JBC RESEARCH ARTICLE

# A distinct class of ferredoxin:NADP$^+$ oxidoreductase enzymes driving thermophilic ethanol production

Received for publication, December 13, 2024, and in revised form, May 5, 2025 Published, Papers in Press, May 21, 2025,
https://doi.org/10.1016/j.jbc.2025.110263

**Shu Huang**[1,2], **Syed Muhammad Saad Imran**[3], **Anthony A. Lanahan**[1,2], **Sarah K. Hammer**[1], **Carolyn E. Lubner**[3], **Lee R. Lynd**[1,2], **and Daniel G. Olson**[1,2,*]

*From the* [1]*Thayer School of Engineering at Dartmouth College, Hanover, New Hampshire, USA;* [2]*Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA;* [3]*Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado, USA*

Reviewed by members of the JBC Editorial Board. Edited by Joseph Jez

**Biofuel production from lignocellulosic biomass offers a transformative solution to reduce global fossil fuel dependency. Certain thermophilic anaerobes, including *Clostridium thermocellum*, show promise for renewable ethanol production due to their ability to break down plant material at high temperatures. However, achieving commercially viable ethanol yields has proven challenging despite extensive engineering efforts. Here, we characterized 27 ferredoxin:NADP$^+$ oxidoreductase (Fnor) enzymes for their enzyme activity, nicotinamide cofactor specificity, thermotolerance, and functional expression in *C. thermocellum*. We identified a subset of 10 of these enzymes as a novel class of Fnor enzymes suited for metabolic pathways aimed at high-titer ethanol production. When expressed in engineered *C. thermocellum*, these enzymes increased ethanol production up to 2.2-fold. These findings establish a novel ethanol pathway and provide insights into physiological roles and biotechnological applications of this new class of Fnor enzymes.**

The transportation sector is a significant contributor to the production of carbon dioxide and other greenhouse gasses that cause global warming and climate change. Production of biofuels from lignocellulosic biomass is one of the most promising approaches to decarbonize the transportation sector (1). In the past 2 decades, there have been several unsuccessful attempts to commercialize cellulosic biofuels. A common theme of these efforts is that they started with organisms with a strong native ability to produce ethanol from soluble sugars (*i.e. Saccharomyces cerevisiae* or *Zymomonas mobilis*) and use a combination of thermochemical pretreatment and added fungal enzymes to generate soluble sugars from lignocellulose (2). An alternative approach to lignocellulosic biofuel production is to start with organisms that natively ferment lignocellulose and engineer them for increased biofuel production. This approach, known as consolidated bioprocessing (CBP), avoids many of the costs associated with the traditional approach such as added enzymes and thermochemical pretreatment (3). However, CBP has its own challenges, primarily related to engineering cellulolytic organisms for increased ethanol production.

*Clostridium thermocellum*, a thermophilic anaerobe, is a promising candidate for CBP because of its ability to efficiently degrade various biomass substrates and produce high ethanol yields (4, 5). Efforts to improve ethanol yield and titer in *C. thermocellum* have included deleting competing pathways, adaptive laboratory evolution, and introduction of heterologous genes from *Thermoanaerobacterium saccharolyticum* (6, 7) since *T. saccharolyticum*, although unable to ferment cellulose, can ferment a wide range of C5 sugars (8) and several engineered variants have the ability to produce ethanol at near-theoretical yields (9, 10). Previously, *C. thermocellum* has been engineered to produce ethanol at yields above 80% of the theoretical maximum (11) and titers near 30 g/L (6, 12) under laboratory conditions. However, for commercial application, ethanol titers of 40 to 50 g/L are needed (13).

In *C. thermocellum*, ethanol production from pyruvate proceeds *via* a set of three reactions (Fig. 1). First, pyruvate is converted to acetyl-coA by the pyruvate:ferredoxin oxidoreductase reaction (PFOR[1]). Then, acetyl-coA is converted to acetaldehyde (ALDH reaction), and finally, the acetaldehyde is converted to ethanol (ADH reaction), with both ALDH and ADH reactions catalyzed by a bifunctional enzyme, AdhE (14). However, ethanol production requires electrons in addition to carbon. In the WT organism, the electrons released by the PFOR reaction reduce ferredoxin (Fd), which in turn reduces NAD$^+$ to NADH *via* the RNF reaction (*see below*) (15). NADH, generated both by glycolysis and the RNF reaction, is then used as a reductant by AdhE for consecutive acetyl-CoA and acetaldehyde reduction to ethanol.

However, the presence of NADH-linked ADH activity makes *C. thermocellum* highly sensitive to inhibition by ethanol. Adaptive evolution in the presence of added ethanol consistently results in mutations in the *adhE* gene that disrupt ADH activity (16–18). This has been further confirmed by showing that expression of NADH-linked Adh enzymes (either

[1] We use the following formatting conventions to distinguish between genes, proteins, and enzymatic reactions. Genes are shown in lower case italics (*i.e., gene*). Proteins are shown in title case roman (*i.e.*, Protein), and enzymatic reactions are shown in all caps (*i.e.*, REACTION).

* For correspondence: Daniel G. Olson, Daniel.G.Olson@dartmouth.edu.



© 2025 THE AUTHORS. Published by Elsevier Inc on behalf of American Society for Biochemistry and Molecular Biology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

