having established that Hyd-4 was active, the next task was to test the hypothesis that Hyd-4 could be part of a wider FHL-2 complex (Fig. 1). First, formate dependence on H₂ production was tested by growing the wild-type parental strain, the PH001 ($\Delta hyfG$) strain, and the PH002 ($\Delta hybC$) strain anaerobically in the presence of

increasing amounts of exogenous formate (Fig. 4A). A correlation was observed between the amount of H₂ produced and the amount of formate added to the growth medium, and this was particularly clear when the uptake hydrogenase activity was inactivated (Fig. 4A). High levels of H₂ production remained dependent upon the presence of an active Hyd-4 (Fig. 4A), providing initial evidence for a link between formate and H₂ metabolism.

The *P. atrosepticum* SCRI1043 genome contains a gene encoding a putative formate dehydrogenase close to those for Hyd-4 (Fig. 1C). The gene product shares 85% overall sequence identity with *E. coli* FdhF but interestingly contains a cysteine residue at position 140 (Fig. S1), which is occupied by a critical selenocysteine in the *E. coli* version and other related enzymes (Axley *et al.*, 1991). A mutant strain was therefore constructed (PH004, Table 1) carrying a $\Delta fdhF$ allele. The PH004 ($\Delta fdhF$) strain produced very low, but detectable, levels of H₂ gas under fermentative conditions (Fig. 4B). Addition of a $\Delta hybC$ allele to the $\Delta fdhF$ strain to generate a double mutant (PH005, Table 1) had no further effect on the amount of H₂ that could be produced (Fig. 4B).

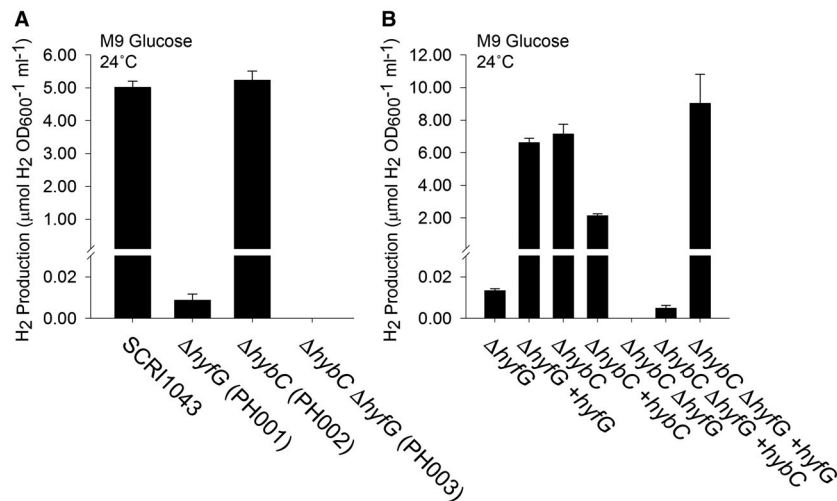


Fig. 3. Hydrogen gas is produced by the activity of [NiFe]-Hydrogenase-4.

A. Hyd-4 is responsible for fermentative H₂ production. *P. atrosepticum* parental strain SCRI1043 and mutants PH001 (Δ hyfG), PH002 (Δ hybC) and PH003 (Δ hybC Δ hyfG) were incubated in M9 medium supplemented with 0.8% (w/v) glucose at 24°C for 48 h.

B. Complementation of the mutant phenotype *in trans*. Strains PH001 (Δ hyfG), PH002 (Δ hybC) and PH003 (Δ hybC Δ hyfG) were separately transformed with plasmids encoding either HyfG or HybC under the control of constitutive promoters. Levels of molecular H₂ in the culture headspace were quantified by GC and normalised to OD₆₀₀ and culture volume. Error bars represent SD ($n = 3$).

It is interesting, however, that a strain devoid of both *hybC* (Hyd-2 activity) and *hyfG* (Hyd-4 activity) could not produce any H₂ gas (Fig. 3B). This suggests that the low levels of H₂ evolved from the Δ hybC, Δ fdhF strain are derived from Hyd-4 but that an alternative electron donor may be operating. Notably, two further genes encoding homologs of FdhF are encoded on the *P. atrosepticum* SCRI1043 chromosome (Bell *et al.*, 2004). The *ECA1507* gene encodes a protein with 65% overall sequence identity with FdhF, and the *ECA1964* gene encodes a protein with 22% overall sequence identity with FdhF. Deletion of the genes encoding *ECA1507* or *ECA1964* alone (Table 1) had no influence on the H₂ production capability of the bacterium (Fig. 4C). Moreover, when the genes were supplied in multicopy on an expression vector, neither was able to rescue the phenotype of the Δ fdhF mutant back to native levels of H₂ production (Fig. 4D). However, it is clear that extra levels of *ECA1507* in the cell result in a slight increase in H₂ accumulation over the time course of the experiment (Fig. 4D). The increase in H₂ production is statistically significant ($P < 0.0001$) and suggests *ECA1507* could function as an alternative electron donor subunit for *P. atrosepticum* Hyd-4. Interestingly, extra copies of *ECA1964* on a plasmid had the opposite effect (Fig. 4D). In this case, H₂ production in the Δ fdhF strain was pushed to a statistically significant ($P < 0.0001$) even lower level (Fig. 4D), suggesting *ECA1964* was either largely inactive or interfering with the interaction of Hyd-4 with other redox partners.

Taken altogether, these data establish that *P. atrosepticum* SCRI1043 has functional formate hydrogenlyase activity

where molecular hydrogen production is clearly linked to both formate availability and the presence of a formate dehydrogenase gene. Importantly, the predominant electron donor for the reaction is an unusual version of formate dehydrogenase that does not require selenocysteine at its active site, and the enzyme responsible for proton reduction is a [NiFe]-hydrogenase-4.

The role of the FHL-2 membrane arm in hydrogen production

One clear defining structural difference between the FHL-1 type formate hydrogenlyase found in *E. coli* and the FHL-2 type of *P. atrosepticum* SCRI1043 is the number of genes encoding components of the membrane arms (Fig. 1). An FHL-1 enzyme is predicted to contain two different membrane proteins, HycC (related to HyfB in FHL-2) and HycD (related to HyfC) (Fig. 1A). Alternatively, an FHL-2 enzyme is predicted to contain three additional membrane proteins, including HyfE (not present in FHL-1) and two further homologs of HycC/HyfB, namely HyfD and HyfF (Fig. 1B).