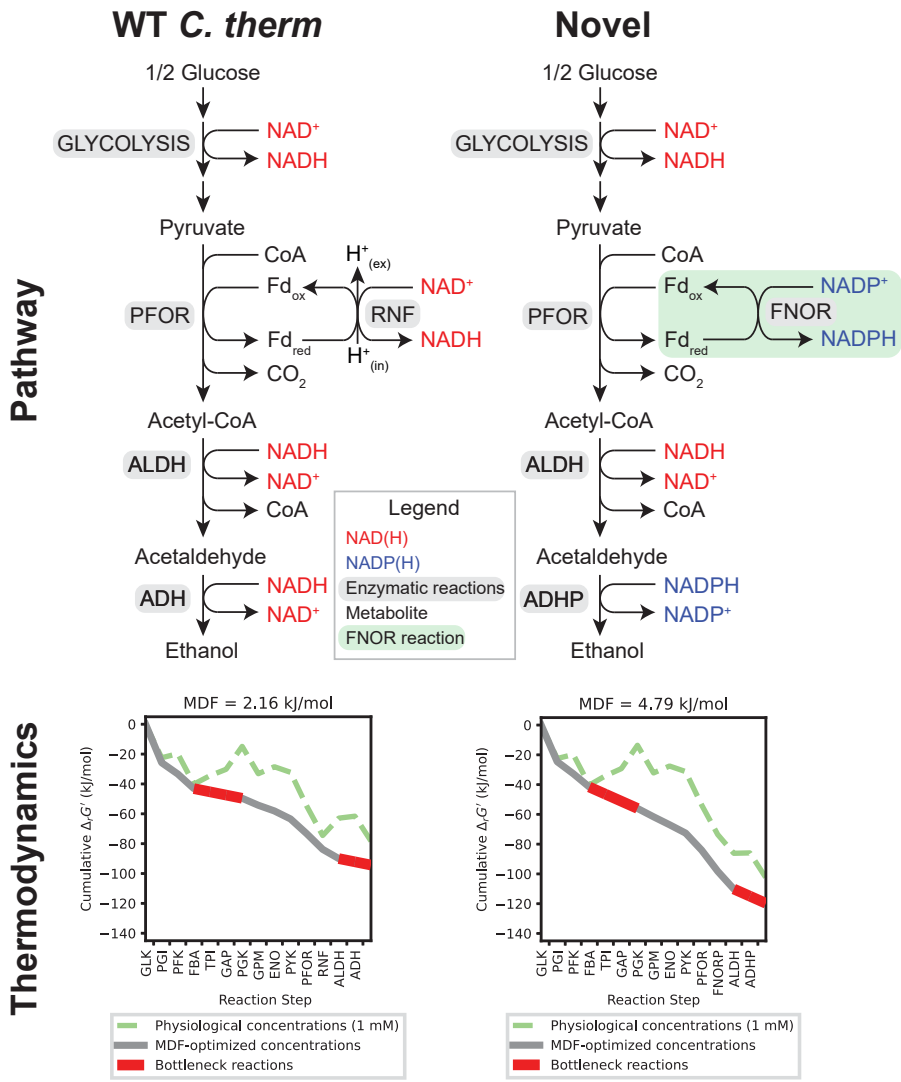



**Figure 1. Comparison of two ethanol production pathways.** The ethanol production pathway in WT *C. thermocellum* is compared with a novel pathway incorporating FNOR (ferredoxin:NADP$^+$ oxidoreductase) and ADHP (NADPH-linked alcohol dehydrogenase) reactions. The *top panels* show the stoichiometry of the different pathways. Nicotinamide adenine dinucleotide cofactors (NADH and NAD$^+$) are highlighted in *red*, and nicotinamide adenine dinucleotide phosphate cofactors (NADPH and NADP$^+$) are highlighted in *blue*. Enzymatic reaction names are shown in *gray boxes*, metabolites are represented by *black letters*, and the FNOR reaction (the central focus of the current work) is highlighted with a *green box*. The *bottom panels* show the negative Gibbs free energy change ($-\Delta_r G'$) of each enzymatic reaction in both pathways when the Max-min Driving Force (MDF) of the pathway is maximized, following the approach described previously (57, 97). The glycolysis reactions depicted include GLK (glucokinase), PGI (glucose-6-phosphate isomerase), PFK (6-phosphofructokinase), FBA (fructose-bisphosphate aldolase), TPI (triosephosphate isomerase), GAP (glyceraldehyde-3-phosphate dehydrogenase), PGK (phosphoglycerate kinase), GPM (phosphoglycerate mutase), ENO (enolase), and PYK (pyruvate kinase).

bifunctional AdhE enzymes (19) or monofunctional enzymes (20)) reduces ethanol tolerance. The mechanism for this appears to be due to a thermodynamic linkage between the ADH reaction and the GAPDH reaction due to use of the same cofactor pair, NADH/NAD$^+$ (21, 22) (more information in the discussion section).

Therefore, one potential approach to increase ethanol titer is to change the cofactor specificity of the ADH reaction from NADH to NADPH, either by protein engineering of AdhE (19) or by expression of a monofunctional enzyme with NADPH-linked ADH activity (*i.e.* ADHP activity) (23). However, this is not compatible with the WT glucose-to-ethanol pathway because no enzymes involved generate NADPH. Therefore, in conjunction with an Adhp enzyme, a ferredoxin:NAD(P)$^+$ oxidoreductase (Fnor) enzyme must be utilized, which can catalyze the following NADPH-linked FNOR reaction:

$$\mathrm{NAD(P)^+ + H^+ + 2\ Fd_{red} \rightleftharpoons NAD(P)H + 2\ Fd_{ox}} \qquad \text{Eq. 1}$$

Several enzymes exhibit FNOR activity, some of which have monofunctional FNOR activity while others couple the catalysis of FNOR reaction to another activity. The two native enzymes in *C. thermocellum* with FNOR-type reactivity are Rnf and Nfn (15). The RNF reaction (originally named from ***rhodobacter* nitrogen fixation** (15, 24)) couples NADH-linked

FNOR activity with $Na^+/H^+$ pumping to generate an electrochemical gradient (25–28):

$$NAD^+ + H^+ + 2\ Fd_{red} + H^+/Na^+_{in} \rightleftharpoons NADH + 2\ Fd_{ox} + H^+/Na^+_{out} \quad \text{Eq. 2}$$

The NFN reaction couples NADPH-linked FNOR activity with transhydrogenation of NADH to NADPH (29):

$$2\ NADP^+ + H^+ + 2\ Fd_{red} + NADH \rightleftharpoons 2\ NADPH + 2\ Fd_{ox} + NAD^+ \quad \text{Eq. 3}$$

However, neither the RNF reaction nor the NFN reaction are suitable for the pathway described in Figure 1, as their reaction stoichiometry does not balance the cofactors appropriately. For the novel pathway, one equivalent each of NADH (for ALDH) and NADPH (for ADHP) are required, and the oxidized cofactors ($NAD^+$ and $NADP^+$) must be continuously re-reduced to keep the pathway functional. ALDH, together with glycolysis and/or RNF, can continuously cycle the $NAD^+$/NADH cofactor pair. However, neither glycolysis nor Rnf is NADPH-linked, so $NADP^+$/NADPH pair cannot be cycled and ADHP cannot function. Nfn does produce NADPH, and it can cycle $NADP^+$/NADPH cofactor pair with ADHP. However, the NFN reaction generates two equivalents of NADPH while ADHP requires only one. Moreover, the NFN reaction also consumes one equivalent of NADH generated by glycolysis, which decreases the amount of NADH available for the ALDH reaction to continue. The imbalanced reactant/product stoichiometries described above can prevent ethanol production due to insufficient flux of intermediates or correct cofactor through specific steps of the pathway.

A promising alternative is Cac_0764, an NADPH-linked Fnor enzyme recently identified in the mesophilic bacterium *Clostridium acetobutylicum* (30), which has the exact reaction stoichiometry required (Equation 1) to provide one equivalent of NADPH to ADHP. This FNOR and ADHP combination can cycle the $NADP^+$/NADPH cofactor pair, and therefore, together with the glycolysis and ALDH reactions, all cofactors and their stoichiometries can be appropriately balanced in the novel pathway by the introduction of this Fnor. However, key characteristics of this novel Fnor relevant to its incorporation in CBP workflows such as thermostability have not yet been assessed.

Thus, the primary goal of this work was to identify thermostable Fnor enzymes, characterize their enzyme activity, and demonstrate their ability to be functionally expressed in *C. thermocellum*, to enable a novel ethanol production pathway not present in the WT organism.

## Results

### Identification of Fnor candidates

For our initial characterization, we identified candidate Fnor enzymes using the following criteria. 1) enzymes that had been previously characterized for FNOR-type activity. This included Pfu_06615 and Pfu_09610 from *Pyrococcus furiosus*, known as NfnI-B and NfnII-B, respectively (31, 32), and Sov_03740 from *Sporomusa ovata* (StnC) (33), along with Telo_06135 from *Thermosynechococcus vestitus* (Fnr) (34, 35) and Hth_1218 from *Hydrogenobacter thermophilus* (36) due to their documented enzymatic activities and roles in ferredoxin metabolism. Additionally, we incorporated Tsac_2085 and Tsac_2086 from *T. saccharolyticum* (NfnB and NfnA), as well as another NfnB homolog Tma_1681 from *Thermotoga maritima*. Tsac_1705 was also included, previously identified as a putative NADH-linked Fnor in *T. saccharolyticum* (37). 2) enzymes with high similarity to Cac_0764 at the amino acid level from thermophilic organisms. This set was identified by searching the JGI IMG database (38) for genomes annotated as either "thermophilic" or "hyperthermophilic" and then performing a BLAST search (39) within those genomes for similarity to Cac_0764. This resulted in 93 hits. After eliminating highly similar sequences, nine were chosen for synthesis. 3) an additional set of enzymes was identified by relaxing the thermophilic constraint (*i.e.* by searching through all of the genomes in the JGI IMG database) and searching for those with higher sequence identity than those from the thermophilic set. In total, across the three different approaches, 27 candidates were identified (Fig. 2). These candidates were codon harmonized for expression in *C. thermocellum* using the JGI BOOST pipeline (40).

### Presence of two types of [4Fe-4S] and FAD-dependent Fnors revealed by sequence alignments

When aligned, most of the 27 putative Fnor sequences group into two phylogenetic clusters, clusters 1 and 2 (Fig. 2). One protein (Sov_03740) shows distant homology to both clusters. The defining feature of both phylogenetic clusters is the presence of Pyr_redox_2 domain in all proteins and Fer4_20 domain in all proteins except one. Pyr_redox_2 consists of a pyridine nucleotide (NAD(P) (H))-binding domain nested within an FAD-binding domain (41). Fer4_20 (42) is the characteristic domain from dihydropyrimidine dehydrogenase that binds two [4Fe-4S] clusters, one with only cysteines and the other with cysteines and a noncysteine (site-differentiated) residue (43). A smaller group, cluster 3, also emerged which comprises conventional plant-type Fnors (better known as Fnrs) (44), with domains that bind FAD and NAD(P) (H). Cluster 3 includes Sole_110801208, the well-characterized Fnr from spinach (45, 46) (*Spinacia oleracea*; which was only included for sequence alignments and not assayed for activity), Telo_06135, and Hth_1218, which have known NADPH-linked FNOR and/or diaphorase activities (36, 45).

Cluster 1 contains five known NfnB proteins (Table S1), of which Pfu_06615 (*P. furiosus* NfnI-B) is the best characterized biochemically, biophysically, and structurally (32, 47, 48) (Fig. 3*A*). Therefore, its sequence is used as reference for all alignments and residue numbering. Overall, there is a lot of homology within and between the cluster 1 and cluster 2 proteins (Figs. S1–S3). All residues which, in *Pf*NfnI-B (Fig. 3*B*), interact with the FAD and bind the two [4Fe-4S] (including Glu126, the site-differentiated ligand for one Fe in the [4Fe-4S] cluster proximal to the FAD) are fully conserved in cluster 1 proteins (Figs. 3*C*, S2). Residues homologous to

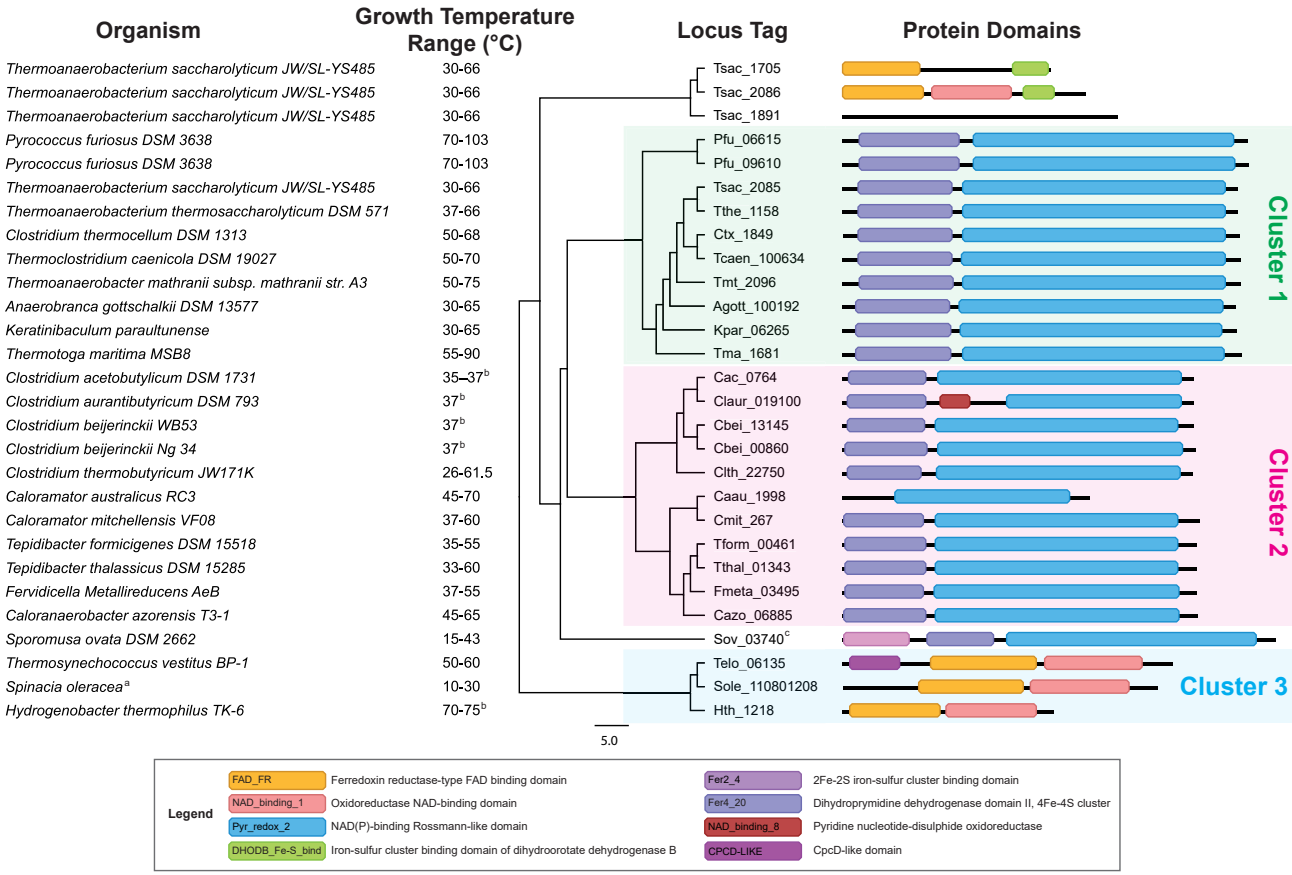



**Figure 2. Similarity analysis of Fnor amino acid sequences.** For each candidate Fnor protein, the source organism and its growth temperature range are indicated. Analysis of functionally important PFAM and PROSITE domains (98) revealed substantial structural similarity among most of the Fnor proteins. [a] *Spinacia oleracea* Fnr Sole_110801208 was included only for sequence analysis. [b] Growth temperature range was not found in the literature; therefore, only the optimal growth temperature was listed. [c] The rest of the PFAM domains are not shown.

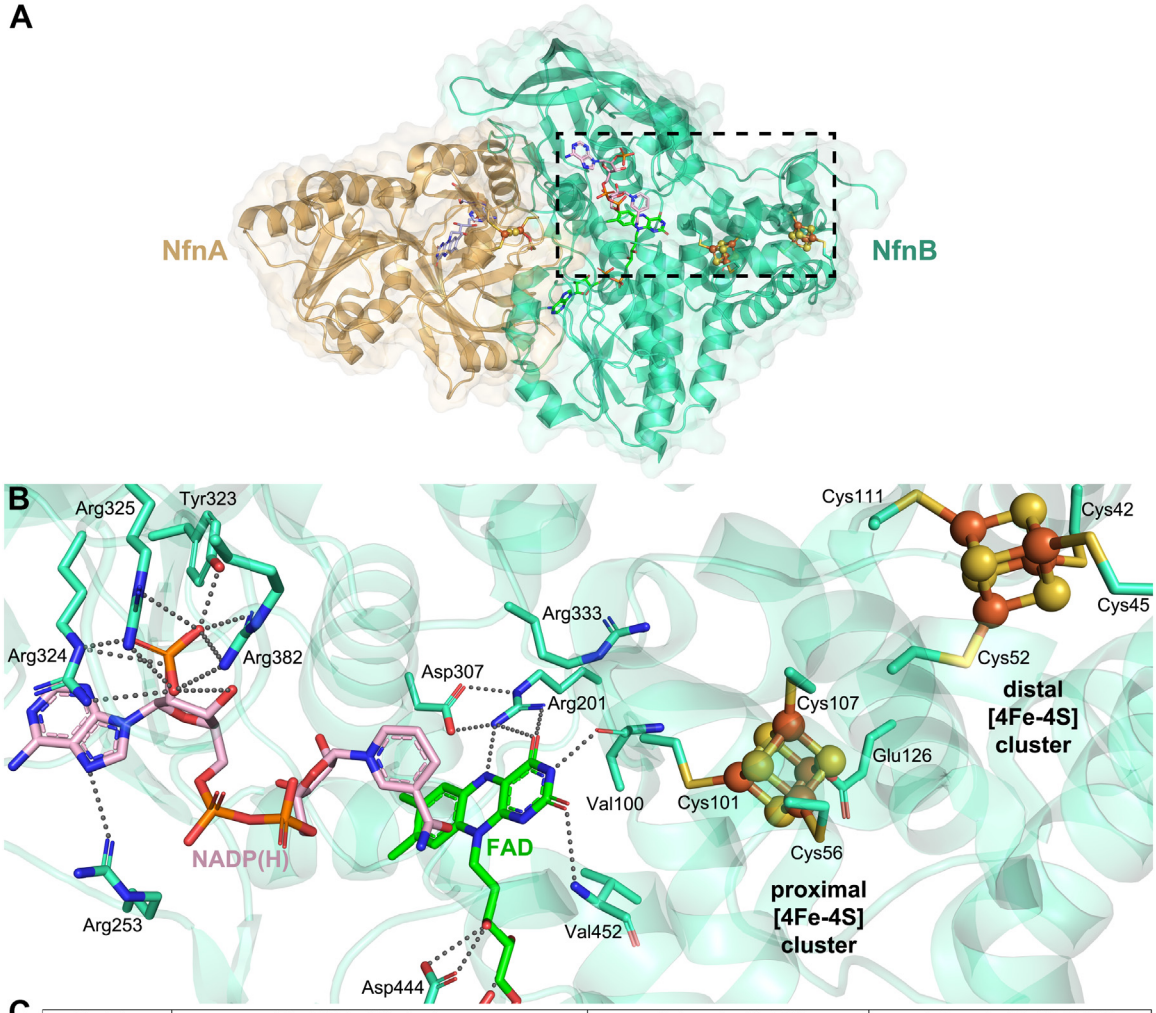
those which interact with NADP(H) in *Pf*NfnI-B show slightly more variability (especially Arg253 which interacts with the adenine moiety of NADP(H)) but are still highly conserved. Ten out of eleven proteins in cluster 2 also follow the same pattern, with high homology to each other and the reference *Pf*NfnI-B protein sequence (Figs. 3*C*, S3). The exception is Caau_1998, which is missing the Fer4_20 domain and hence lacks all cluster-binding residues. More variability is observed at positions Arg253, Arg325, and Arg382, but the residues still largely remain positively charged (Lys) and/or polar (Asn, Glu, Tyr). The presence of Fer4_20 and Pyr_redox_2 domains and strong conservation of residues which interact with NADP(H), FAD, and [4Fe-4S] clusters indicate that, like the known NfnB proteins, proteins in both phylogenetic clusters are involved in electron transfer to/from NADP(H), which is mediated by FAD and, except for one protein, two [4Fe-4S] clusters.