To explore the roles of the extra *hyfDEF* genes located within the FHL-2 locus, mutant strains were constructed (Table 1). First, versions of the Δ hybC strain PH002, lacking either the genes encoding the entire FHL-2 membrane arm (PH007: Δ hybC, Δ hyfBCDEF – Table 1) or lacking the extra membrane components not found in FHL-1 (PH008: Δ hybC, Δ hyfDEF – Table 1) were constructed. In addition, the Δ hybC strain PH002, producing Hyd-4 as the only active hydrogenase, was modified by addition of a 10-His sequence between codons 82 and

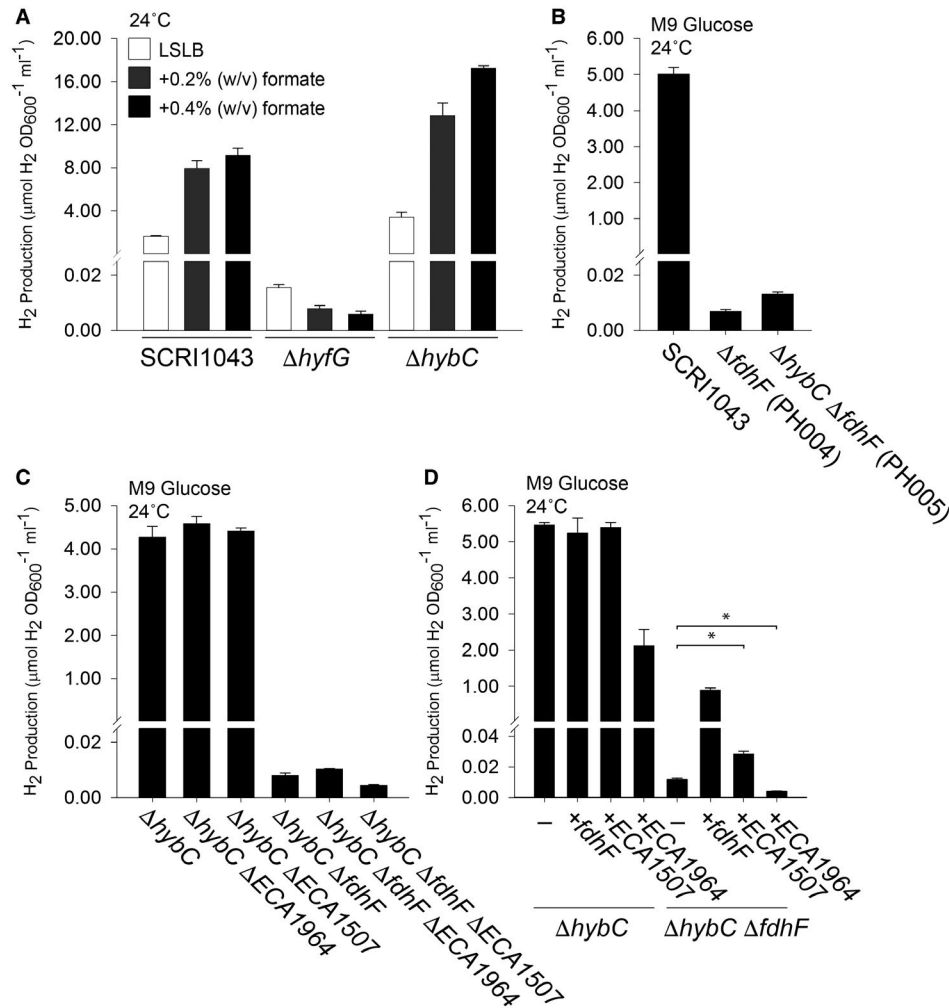


Fig. 4. Hydrogen gas is produced by the activity of a selenium-free formate dehydrogenase.

A. Addition of exogenous formate increases H_2 production. *P. atrosepticum* parental strain SCRI1043 and mutants PH001 ($\Delta hyfG$) and PH002 ($\Delta hybC$) were incubated in low-salt (5 g l^{-1}) LB (LSLB) rich medium supplemented with 0.2% or 0.4% (w/v) formate at 24°C for 48 h.
 B. The formate dehydrogenase encoded within the gene cluster is responsible for FHL-2 activity. Strains SCRI1043, PH004 ($\Delta fdhF$), PH005 ($\Delta hybC \Delta fdhF$) were incubated in M9 medium supplemented with 0.8% (w/v) glucose at 24°C for 48 h.
 C. Alternative formate dehydrogenase homologues do not have a major role in H_2 production. Strains SCRI1043, PH002 ($\Delta hybC$), PH019 ($\Delta hybC \Delta ECA1964$), PH028 ($\Delta hybC \Delta ECA1507$) and PH005 ($\Delta hybC \Delta fdhF$) were incubated in M9 medium supplemented with 0.8% (w/v) glucose at 24°C for 48 h.
 D. Complementation of the mutant phenotype *in trans*. Strains PH002 ($\Delta hybC$) and PH005 ($\Delta hybC \Delta fdhF$) were separately transformed with plasmids encoding either FdhF, ECA1964 or ECA1507 under the control of constitutive promoters. In all cases, the levels of molecular H_2 in the culture headspace were quantified by GC and normalised to OD_{600} and culture volume. Error bars represent SD ($n = 3$). In panel (D) a one-tailed *t*-test was used to determine statistical significance ($*P < 0.0001$).

83 of the *hyfG* gene. This new epitope-tagged strain was called PH009 ($\Delta hybC$, *hyfG*^{His} – Table 1). Finally, versions of PH009 lacking either the genes encoding the entire FHL-2 membrane arm (PH020: $\Delta hybC$, *hyfG*^{His}, $\Delta hyfBCDEF$ – Table 1) or lacking the extra membrane components not found in FHL-1 (PH021: $\Delta hybC$, *hyfG*^{His}, $\Delta hyfDEF$ – Table 1) were constructed.