### *Enzyme assay with purified enzymes*

To determine the enzyme activities and the nicotinamide cofactor specificities of the Fnor enzymes, we performed benzyl viologen (BV) reduction assay and ferredoxin oxidation assay with enzymes expressed and purified from *Escherichia coli*. The BV reduction assay is an effective screening tool because it is simple to set up, can be performed with commercially available reagents, and allows for easy observation of the reaction through a visible color change. However, the broad reactivity of BV with a wide range of enzymes such as NADP-dependent formate dehydrogenase (49) and thioredoxin reductase (50) can be a drawback, as it may not accurately reflect the enzyme's native activity with specific electron carriers like ferredoxin. Additionally, the BV reaction in the assay does not proceed in the physiological direction expected within the cell during ethanol formation. Therefore, the BV assay was used for initial screening. Enzymes exhibiting BV activity were then assayed using the ferredoxin-linked assay.

With the BV assay, NADH-linked activity was below 10 U/mg among all proteins tested (Fig. 4*A*). NADPH-linked activity was generally higher, with average activities ranging from 33 U/mg protein (Fmeta_03495) to 242 U/mg protein (Claur_019100) (Fig. 4*C*).

For the ferredoxin oxidation assay, we coupled the FNOR reaction with the PFOR reaction to regenerate reduced ferredoxin. Several Fnor enzymes—Claur_019100, Telo_06135, Hth_1218, Cazo_06885, Clth_22750, Cac_0764, Cmit_267, Tform_00461, Fmeta_03495, and Tthal_01343—exhibited NADPH-linked activities higher than that of Ctx_1849, the native *C. thermocellum* NfnB (Fig. 4*D*), confirming that these enzymes transfer electrons from ferredoxin to $NADP^+$. Among




C

| | Proteins | Bind [4Fe-4S] clusters | | | | | | | | Interact with FAD | | | | | Interact with NADP(H) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | C42 | C45 | C52 | C56 | C101 | C107 | C111 | E126 | V100 | R201 | D307 | D444 | V452 | R253 | Y323 | R324 | R325 | R382 |
| Cluster 1 | Pfu_06615 | C42 | C45 | C52 | C56 | C101 | C107 | C111 | E126 | V100 | R201 | D307 | D444 | V452 | R253 | Y323 | R324 | R325 | R382 |
| | Pfu_09610 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | ✓ | K |
| | Tsac_2085 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | ✓ | ✓ |
| | Tthe_1158 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | ✓ | ✓ |
| | Ctx_1849 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | S | ✓ | ✓ | ✓ | ✓ |
| | Tcaen_100634 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | M | ✓ | ✓ | ✓ | ✓ |
| | Tmt_2096 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | ✓ | ✓ |
| | Agott_100192 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | Y | ✓ | ✓ | ✓ | ✓ |
| | Kpar_06265 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | ✓ | ✓ |
| | Tma_1681 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | ✓ | ✓ |
| Cluster 2 | Cac_0764 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | K | K |
| | Claur_019100 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | K | K |
| | Cbei_13145 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | K | K |
| | Cbei_00860 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | K |
| | Clth_22750 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | K | K |
| | Caau_1998 | × | × | × | × | × | × | × | × | × | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | ✓ | K |
| | Cmit_267 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K |
| | Tform_00461 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | N | ✓ | ✓ | K | Y |
| | Tthal_01343 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | N | ✓ | ✓ | K | N |
| | Fmeta_03495 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | N | ✓ | ✓ | K | K |
| | Cazo_06885 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | K | ✓ | ✓ | ✓ | E |

Legend: ✓ fully conserved residue; × missing residue; residue of a similar chemical nature (polar/non-polar); residue of a dissimilar chemical nature (polar/non-polar)

**Figure 3. Structural analysis of PfNfnI highlighting substrate/cofactor binding and residue conservation across protein clusters.** *A*, structure of *Pf*NfnI (PDB ID: 5JCA (32)), which consists of the small subunit NfnA (*orange*) and large subunit NfnB (*green*). The *dashed box* indicates the regions binding the substrates and cofactors shown in detail in (*B*). *B*, key residues in *Pf*NfnI-B (Pfu_06615) involved in binding/interacting with the substrate NADP(H) and the cofactors FAD and [4Fe-4S] clusters. Arg333 is also shown, which does not form observed contacts with cofactors, and its hypothesized role is discussed later in the text (see also Fig. 8). *C*, conservation of important residues in cluster 1 and cluster 2 proteins, with *Pf*NfnI-B (Pfu_06615) protein sequence and numbering used as reference.

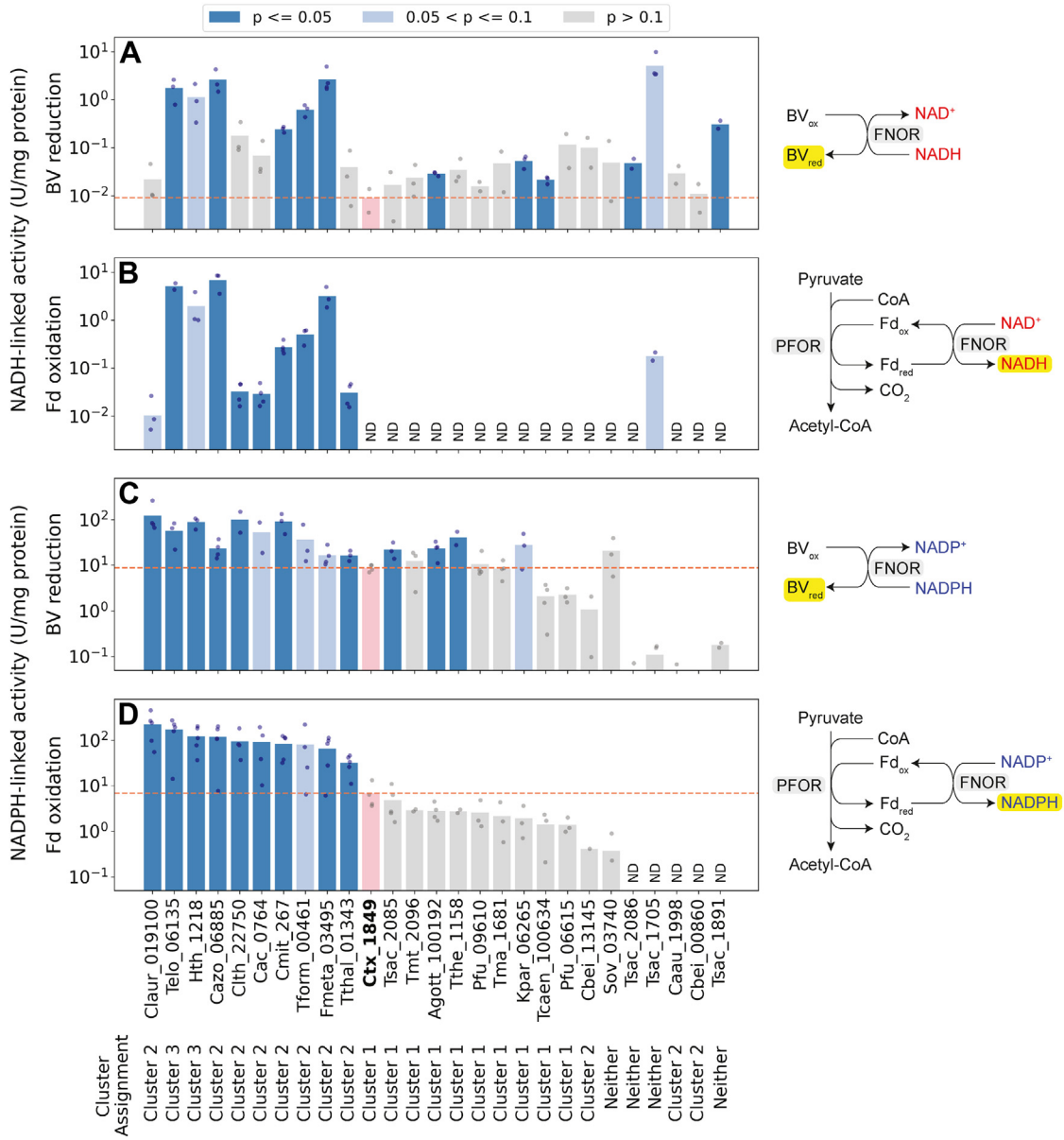



**Figure 4. Enzymatic activities of the Fnor enzymes with various cofactors and electron acceptors/donors.** *A*, NADH-linked BV reduction reactivity, (*B*) NADH-linked Fd oxidation reactivity, (*C*) NADPH-linked BV reduction reactivity, (*D*) NADPH-linked Fd oxidation reactivity. Reaction schemes are provided on the *right*, with the metabolites being measured highlighted in *yellow*. The *arrows* indicate the direction of the reactions. The control is Ctx_1849, a native *C. thermocellum* NfnB, and its activities were highlighted in *pink*. Each point represents an individual biological replicate, and *bars* represent the mean value of all replicates. A one-sided *t* test was performed to determine if the mean activity of each enzyme is greater than that of the Ctx_1849 control (or 0 for NADH-linked ferredoxin oxidation reaction). The color of the *bars* indicates statistical significance of increased FNOR activity, with *dark blue* indicating the most significant increase compared to the control, *light blue* indicating a less significant increase, and *gray* indicating no significant difference or decreased activity. The *orange dashed line* represents the activity level of Ctx_1849. "ND" signifies not detected. Enzymes are arranged in the order of their NADPH-linked ferredoxin oxidation activity levels. The phylogenetic cluster assignment is indicated below each Fnor.

them, Telo_06135, Hth_1218, Cazo_06885, Tform_00461, and Fmeta_03495 also demonstrated NADH-linked activities, although these activities were at least 10 fold lower than their NADPH-linked activities (Fig. 4*B*).

### *Thermostability*

To function as part of an ethanol production pathway in *C. thermocellum*, the enzymes must withstand the growth temperature of 55 °C. To assess the thermostability of the Fnor candidates, we measured their Fd oxidation activity before and after an 80-min incubation at 55 °C. We tested the highly active enzymes identified in previous assays: Claur_019100, Telo_06135, Hth_1218, Cazo_06885, Clth_22750, Cac_0764, Cmit_267, Tform_00461, Fmeta_03495, and Tthal_01343. These enzymes were selected not only for their high activity but also for their favorable expression characteristics and ability to be purified to relatively high purity using a single-step histidine tag purification method (Fig. S8). Additionally, we

included Ctx_1849 and Tsac_2085, the native NfnB enzymes from *C. thermocellum* and *T. saccharolyticum*, respectively, as positive controls.

Several enzymes (Claur_019100, Telo_06135, Hth_1218, Cazo_06885, Ctx_1849, Tsac_2085) retained their activity relatively well after incubation (Fig. 5). The thermostabilities of the Fnor enzymes were largely—but not always—consistent with the growth temperature of the host organisms. As expected due to its mesophilic origin, Cac_0764 showed a significant decrease in activity after incubation. Claur_019100, despite originating from the mesophile *Clostridium aurantibutyricum*, exhibited high thermostability. Conversely, some enzymes from thermophiles did not maintain their activities after incubation. This could be due to suboptimal buffer conditions (for example, the presence of 250 mM imidazole from the elution buffer), which may destabilize the enzymes even at temperatures they would typically tolerate in their native environment. Another possible reason is that the presence of the His-tag might decrease the thermostability of the protein (51).

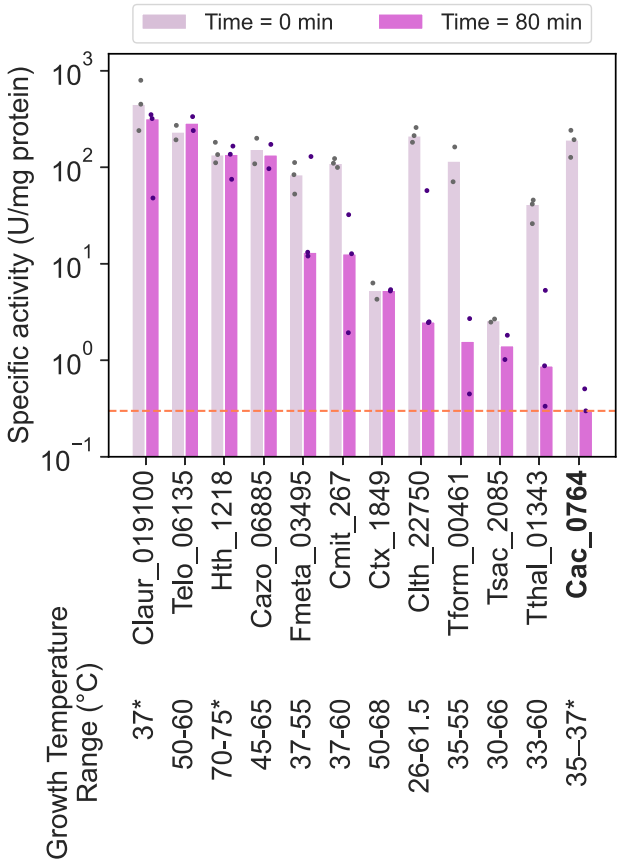



**Figure 5. Thermostability of purified Fnor enzymes.** Thermostability of the purified enzymes was assessed by measuring their enzymatic Fd oxidation activities before and after an 80-min incubation at 55 °C. Each point represents an individual biological replicate, with bars indicating the median value across all replicates (2–3 replicates per sample). The Cac_0764 enzyme is used as a control due to its lack of thermostability. The *orange dashed line* denotes the level of activity after incubation for the control. The enzymes are plotted in the order based on the activity levels following the incubation. *: Growth temperature range was not found in the literature; therefore, only the optimal growth temperature was listed.

### *Functional expression in* C. thermocellum

After measuring the thermostability of our Fnor candidates, we cloned them into expression vectors and transformed the plasmids into *C. thermocellum* strain LL1784. LL1784 is an engineered strain in which its native ethanol production pathway has been replaced with that of *T. saccharolyticum*. In this strain, we deleted *adhE*, native *nfnAB*, and *rnf* genes from the parent strain, LL1592 (52), which led to a significant reduction in carbon flux through glycolysis and a drop in ethanol production from 311.8 mM to 53.8 mM during fermentation with 50 g/L cellobiose (Fig. S6). The ADH and ALDH activities of LL1592 and LL1784 can be found in Figure S7.

Additionally, this strain expresses *T. saccharolyticum* AdhA, an NADPH-dependent alcohol dehydrogenase that enables a stoichiometrically balanced pathway when coupled with an NADPH-dependent Fnor. These characteristics make LL1784 an ideal strain for evaluating the impact of expressing a functional Fnor on ethanol production.

To determine if each Fnor enzyme was functionally expressed, we measured the NADPH-linked FNOR activity by performing the ferredoxin oxidation assay with transformed *C. thermocellum* cell lysates. The enzyme assays showed an increase in Fnor activity in the cell lysates of the strains expressing Cazo_06885, Tform_00461, Fmeta_03495, Claur_019100, and Tsac_2085, compared to the empty vector negative control based on the one-side statistical test (Fig. 6).

For some Fnor enzymes that failed the statistical test, we observed significant colony-to-colony variation in activity. Suspecting that mutations might have occurred, which could lead to a loss in enzyme function, we performed PCR amplification on all transformants to amplify the expressed *fnor* genes and their flanking regions. Sequencing analysis of purified PCR products revealed that many of the *fnor* genes had undergone truncations or deletions, causing a partial decrease or complete elimination of FNOR activity (Fig. 6). The presence of inactivating mutations suggests that *fnor* expression may interfere with growth.