Deletion of the genes encoding the entire membrane arm reduced the FHL-2-dependent H_2 accumulation levels to around 5% of that observed in the parent strain (Fig. 5A). The addition of the 10-His tag to HyfG allowed the Hyd-4 catalytic subunit to be visualised in whole cell

extracts by western immunoblotting (Fig. 5B). The polypeptide was clearly detectable when *P. atrosepticum* was cultured under anaerobic fermentative conditions (Fig. 5C). Interestingly, the amount of cellular HyfG^{His} was seen to increase when the genes encoding the membrane arm were removed (Fig. 5). This is particularly pertinent for the PH020 strain ($\Delta hybC$, $\Delta hyfBCDEF$), which is essentially devoid of FHL-2 activity (Fig. 5A), since it can be concluded that genetic removal of the complete membrane arm does not destabilise the Hyd-4 catalytic subunit, but instead leads to a physiologically inactive enzyme. It is also notable that in the absence of the genes

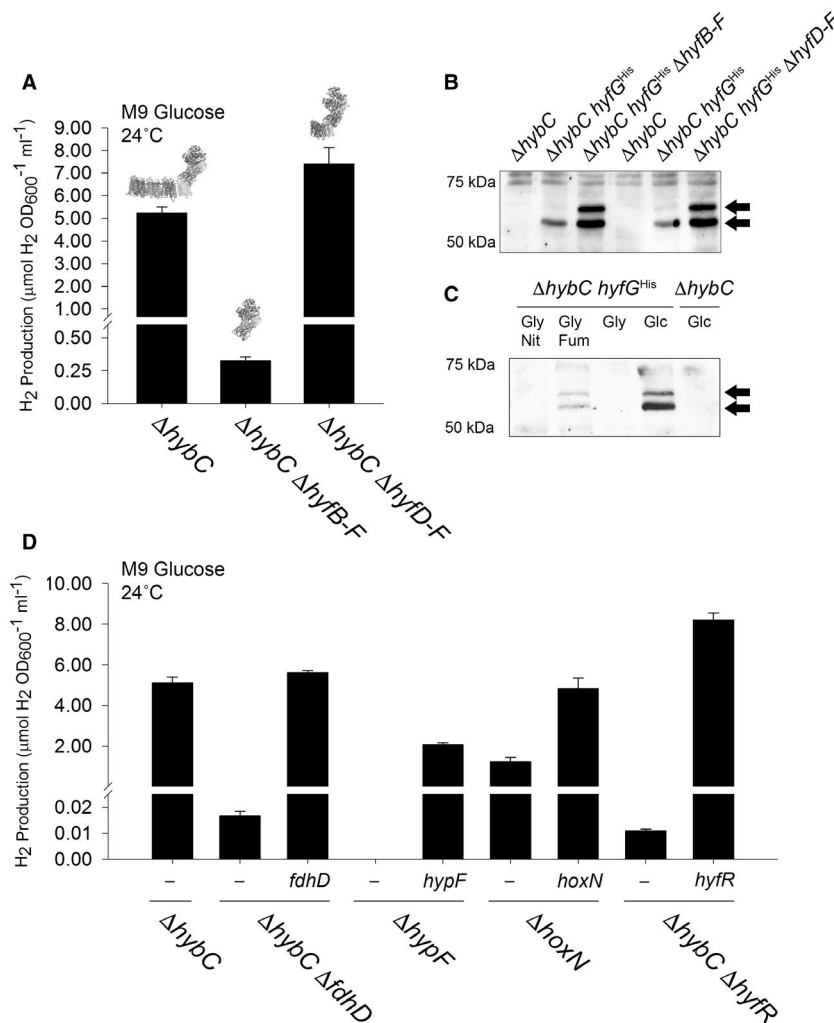


Fig. 5. Genetic dissection of FHL-2 activity.

A. Some genes encoding the membrane arm are not essential for FHL-2 activity. *P. atrosepticum* strains PH002 (Δ*hybC*), PH009 (Δ*hybC* Δ*hyfB-F*) and PH08 (Δ*hybC* Δ*hyfD-F*) were incubated in M9 medium supplemented with 0.8% (w/v) glucose at 24°C for 48 h.

B. HyfG^{His} can be detected in strains devoid of membrane subunits. *P. atrosepticum* strains PH002 (Δ*hybC*), PH009 (Δ*hybC*, *hyfG*^{His}), PH020 (Δ*hybC* *hyfG*^{His}, Δ*hyfB-F*) and PH021 (Δ*hybC* *hyfG*^{His}, Δ*hyfD-F*) were incubated in M9 medium supplemented with 0.8% (w/v) glucose at 24°C for 48 h. Whole cell samples were then prepared by centrifugation, separation of proteins by SDS-PAGE, transfer to nitrocellulose and HyfG^{His} probed with an anti-HIS-HRP antibody.

C. HyfG^{His} is induced upon glucose fermentation. Strains PH002 (Δ*hybC*) and PH009 (Δ*hybC*, *hyfG*^{His}) were incubated in M9 medium supplemented with either 0.5% (v/v) glycerol and 0.4% (w/v) nitrate ('Gly Nit'); 0.5% (v/v) glycerol and 0.4% (w/v) fumarate ('Gly Fum'); 0.5% (v/v) glycerol only (Gly); or 0.8% (w/v) glucose only ('Glc') at 24°C for 48 h. Whole cell samples were probed for HyfG^{His} with an anti-HIS-HRP antibody.

D. The role of accessory genes in FHL-2 activity. Strains PH002 (Δ*hybC*), PH013 (Δ*hybC* Δ*fdhD*), PH010 (Δ*hypF*), PH011 (Δ*hoxN*), and PH015 (Δ*hybC* Δ*hyfR*) were transformed with plasmids encoding either FdhD, HypF, HoxN or HyfR under the control of constitutive promoters. In all cases, the levels of molecular H₂ in the culture headspace were quantified by GC and normalised to OD₆₀₀ and culture volume. Error bars represent SD (*n* = 3).

encoding membrane proteins that the HyfG^{His} protein migrates as two electrophoretic species during SDS-PAGE (Fig. 5B). This is a common observation for catalytic subunits of [NiFe]-hydrogenases as they are synthesised as precursors that undergo proteolytic processing at the C-terminus following cofactor insertion (Bock *et al.*, 2006). In this case, the faster migrating species was calculated as 56.4 kDa, while the slower migrating species was estimated as 62.6 kDa by SDS-PAGE. The predicted molecular mass of HyfG^{His} prior to proteolytic processing

is 67,559 Da, and the predicted molecular weight of the 32-residue C-terminal tail that is removed is 3,821 Da.

Conversely, partial modification of the FHL-2 membrane arm to leave only those subunits present in FHL-1 (Δ*hybC*, Δ*hyfDEF*) had no negative effect on hydrogen production levels (Fig. 5A), rather a slight increase was observed. This is consistent with a noticeable enhancement of HyfG^{His} levels in the cells upon removal of the *hyfDEF* genes (Fig. 5B). The available evidence suggests that HyfD, HyfE and HyfF have no essential roles

in the biosynthesis and hydrogen production activity of FHL-2.