### *Effect on fermentation*

Having identified a problem with genetic instability of our *fnor* expression plasmids, we sought to minimize the effects of this problem by reducing the number of cell divisions between transformation and measurement of fermentation products. To achieve this, we repeated the transformation and inoculated colonies directly from the initial transformation plates (CTFUD-rich medium) into liquid culture on defined medium (MTC) without prior subculturing. Subsequently, we quantified the fermentation products through HPLC at the end of the fermentation to evaluate the effect of expression of Fnor enzymes on fermentation products. At several points during the fermentation, PCR amplification and sequencing were again performed to confirm the integrity of the transformed *fnor* genes. The sequencing results revealed that a much smaller portion of the strains were mutated compared to the previous experiment, as expected (Fig. 7).

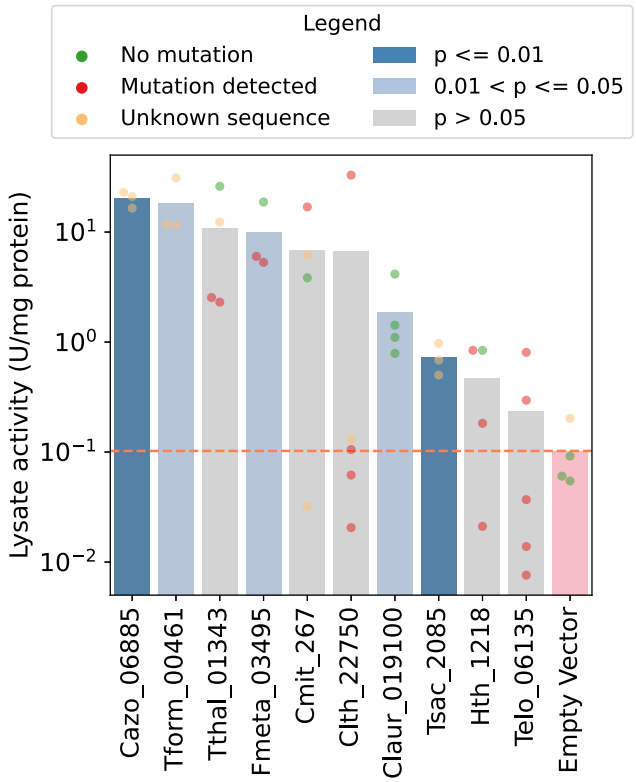



**Figure 6. Functional expression of Fnor enzymes in *C. thermocellum*.** Fnor enzymes were expressed in *C. thermocellum* strain where most of the native Fnors were deleted (LL1784). Each point represents the lysate activity of an individual colony, and *bars* represent the mean of values from individual colonies. The control is strain LL1784 transformed with plasmid pDGO143 (empty vector control). A one-sided *t* test was performed to determine if the mean activity of each sample is greater than that of the control. The color of the *bars* indicates statistical significance of increased FNOR activity, with *dark blue* indicating the most significant increase compared to the control, *light blue* indicating a less significant increase, and *gray* indicating no significant difference. *p*-value cutoffs are shown in the graphical legend. The *orange dashed line* represents the level of lysate activity for the control. Enzymes are arranged in the order of their lysate activities. The points are colored *red* if mutation is detected, *green* if there is no mutation, and *yellow* if the status is unknown.

Colonies expressing Hth_1218 produced an average of 56.7 ± 4.7 mM ethanol, representing a 2.2-fold increase compared to the empty vector control (which produced an average of 25.7 mM ethanol). Colonies expressing Claur_019100 produced an average of 46.3 ± 10 mM ethanol, corresponding to a 1.8-fold increase, with a slightly larger variability than Hth_1218. Two of the Tthal_01343-expressing colonies exhibited a significant increase in ethanol production, with one producing 71 mM and the other 81.6 mM ethanol. However, we detected a partial gene deletion in one and insertions in the terminator region in the other, and it remains unclear how these mutations contribute to the observed increase in ethanol production (Sequencing data can be found in Data S2).

In our analysis of metabolic flux, we use three-carbon (C3) equivalents for accounting purposes, as we have done previously (23, 52–54). This approach simplifies carbon balance calculations in two important ways. First, by treating all carbon-containing molecules in terms of C3 units, we eliminate the need to repeatedly adjust for these conversion ratios across different steps of metabolism. Second, it is a more transparent way to account for flux to $CO_2$, since it avoids the need to estimate $CO_2$ production (*i.e.* by treating ethanol as a C3 equivalent, since ethanol is always produced from the C3 compound pyruvate).

Most of the Fnor-expressing strains had lower cellobiose uptake than the empty vector control, with the exception of Cac_0764 and Tsac_2085 (Fig. 7*A*). This reduced uptake likely explains the reduced amount of cellobiose entering glycolysis observed for Cazo_06885, Tthal_01343, Tform_00461, and Cmit_267 (Fig. 7*E*). Acetate production, as well as pyruvate and lactate levels (not shown), were similar across the Fnor-expressing strains, except for Tthal_01343, indicating that the increase in ethanol production is not derived from these competing fermentation products (Fig. 7*B*).

"Missing carbon" refers to the carbon entering glycolysis that cannot be accounted for in measured products (pyruvate, lactate, acetate, or ethanol). When examining the missing C3 carbon, calculated as the difference between the C3 equivalents entering glycolysis and the sum of the known C3 products, we observed a significant decrease in missing carbon compared to the empty vector control in all strains except Tsac_2085 (Fig. 7*F*), which suggests that previously unaccounted-for C3 intermediates are now being redirected into ethanol production, explaining the observed increase in ethanol concentration.

## Discussion

### *The novelty of the proposed pathway*

We have been particularly interested in exploring the metabolic pathway of *T. saccharolyticum* due to its high ethanol production yield and titer, with the goal of transferring this capability to *C. thermocellum* to achieve a similar effect. Shaw *et al.* (55) first identified both NADH- and NADPH-linked FNOR activity in *T. saccharolyticum* cell lysates using BV assay and proposed that the ethanol production pathway involves a monofunctional Fnor enzyme. However, they did not specify whether FNOR, ALDH, or ADH reactions were linked to either NADH or NADPH and did not measure ferredoxin-linked activity directly.

After discovery of the NfnAB complex and associated electron-bifurcating stoichiometry by Wang *et al.* in *Clostridium kluyveri* (29), we identified *Tsac_2085-2086* coding sequence as the homologous protein complex in *T. saccharolyticum* (56). However, the stoichiometry of the NFN reaction (Equation 3) requires NADPH-linked activity for the final two steps in ethanol production (acetyl-CoA to acetaldehyde and then acetaldehyde to ethanol) to achieve high yields. Although there are many options for NADPH-linked Adh enzymes, there are no known Aldh enzymes with NADPH-linked activity, making the NFN reaction stoichiometry incompatible with the final two steps of ethanol biosynthesis.

Therefore, we kept looking for a monofunctional NADH-linked Fnor enzyme that could account for the

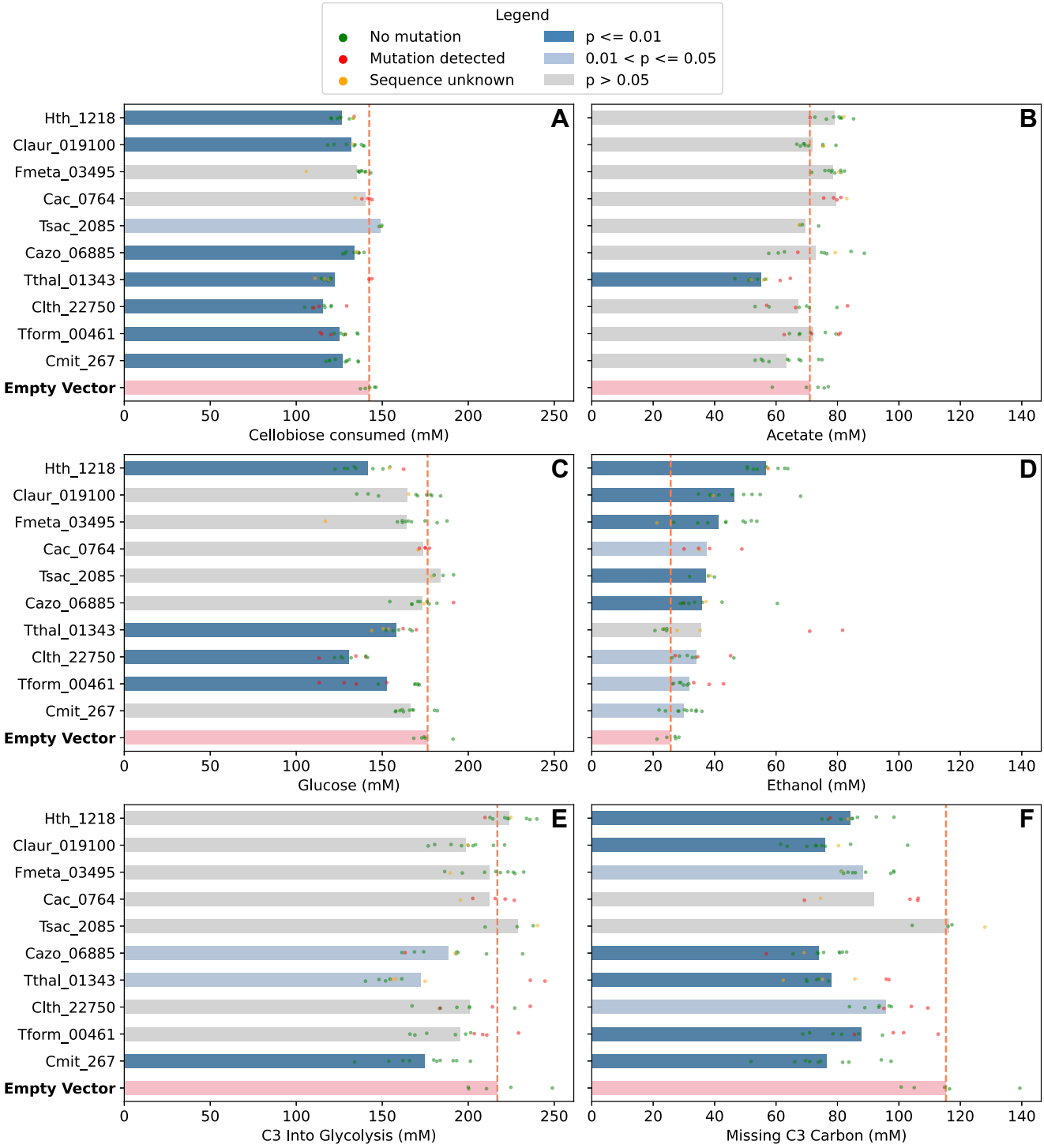



**Figure 7. Impact of Fnor expression in *C. thermocellum* on fermentation profiles.** *A*, cellobiose remaining in the fermentation broth, (*B*) acetate production, (*C*) glucose remaining in the fermentation broth, (*D*) ethanol production, (*E*) C3 equivalents of glucose entering the glycolysis, (*F*) amount of missing C3 carbon, calculated as the difference between the C3 equivalents of glucose entering glycolysis and the sum of C3 equivalents of acetate, ethanol, pyruvate, and lactate. Each data point represents the concentration for an individual colony, with *bars* indicating the mean values across colonies. Data points are colored *red* if a mutation is detected, *green* if no mutation is present, and *yellow* if the mutation status is unknown. A two-sided *t* test was conducted to assess whether the mean concentration of each Fnor-expressing strain significantly differs from that of the empty vector control. The color of the bars indicates statistical significance with *dark blue* indicating the most significant differences from the control, *light blue* indicating less significant differences, and *gray* indicating no significant differences. *p*-value cutoffs are shown in the graphical legend. The *orange dashed line* represents the concentrations for the empty vector control. Enzymes are arranged in the order of their ethanol production.

observed high ethanol yields (9, 10, 56) in strains of *T. saccharolyticum* engineered for homo-ethanol production. Tian *et al.* (37) identified Tsac_1705 as a monofunctional NADH-linked Fnor enzyme and suggested that the ethanol production pathway involves Tsac_1705 and NADH-linked ALDH and ADH reactions. However, the extremely low

activity we measured for this protein (Fig. 4), its annotation as a dihydroorotate dehydrogenase subunit (*see below*), and presence within the pyrimidine biosynthesis operon of *T. saccharolyticum* suggest that its FNOR activity may not be physiologically relevant.

Recently, a new type of Fnor enzyme was identified in *C. acetobutylicum*, which allows electron transfer from ferredoxin to $NADP^+$ for butanol production (30). We realized that the stoichiometry of this enzyme was a good fit for the requirements of ethanol production in thermophilic bacteria—in particular, the need for compatibility with NADPH-linked alcohol dehydrogenase reactions (19). Building on these findings (19, 30), we identified thermophilic Cac_0764-like Fnor enzymes and showed that they could enable ethanol production as part of a novel metabolic pathway.

### Theoretical analysis of metabolic pathways using Fnor

The most straightforward argument for why NADPH-linked FNOR activity (Equation 1) is necessary is based on stoichiometry. If we accept that NADH-linked ADH activity is incompatible with high ethanol tolerance (which is generally needed for high titer) in anaerobic bacteria (20), then NADPH-linked ADH activity is required for ethanol production. A source of NADPH is needed, which is provided by the NADPH-linked FNOR reaction.

At a deeper level, we can use thermodynamic analysis to understand why this might be so. For a given enzymatic reaction, it can only proceed in the forward direction if its free energy change ($\Delta rG'$) is negative. The free energy change is partially determined by the nature of the chemical reaction ($\Delta rG'^{\circ}$) and partly by the concentration of substrates and products. For a metabolic pathway, all of the reactions in that pathway must have a negative free energy change. Cellular metabolism adjusts the concentrations of these species to allow pathway operation, but the thermodynamic landscape provides constraints on the allowable concentration ranges, and these can differ depending on the metabolic pathway (57).

This kind of analysis can be turned into an optimization problem using the max-min driving force (MDF) framework (57). This framework tries to find the optimal set of metabolite concentrations to maximize the thermodynamic driving force of the least favorable metabolic reaction, subject to constraints on maximum and minimum concentrations for each metabolite, and therefore enables objective ranking of pathways by the extent to which their flux is constrained by thermodynamic driving force (57). In general, pathways with a higher MDF value will be less thermodynamically constrained as product titers increase.

For both the WT pathway and the novel pathway we proposed, we calculated MDF values in the presence of 2 M (*i.e.* 92 g/L) ethanol, which is at the upper end of titers we expect to be necessary for commercial application. The novel pathway we proposed, incorporating NADPH-linked FNOR and ADH reactions, has a higher MDF score (4.79 kJ/mol) than the *C. thermocellum* WT pathway (2.16 kJ/mol) (Fig. 1). The reason for this increased MDF value is that changing the cofactor specificity of the ADH reaction from NADH to NADPH breaks a thermodynamic coupling between the ADH and GAPDH reactions (this coupling is only present if both use the same cofactor pair, $NADH/NAD^+$). To date, this kind of thermodynamic analysis has mainly been used to explain features found in naturally occurring pathways (57, 58), but additional experimentation with the pathway described in this work provides new opportunities to test the applicability of MDF analysis to improving product titer.

### Growth toxicity of Fnor enzymes

As noted, during serial transfers, strains expressing Fnor enzymes frequently develop mutations in the *fnor* gene, leading to its partial or complete inactivation. This may occur because the Fnor is not essential for growth or survival under the selective pressure, favoring mutations that reduce or eliminate its expression. Additionally, if the downstream AdhA enzyme, which utilizes NADPH, is not sufficiently efficient or its expression levels are not well-matched to Fnor, NADPH may accumulate. This imbalance could create stress conditions that cells might try to mitigate by inactivating the *fnor* gene through mutations.

### Differences between cluster 1 and 2 proteins in terms of protein sequences and genomic contexts

Despite the similarity in the sequences of proteins, there are differences between phylogenetic clusters 1 and 2, which are likely to be functionally significant. The most prominent discriminant is the length of N-terminus in each cluster. Cluster 1 proteins have N-termini that are 25 to 31 residues longer than those of cluster 2 proteins (Figs. 8*A*, S1). While there is variability in the conservation of individual residues, all cluster 1 proteins have an extended N terminus. In the *Pf*NfnI crystal structure (32), there are specific hydrogen bonding and ionic contacts between the N-terminal residues and the rest of the *Pf*NfnI-B protein, especially the helices that coordinate the two [4Fe-4S] clusters (Fig. 8*B*). While the function of this extended N terminus is unknown, a similarly long and conserved N terminus has also been observed in the electron transfer protein BchL (59). The N terminus of BchL is thought to gate electron transfer from BchL unless the protein binds its substrate and potentially protects against oxidative damage to its [4Fe-4S] cluster. N-termini of cluster 1 proteins could be similarly protective—especially to the distal [4Fe-4S] cluster more exposed to the protein surface—and provide specificity both in protein–protein interactions and electron transfer involving particular ferredoxins.