A requirement for accessory genes in anaerobic hydrogen production

FHL-2 is a multi-subunit metalloenzyme and assembly of such enzymes is often carefully co-ordinated by dedicated chaperones, sometimes called accessory proteins or 'maturases'. Maturation of molybdenum-dependent formate dehydrogenases has been reported to require the action of an FdhD protein, which is believed to supply an essential sulfur ligand to the active site metal (Arnoux *et al.*, 2015). In *P. atrosepticum* SCRI1043, *fdhD* (ECA0093) is not part of the FHL-2 locus but is located elsewhere on the chromosome next to a gene encoding superoxide dismutase (*sodA* or ECA0092) (Bell *et al.*, 2004). Genetic modification of the PH002 strain, containing only Hyd-4 and FHL-2 activities, by the incorporation of a $\Delta fdhD$ allele (PH013: $\Delta hybC$, $\Delta fdhD$ – Table 1) led to a defect in physiological H₂ production under fermentative conditions (Fig. 5D). This phenotype could be rescued by the provision of extra copies of *fdhD* *in trans* (Fig. 5D).

Maturation of [NiFe]-hydrogenases requires the activity of a network of proteins involved in metal homeostasis and cofactor maturation and insertion (Sargent, 2016). The *P. atrosepticum* FHL-2 locus (Fig. 1C) contains a *hoxN* gene (ECA1252) encoding a putative membrane-bound nickel ion transporter (Eitinger and Mandrand-Berthelot, 2000). Deletion of the *hoxN* gene in *P. atrosepticum* SCRI1043 (strain PH011, Table 1) reduced hydrogen evolution levels to around 50% of that observed for the parental strain (Fig. 5D). Note that there is no other homologue of *hoxN* encoded on the *P. atrosepticum* SCRI1043 genome, but there are several uncharacterised ABC transporters that could be related to the high-affinity *nikA* system (Wu *et al.*, 1991), which could account for the continued availability of nickel for hydrogenase biosynthesis in this experiment.

Once inside the cell, nickel is processed into the Ni–Fe–CO–2CN[−] cofactor through the action of several enzymes and chaperones. One key step in the biosynthesis of the cofactor is the first step in the generation of CN[−] from carbamoyl phosphate by HypF (Sargent, 2016). Deletion of the *hypF* gene from *P. atrosepticum* (PH010, Table 1), which is located in the hydrogen metabolism gene cluster under investigation here (Fig. 1C), led to the complete abolition of all detectable H₂ evolution (Fig. 5D). The mutant phenotype could be rescued by supply of *hypF* *in trans*, but note that full H₂ evolution levels were not restored (Fig. 5D).

Finally, it was observed that a member of the HyfR family of transcriptional regulators was encoded in the hydrogen metabolism gene cluster (Fig. 1C). The HyfR protein is predicted to be related to FhlA, which is a

formate-sensing transcriptional activator (Skibinski *et al.*, 2002). A $\Delta hyfR$ strain devoid of the HyfR protein has very low formate hydrogenlyase-2 activity (Fig. 5D).

Taken altogether, it can be concluded that all of the genes required for biosynthesis of FHL-2 are functional in *P. atrosepticum* SCRI1043, which is entirely consistent with the physiological data reported here.

Discussion

Key differences between FHL-2 and FHL-1

Formate hydrogenlyases can be classified into two structural classes, FHL-1 and FHL-2 (Finney and Sargent, 2019). The most obvious structural difference between an FHL-1, such as the best-characterised *E. coli* enzyme (McDowall *et al.*, 2014; Pinske and Sargent, 2016), and an FHL-2, such as the *P. atrosepticum* enzyme characterised here, is the predicted size and composition of the membrane arm (Fig. 1B). Indeed, this large membrane arm is thought to be the ancient progenitor to the ion-pumping membrane arm of respiratory Complex I (Batista *et al.*, 2013; Marreiros *et al.*, 2013; Yu *et al.*, 2018). Although eukaryotic Complex I, prokaryotic Complex I, and Group 4 hydrogenases such as FHL-1, FHL-2, Ech and MBH are clearly evolutionarily related, the gene and protein names for each type of enzyme are different. Some review articles contain useful tables to highlight the relatedness of the individual subunits (Friedrich and Scheide, 2000; Marreiros *et al.*, 2013; Schut *et al.*, 2016). FHL-1 includes only two membrane proteins, which are a single HycD/HyfC-type protein together with a single HycC/HyfB. This is sufficient to anchor the peripheral arm close to the membrane and, in the case of *Thermococcus onnurineus* FHL-1 (Lim *et al.*, 2014) and the related Ech hydrogenase from *Methanosarcina mazei* (Welte *et al.*, 2010), will also allow initial generation of a proton gradient (Yu *et al.*, 2018). Operons encoding FHL-2 complexes encode at least three further integral membrane proteins. In *P. atrosepticum* these are HyfD and HyfF, which are extra versions of the HycC/HyfB putative ion channels, and the HyfE protein, which is more closely related to a region of NuoK in Complex I. Interestingly, if FHL-2 is modelled based on the Complex I structure (Fig. S2), the extra HyfDEF proteins would be placed between HyfBC, thus separating them and pushing HyfB to the most distal point in the peripheral arm (Marreiros *et al.*, 2013). Alternatively, if FHL-2 is modelled based on the Hyd-4-like MBH structure from *Pyrococcus furiosus* (Yu *et al.*, 2018), then HyfBC remain in contact with each other and HyfDEF form the distal region of the membrane arm (Figs 1 and S2). The experimental evidence presented in this work suggests *P. atrosepticum* FHL-2 adopts a membrane arm architecture similar to the *Pyrococcus furiosus*

MBH hydrogenase (Fig. 1). This is because removal of all of the extra HyfDEF membrane proteins from FHL-2 had no discernible effect on the physiological activity of the *P. atrosepticum* system (Fig. 5A), suggesting an active FHL-1-like core enzyme remains. Clearly if HyfB was normally separated from HyfC by the extra proteins they would be unlikely to come together to form a complex when placed in a Δ hyfDEF background, and *E. coli* FHL-1, for instance, is completely inactive when lacking its HyfB homolog HycC (Pinske and Sargent, 2016). This highlights the principle of modularity in metalloenzyme evolution, since it is clear that the HyfDEF module may be added or removed depending on both selective pressure and also the, as yet undefined in terms of hydrogenases, biochemical function of these membrane proteins (Friedrich and Scheide, 2000). Indeed, it is notable that distal components of the *Pyrococcus furiosus* MBH membrane arm (MbhABC) could also be genetically removed with only minor effects on cellular hydrogenase activity (Yu *et al.*, 2018). Taken together, this perhaps points to Hyd-3 from FHL-1 as the minimal module of a Group 4 hydrogenase.

Western immunoblotting pointed towards either stabilisation or upregulation of the catalytic subunit HyfG in the absence of *hyfDEF* or *hyfBCDEF* (Fig. 5B). This is unlikely to be caused by an accumulation of formate in the cells, perhaps leading to maximal transcription, because the Δ hyfDEF strain retained normal levels of formate hydrogenase activity (Fig. 5A). It is more likely that the removal of genes encoding large membrane proteins from immediately upstream of *hyfG* relaxes some restrictions on the rates of transcription and translation. In bacteria, transcription, translation and membrane insertion of the nascent chain are thought to be coupled together in a process called transertion (Roggiani and Goulian, 2015), and removal of some or all of the elaborate membrane integration step could have an effect on translation of downstream genes.

At native levels, the HyfG^{His} protein can be detected as a single species migrating at 56.4 kDa in SDS-PAGE (Fig. 5B), and occasionally a slower migrating form is detectable migrating at 62.6 kDa (Fig. 5C). These two forms of HyfG^{His} become prominent when the membrane arm of FHL-2 is genetically modified (Fig. 5B). It is known that almost all [NiFe]-hydrogenases are proteolytically processed at their C-termini following successful insertion of the Ni–Fe–CO–2CN⁻ cofactor (Bock *et al.*, 2006). In *P. atrosepticum* HyfG, processing is expected to occur at Arg-546 and would remove 32 amino acids. Thus, in theory, HyfG should be processed from a 67.6 kDa inactive precursor to a 63.8 kDa active mature form. In practice, the motility of HyfG in SDS-PAGE does not match precisely the theoretical values (Fig. 5B and C); however, only the mature form of HyfG

could contribute to physiological formate hydrogenase activity.

The *P. atrosepticum* HyfG catalytic subunit from the Hydrogenase-4 component of FHL-2 shares 74% overall sequence identity (85% similarity) with the *E. coli* HycE protein from Hydrogenase-3/FHL-1. The sequence variation between these two Group 4A hydrogenases is therefore small with only subtle notable differences. For instance, each protein is known or predicted to undergo cleavage during cofactor insertion and maturation leaving a C-terminal arginine residue in the mature form of the proteins. The cleavage sites themselves are slightly differently conserved in an FHL-1-type enzyme compared to an FHL-2, for example, ...R*MTVV... for HycE-like proteins compared to ...R*VTLV... for HyfG. This may reflect the need for a different maturation protease for each type of hydrogenase, however, this remains to be tested experimentally. In addition, it is notable that both *E. coli* and *P. atrosepticum* *hyfG* initiate translation with a GUG start codon, which may have a role in controlling cellular levels of the enzyme (Belinky *et al.*, 2017).

Phylogenetic analysis of the Group 4A [NiFe]-Hydrogenase subunits, including HycE and HyfG, shows that the enzymes associated with FHL-1 separate into a clearly distinct evolutionary clade from those associated with FHL-2, which form their own distinct clade (Fig. S3). Examples of species that encode both FHL-1 and FHL-2 are rare (Fig. S3).

A selenium-free formate dehydrogenase

Arguably one of the best-studied FdhF enzymes is the *E. coli* version, which contains selenocysteine at its active site (Axley *et al.*, 1991; Gladyshev *et al.*, 1994; Boyington *et al.*, 1997). Selenocysteine is incorporated co-translationally at a special UGA 'nonsense' codon within the coding sequence (Zinoni *et al.*, 1987), and replacement of selenocysteine with cysteine in the *E. coli* enzyme resulted in a dramatically reduced turnover number (Axley *et al.*, 1991). One surprising aspect of *P. atrosepticum* SCRI1043 is that it contains none of the biosynthetic machinery to synthesise selenocysteine (Babujee *et al.*, 2012) and the *fdhF* gene studied in this work contains a cysteine codon where selenocysteine would be encoded in the *E. coli* enzyme (Fig. S1). Certainly, the discovery of an active FHL-2 with no need for selenocysteine would benefit scientists interested in engineering this activity into other biological systems. Indeed, an in-frame deletion in the *fdhF* gene located in the FHL-2 gene cluster (Fig. 1C) resulted in a ~ 500 times reduction in H₂ production (Fig. 4), indicating the majority of H₂ production from *P. atrosepticum* is dependent on this formate dehydrogenase engaging with Hyd-4 to

form an FHL-2 complex. However, the $\Delta\text{hybC } \Delta\text{fdhF}$ double mutant still produced low, but quantifiable, levels of H_2 (Fig. 4). Compare that with the behaviours of the $\Delta\text{hybC } \Delta\text{fdhF}$ strain (Fig. 3B) and the ΔhypF mutant (Fig. 5C), neither of which produced any detectable H_2 gas. This genetic approach points to the residual H_2 emitting from Hyd-4, perhaps through the activity of alternative electron donors. Certainly for the *E. coli* FHL, it is known that FdhF is only loosely attached (Boyington *et al.*, 1997) and this may be because the enzyme is 'moonlighting' in other biochemical pathways (Iwadata and Kato, 2019). It raises the possibility that other FdhF-like enzymes in particular could 'plug in' to Hyd-4 and pass excess reducing electrons on to protons. In this work, ECA1507 was found to partially rescue the phenotype of a ΔfdhF strain (Fig. 4D) suggesting it could be an alternative redox partner: note well, however, that the potential substrates and kinetics of ECA1507 cannot be reliably predicted and should be determined empirically.

The FdhF formate dehydrogenase from *P. atrosepticum* shares 65% overall sequence identity (and 85% similarity) with the well-known *E. coli* enzyme (Fig. S1). Interestingly, phylogenetic analysis suggests that > 50% of bacterial species that contain FHL genes utilise a cysteine-dependent, rather than selenocysteine-dependent, formate dehydrogenase (Fig. S4). *P. atrosepticum* ECA1507 and ECA1964 were identified here as two FdhF-like proteins that could potentially interact with Hydrogenase-4 to generate novel FHL-like complexes. Sequence analysis revealed ECA1507 and ECA1964 share 65% and 22% overall sequence identity with FdhF, respectively, and phylogenetic analysis determined that ECA1964 is more similar to *E. coli* YdeP than any other predicted molybdenum dependent oxidoreductases in *P. atrosepticum* (Fig. S5). YdeP has a putative role in acid resistance in *E. coli* (Masuda and Church, 2003).

A role for formate metabolism in a plant pathogen

In the potato pathogen *P. atrosepticum*, FHL-2 activity was found to be expressed at lower growth temperatures (Fig. 2). This suggests that FHL-2 may be produced *in planta* during the infection or colonisation event. Formate is produced endogenously by enteric bacteria under fermentative conditions, but plants and tubers have multiple metabolic pathways that generate and consume formate. Potato tubers produce a NAD^+ -dependent formate dehydrogenase (FDH), and the levels of this enzyme are boosted under stress conditions (Hourton-Cabassa *et al.*, 1998). Indeed, proteomic experiments have identified FDH as a differentially produced protein during wound healing in potato tuber slices, with order of magnitude level changes in protein during this process (Chaves *et al.*, 2009). It could be hypothesised that the

expression of FDH in the potato tuber could be coordinated with the initial secretion of formate by a fermenting pathogen. Potentially, this would generate NADH from formate in stressed or damaged plant tissues. Recently, it was shown that FDH co-ordinates cell death and defence responses to phytopathogens in *Capsicum annuum* (Bell pepper) (Choi *et al.*, 2014). There is also indication that formate and other molecules that lead to the generation of formate, such as methanol and formaldehyde, induce the production of the NAD^+ -dependent FDH, perhaps suggesting there is a signalling response to these C1 compounds in plants (Hourton-Cabassa *et al.*, 1998).

Concluding remarks

In this work, *P. atrosepticum* SCRI1043 has been established as a tractable new model organism for studying hydrogen metabolism in general and FHL function in particular. The organism is a rare example of a bacterium with an active hydrogenase-4-containing FHL-2 complex, however, in the course of this work, hydrogenase-4 activity was reported in *T. guamensis*, another γ -proteobacterium (Lindenstrauß and Pinske, 2019). Interesting, the *T. guamensis* Hyd-4 was found to be active *in vivo* but very poorly reactive *in vitro* in standard enzymatic assays with redox-active dyes (Lindenstrauß and Pinske, 2019). This again highlights the need for development of new approaches to characterise FHL-2 and its component parts. In *P. atrosepticum*, the active hydrogenase-4 enzyme operates in tandem with an unusual selenium-free formate dehydrogenase, which may be more amenable to biotechnological engineering than selenium-dependent isoenzymes. In evolutionary terms, the FHL-2 complex has been discussed as a key intermediate in the evolution of the NADH dehydrogenase (Complex I) from a structurally simpler membrane-bound hydrogenase (Friedrich and Scheide, 2000; Marreiros *et al.*, 2013; Schut *et al.*, 2016). The most obvious difference in the predicted quaternary structures inferred from the genetics is the large membrane arm present in FHL-2 compared to FHL-1, and data presented here points to the extra membrane proteins being not essential for formate-dependent hydrogen evolution *in vivo*. The role of the FHL membrane arm in generating a transmembrane ion gradient remains to be fully explored in enteric bacteria.

Experimental procedures

Bacterial strains

The parental *P. atrosepticum* strain used in this study was SCRI1043 (Bell *et al.*, 2004). In-frame deletion and insertion mutants were constructed using pKNG101 suicide vector

in *E. coli* strain CC118 λ *pir* (Kaniga *et al.*, 1991; Coulthurst *et al.*, 2006). Briefly, upstream and downstream regions (≥ 600 bp) of the target gene(s) was amplified and inserted into pKNG101 using a three fragment Gibson assembly reaction (HiFi Assembly, NEB). For the insertion of a deca-His encoding sequence into *hyfG*, primers were designed using the NEBuilder online tool to include the deca-His encoding sequence in the overlapping region of the two fragments containing the respective 3' and 5' sequences of *hyfG*. After successfully assembly and sequencing of pKNG101 plasmids, the CC118 λ *pir* strain with desired plasmid, a HH26 pNJ5000 helper strain, and the desired *P. atrosepticum* strain were grown in rich media, with antibiotics as necessary. Equal volumes of the stationary phase cultures were mixed and 30 μ l was spotted on a non-selective rich media plate for 24 h at 24°C. *P. atrosepticum* cells with the pKNG101 plasmid were initially selected for on minimal media agar with streptomycin (100 μ g ml⁻¹). After this, single colonies were re-streaked on the fresh minimal media agar with streptomycin. Co-integrants were then grown to stationary phase in rich medium with no selection before the culture was diluted 1/500 with phosphate buffer. Then, 30 μ l of this diluted culture was plated on minimal media agar with sucrose. These colonies were patch screened for sensitivity to streptomycin before PCR screens were performed to check for presence of the desired mutation(s).

Plasmids and complementation

All plasmids were cloned using Gibson assembly (HiFi Assembly, NEB) using DNA amplified from *P. atrosepticum* SCRI1043 genomic DNA (Table 1). Genes were cloned into pSU-PROM (Kan^R), which includes the constitutive *tatA* promoter from *E. coli* (Jack *et al.*, 2004). Complementation plasmids were used to transform electrocompetent *P. atrosepticum* cells using a 2 mm electroporation cuvette (Molecular BioProduct) with application of an electrical pulse (2.5 kV voltage, 25 μ F capacitance, 200 Ω resistance and 2 mm cuvette length) via a Gene Pulsar Xcell electroporator (BioRad). Post recovery, cells were plated on LB Lennox agar plates with 50 μ g ml⁻¹ kanamycin.

Hydrogen quantification

Hydrogen was directly quantified from 5 ml cultures grown in sealed Hungate tubes (Pinske and Sargent, 2016). Gas headspace samples were collected using a syringe with Luer lock valve (SGE). Samples were analysed using gas chromatography (Shimadzu GC-2014, capillary column, TCD detector). A hydrogen standard curve was used to quantify sample hydrogen content, this was then normalised to optical density (OD₆₀₀) and culture volume (Pinske and Sargent, 2016).

Western immunoblotting

Proteins samples were first separated by SDS-PAGE using the method of Laemmli (1970) before transfer to nitrocellulose (Dunn, 1986). Nitrocellulose membranes were challenged with an anti-His-HRP antibody (Alpha Diagnostics) and a GeneGnome instrument (SynGene) was used to

visualise immunoreactive bands following addition of ECL reagent (Bio-Rad).

Structure modelling and phylogenetic analysis

Structural modelling of the formate hydrogenlyases complexes was performed using Phyre² predictions of respective subunits (Kelley and Sternberg, 2009). Using Chimera (Pettersen *et al.*, 2004), the X-ray crystal structure of *Thermus thermophilus* Respiratory Complex I (4HEA) and the Cryo-EM structure of membrane-bound hydrogenase (6CFW), the individual FHL-2 subunits were manually assembled into a putative complex organisation for FHL-1 and FHL-2. Phylogenetic analysis of *E. coli* FdhF-like proteins from organisms possessing a Group 4A [NiFe]-hydrogenase utilised the HydDB database (Greening *et al.*, 2015) to collect accession numbers for all [NiFe]-hydrogenase subunits. In each organism, the FdhF orthologs were identified before MUSCLE multiple sequence alignment in Jalview (Waterhouse *et al.*, 2009). Through percentage identity tree generation and manual inspection, the closest FdhF-like proteins in each organism were identified. FigTree (<http://tree.bio.ed.ac.uk/software/figtree>) was used to visualise the finalised phylogenetic trees.

Acknowledgements

This research was funded primarily by the BBSRC through award of a 4-year EASTBIO PhD studentship in 2014 to AJF (#1510231). MA was funded by the European Union Horizon 2020 Research & Innovation Programme under the Marie Skłodowska-Curie Grant agreement #654006. SJC is a Wellcome Trust Senior Research Fellow.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

AJF Was a PhD student who designed experiments, analysed data, prepared figures for publication and wrote the paper. RL and MF were undergraduate project students who performed experiments and analysed data. MA was a Marie Skłodowska-Curie Independent Fellow who supervised the research, performed experiments and analysed data. SJC was a Wellcome Trust Senior Research Fellow who designed the research, supervised the research, analysed data and wrote the paper. FS conceived the project, assembled the research team, designed the research, supervised the research, analysed data and wrote the paper.

References

Andrews, S.C., Berks, B.C., McClay, J., Ambler, A., Quail, M.A., Golby, P., *et al.* (1997) A 12-cistron *Escherichia*

- coli* operon (hyf) encoding a putative proton-translocating formate hydrogenlyase system. *Microbiology*, **143**, 3633–3647.
- Arnoux, P., Ruppelt, C., Oudouhou, F., Lavergne, J., Siponen, M.I., Toci, R., *et al.* (2015) Sulphur shuttling across a chaperone during molybdenum cofactor maturation. *Nature Communications*, **6**, 6148.
- Axley, M.J., Bock, A. and Stadtman, T.C. (1991) Catalytic properties of an *Escherichia coli* formate dehydrogenase mutant in which sulfur replaces selenium. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 8450–8454.
- Babujee, L., Apodaca, J., Balakrishnan, V., Liss, P., Kiley, P.J., Charkowski, A.O., *et al.* (2012) Evolution of the metabolic and regulatory networks associated with oxygen availability in two phytopathogenic enterobacteria. *BMC Genomics*, **13**, 110.
- Bae, S.S., Lee, H.S., Jeon, J.H., Lee, J.H., Kang, S.G. and Kim, T.W. (2015) Enhancing bio-hydrogen production from sodium formate by hyperthermophilic archaeon, *Thermococcus onnurineus* NA1. *Bioprocess and Biosystems Engineering*, **38**, 989–993.
- Bagramyan, K., Vassilian, A., Mnatsakanyan, N. and Trchounian, A. (2001) Participation of hyf-encoded hydrogenase 4 in molecular hydrogen release coupled with proton-potassium exchange in *Escherichia coli*. *Membrane Cell Biology*, **14**, 749–763.
- Batista, A.P., Marreiros, B.C. and Pereira, M.M. (2013) The antiporter-like subunit constituent of the universal adaptor of complex I, group 4 membrane-bound [NiFe]-hydrogenases and related complexes. *Biological Chemistry*, **394**, 659–666.
- Belinky, F., Rogozin, I.B. and Koonin, E.V. (2017) Selection on start codons in prokaryotes and potential compensatory nucleotide substitutions. *Scientific Reports*, **7**, 12422.
- Bell, K.S., Sebahia, M., Pritchard, L., Holden, M.T., Hyman, L.J., Holeva, M.C., *et al.* (2004) Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 11105–11110.
- Bock, A., King, P.W., Blokesch, M. and Posewitz, M.C. (2006) Maturation of hydrogenases. *Advances in Microbial Physiology*, **51**, 1–71.
- Bohm, R., Sauter, M. and Bock, A. (1990) Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogenlyase components. *Molecular Microbiology*, **4**, 231–243.
- Boyington, J.C., Gladyshev, V.N., Khangulov, S.V., Stadtman, T.C. and Sun, P.D. (1997) Crystal structure of formate dehydrogenase H: catalysis involving Mo, molybdopterin, selenocysteine, and an Fe₄S₄ cluster. *Science*, **275**, 1305–1308.
- Chaves, I., Pinheiro, C., Paiva, J.A., Planchon, S., Sergeant, K., Renaut, J., *et al.* (2009) Proteomic evaluation of wound-healing processes in potato (*Solanum tuberosum* L.) tuber tissue. *Proteomics*, **9**, 4154–4175.
- Choi, D.S., Kim, N.H. and Hwang, B.K. (2014) Pepper mitochondrial FORMATE DEHYDROGENASE1 regulates cell death and defense responses against bacterial pathogens. *Plant Physiology*, **166**, 1298–1311.
- Coulthurst, S.J., Lilley, K.S. and Salmond, G.P. (2006) Genetic and proteomic analysis of the role of luxS in the enteric phytopathogen, *Erwinia carotovora*. *Molecular Plant Pathology*, **7**, 31–45.
- Dunn, S.D. (1986) Effects of the modification of transfer buffer composition and the renaturation of proteins in gels on the recognition of proteins on Western blots by monoclonal antibodies. *Analytical Biochemistry*, **157**, 144–153.
- Eitinger, T. and Mandrand-Berthelot, M.A. (2000) Nickel transport systems in microorganisms. *Archives of Microbiology*, **173**, 1–9.
- Finney, A.J. and Sargent, F. (2019) Formate hydrogenlyase: a group 4 [NiFe]-hydrogenase in tandem with a formate dehydrogenase. *Advances in Microbial Physiology*, **74**, 465–486.
- 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 Letters*, **479**, 1–5.
- Gladyshev, V.N., Khangulov, S.V., Axley, M.J. and Stadtman, T.C. (1994) Coordination of selenium to molybdenum in formate dehydrogenase H from *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 7708–7711.
- Greening, C., Biswas, A., Carere, C.R., Jackson, C.J., Taylor, M.C., Stott, M.B., *et al.* (2015) Genomic and metagenomic surveys of hydrogenase distribution indicate H is a widely utilised energy source for microbial growth and survival. *ISME Journal*, **10**, 761–777.
- Hourton-Cabassa, C., Ambard-Bretteville, F., Moreau, F., Davy de Virville, J., Remy, R. and Francs-Small, C.C. (1998) Stress induction of mitochondrial formate dehydrogenase in potato leaves. *Plant Physiology*, **116**, 627–635.
- Hunger, D., Doberenz, C. and Sawers, R.G. (2014) Identification of key residues in the formate channel FocA that control import and export of formate. *Biological Chemistry*, **395**, 813–825.
- Iwadate, Y. and Kato, J.I. (2019) Identification of a formate-dependent uric acid degradation pathway in *Escherichia coli*. *Journal of Bacteriology*, **201**, e00573-18.
- Jack, R.L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T. and Sargent, F. (2004) Coordinating assembly and export of complex bacterial proteins. *EMBO Journal*, **23**, 3962–3972.
- Kaniga, K., Delor, I. and Cornelis, G.R. (1991) A wide-host-range suicide vector for improving reverse genetics in gram-negative bacteria: inactivation of the *blaA* gene of *Yersinia enterocolitica*. *Gene*, **109**, 137–141.
- Kelley, L.A. and Sternberg, M.J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols*, **4**, 363–371.
- Kim, Y.J., Lee, H.S., Kim, E.S., Bae, S.S., Lim, J.K., Matsumi, R., *et al.* (2010) Formate-driven growth coupled with H(2) production. *Nature*, **467**, 352–355.
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
- Lim, J.K., Mayer, F., Kang, S.G. and Muller, V. (2014) Energy conservation by oxidation of formate to carbon dioxide and hydrogen via a sodium ion current in a hyperthermophilic archaeon. *Proceedings of the National*

- Academy of Sciences of the United States of America*, **111**, 11497–11502.
- Lindenstrauß, U. and Pinske, C. (2019) Dissection of the hydrogen metabolism of the enterobacterium *Trabulsiella guamensis*: Identification of a formate-dependent and essential formate hydrogenlyase complex exhibiting phylogenetic similarity to complex I. *Journal of Bacteriology*, **201**, e00160-19.
- Maia, L.B., Moura, J.J. and Moura, I. (2015) Molybdenum and tungsten-dependent formate dehydrogenases. *JBIC Journal of Biological Inorganic Chemistry*, **20**, 287–309.
- Marreiros, B.C., Batista, A.P., Duarte, A.M. and Pereira, M.M. (2013) A missing link between complex I and group 4 membrane-bound [NiFe] hydrogenases. *Biochimica et Biophysica Acta*, **1827**, 198–209.
- Masuda, N. and Church, G.M. (2003) Regulatory network of acid resistance genes in *Escherichia coli*. *Molecular Microbiology*, **48**, 699–712.
- McDowall, J.S., Murphy, B.J., Haumann, M., Palmer, T., Armstrong, F.A. and Sargent, F. (2014) Bacterial formate hydrogenlyase complex. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, E3948–E3956.
- McWhorter, A.C., Haddock, R.L., Nocon, F.A., Steigerwalt, A.G., Brenner, D.J., Aleksic, S., et al. (1991) *Trabulsiella guamensis*, a new genus and species of the family *Enterobacteriaceae* that resembles *Salmonella* subgroups 4 and 5. *Journal of Clinical Microbiology*, **29**, 1480–1485.
- Mnatsakanyan, N., Bagramyan, K. and Trchounian, A. (2004) Hydrogenase 3 but not hydrogenase 4 is major in hydrogen gas production by *Escherichia coli* formate hydrogenlyase at acidic pH and in the presence of external formate. *Cell Biochemistry and Biophysics*, **41**, 357–366.
- Mukherjee, M., Vajpai, M. and Sankararamakrishnan, R. (2017) Anion-selective formate/nitrite transporters: taxonomic distribution, phylogenetic analysis and subfamily-specific conservation pattern in prokaryotes. *BMC Genomics*, **18**, 560.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., et al. (2004) UCSF Chimera – a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, **25**, 1605–1612.
- Pinske, C. and Sargent, F. (2016) Exploring the directionality of *Escherichia coli* formate hydrogenlyase: a membrane-bound enzyme capable of fixing carbon dioxide to organic acid. *MicrobiologyOpen*, **5**, 721–737.
- Pinske, C. and Sawers, R.G. (2016) Anaerobic formate and hydrogen metabolism. *EcoSal Plus*, **7**, ESP-0011-2016.
- Roggiani, M. and Goulian, M. (2015) Chromosome-membrane interactions in bacteria. *Annual Review of Genetics*, **49**, 115–129.
- Rossmann, R., Sawers, G. and Bock, A. (1991) Mechanism of regulation of the formate-hydrogenlyase pathway by oxygen, nitrate, and pH: definition of the formate regulon. *Molecular Microbiology*, **5**, 2807–2814.
- Sargent, F. (2016) The model [NiFe]-hydrogenases of *Escherichia coli*. *Advances in Microbial Physiology*, **68**, 433–507.
- Schut, G.J., Zadvornyy, O., Wu, C.H., Peters, J.W., Boyd, E.S. and Adams, M.W. (2016) The role of geochemistry and energetics in the evolution of modern respiratory complexes from a proton-reducing ancestor. *Biochimica et Biophysica Acta*, **1857**, 958–970.
- Self, W.T., Hasona, A. and Shanmugam, K.T. (2004) Expression and regulation of a silent operon, *hyf*, coding for hydrogenase 4 isoenzyme in *Escherichia coli*. *Journal of Bacteriology*, **186**, 580–587.
- Skibinski, D.A., Golby, P., Chang, Y.S., Sargent, F., Hoffman, R., Harper, R., et al. (2002) Regulation of the hydrogenase-4 operon of *Escherichia coli* by the sigma(54)-dependent transcriptional activators FhlA and HyfR. *Journal of Bacteriology*, **184**, 6642–6653.
- Suppmann, B. and Sawers, G. (1994) Isolation and characterization of hypophosphite-resistant mutants of *Escherichia coli*: identification of the FocA protein, encoded by the *pfl* operon, as a putative formate transporter. *Molecular Microbiology*, **11**, 965–982.
- Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. and Barton, G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.
- Welte, C., Kratzer, C. and Deppenmeier, U. (2010) Involvement of Ech hydrogenase in energy conservation of *Methanosarcina mazei*. *FEBS Journal*, **277**, 3396–3403.
- Wu, L.F., Navarro, C. and Mandrand-Berthelot, M.A. (1991) The *hydC* region contains a multi-cistronic operon (*nik*) involved in nickel transport in *Escherichia coli*. *Gene*, **107**, 37–42.
- Yu, H., Wu, C.H., Schut, G.J., Haja, D.K., Zhao, G., Peters, J.W., et al. (2018) Structure of an ancient respiratory system. *Cell*, **173**(1636–1649), e1616.
- Zinoni, F., Birkmann, A., Leinfelder, W. and Bock, A. (1987) Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon. *Proceedings of the National Academy of Sciences of the United States of America*, **84**, 3156–3160.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.