Additionally, we found a distinct change in the protein environment around FAD between the two clusters. In all crystal structures of Nfn, in addition to Arg201 (which is conserved in both clusters), Arg333 is observed as well (Figs. 8*C*, S1–S3, S5). The latter arginine is the only other positively charged residue around the isoalloxazine moiety of FAD. Together with Arg201, Arg333 could be important for the generation and stabilization of the highly reactive anionic semiquinone intermediate observed during and essential to

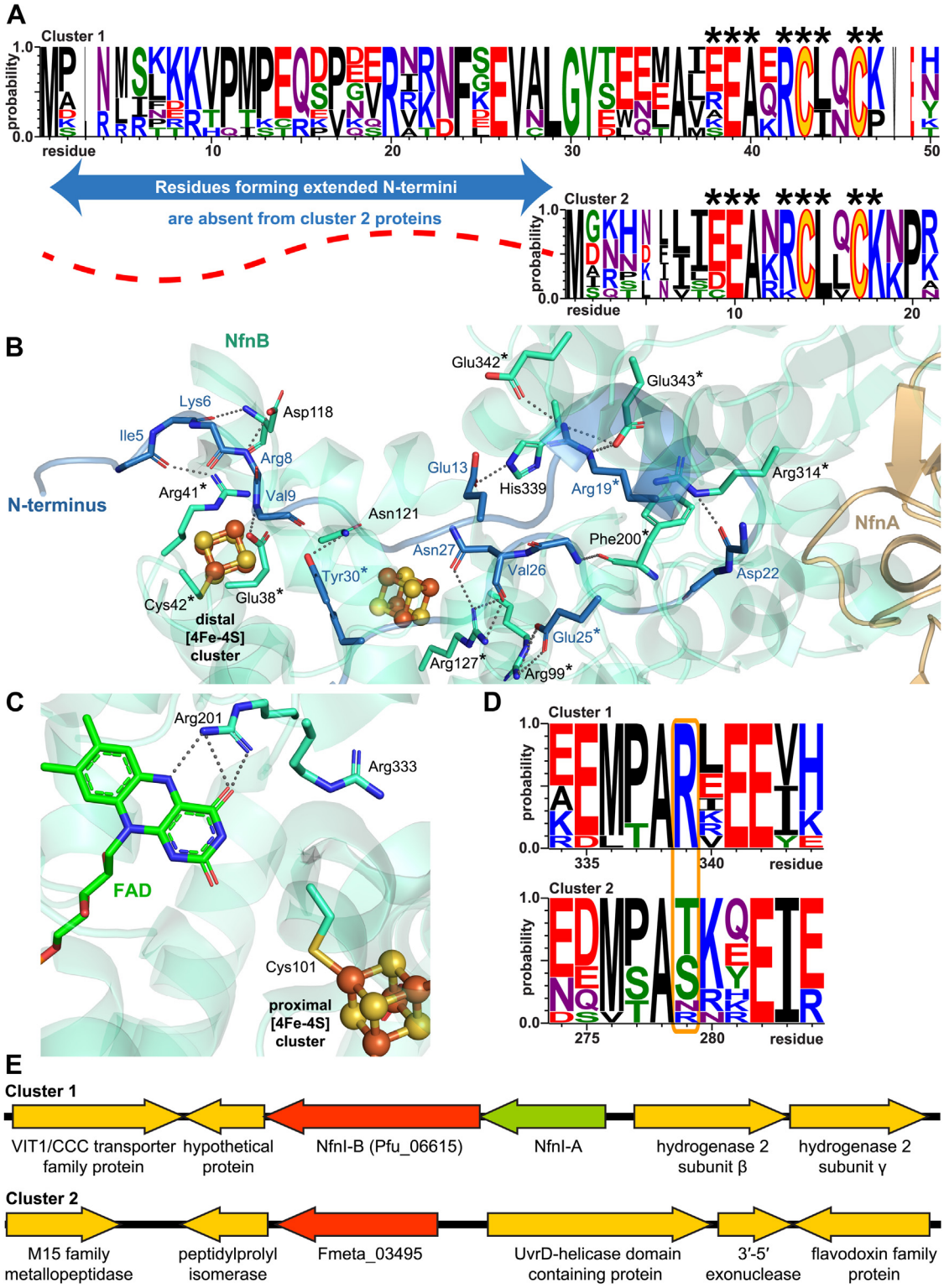



**Figure 8. Differences between proteins in phylogenetic clusters 1 and 2.** *A*, logograms depicting sequence alignments of cluster 1 and cluster 2 (except Caau_1998) proteins at their N-termini. Cluster 1 proteins (*top*) have longer N-termini, depicted by the *blue arrow*, not conserved in cluster 2 proteins (*bottom*). *Asterisks* (*) indicate the position of >50% conserved residues, and *yellow* 'C's indicate the first two [4Fe-4S]-ligating cysteines (Cys42 and Cys45 in *Pf*NfnI-B). *B*, *Pf*NfnI structure (PDB ID: 5JCA (32)) shows specific hydrogen bonding and ionic contacts between N-terminal (*blue*) and other (*green*) residues of *Pf*NfnI-B. FAD and main/side chains of residues not involved in contacts have been omitted for clarity. *Asterisks* (*) indicate fully conserved residues in cluster 1. *C*, Arg333 is located close to bifurcating FAD in *Pf*NfnI-B and could play a role in electron bifurcation. *D*, sequence alignments depicting the residue corresponding to Arg333 in Pfu_06615 (along with five upstream and downstream residues) in cluster 1 and cluster 2 proteins. The relevant residue (*orange box*) is a fully conserved arginine in cluster 1 proteins and is not conserved in cluster 2 proteins. Logograms and residue numbering are based on consensus sequences arising from aligning all cluster 1 (*top*) and cluster 2 (*bottom*) protein sequences (see also Figs. S2, S3). *E*, genome neighborhoods of typical

Nfn electron bifurcation turnover (32, 47). This residue is fully conserved in cluster 1, while it is not conserved in cluster 2 (Figs. 8*D*, S2, S3). One protein (Fmeta_03495) does have an Arg residue, whereas five have Thr, four have Ser, and the Fer4_20 domain-lacking Caau_1998 has an Asn at the corresponding position. These polar-uncharged and smaller residues are unlikely to interact with FAD and contribute to Nfn-like reactivity in cluster 2 proteins.

While alignments of protein sequences have been informative in discriminating between the two types of proteins, genome neighborhoods can more definitively address if a protein is an Fnor or an Nfn. The vast majority of known Nfns—including all characterized to date—are made up of discrete NfnA and NfnB subunits (60). Their corresponding genes, *nfnA* and *nfnB*, are colocalized in microbial genomes (61) and *nfnA* is upstream of *nfnB* (Fig. 8*E*), with a gap or overlap of a few nucleotides between the two (60). Searching the neighborhoods of genes corresponding to the proteins in cluster 1 all revealed an upstream gene, the product of which showed high homology to NfnA (Table S2). Conversely, neighborhoods of genes corresponding to the proteins in cluster 2 revealed no adjacent or nearby genes homologous to *nfnA* (Tables S3–S13).

Within protein sequences, the presence of a longer N terminus and conservation of Arg333 appear to be features unique to NfnBs. In terms of genomic organization, an *nfnA*-like gene upstream of one that encodes a protein comprising Fer4_20 and Pyr_redox_2 domains is the strongest indicator that the downstream gene is *nfnB*. All cluster 1 proteins and genes meet all the above criteria, while none of cluster 2 members do. Fmeta_03495 comes closest with its conservation of Arg at the position corresponding to Arg333 in *Pf*NfnI-B, but its short N terminus and lack of *nfnA*-like gene adjacent to or nearby its gene strongly support it being a Fnor enzyme. This analysis also supports the prediction of thus far uncharacterized Kpar_06265 and Tcaen_100634 as NfnB proteins (Table S1). Taken together, these three factors (an extended N terminus, presence of Arg333, and an *nfnA* homolog upstream of the gene) compellingly differentiate the NfnB-type enzymes in cluster 1 from the Cac_0764-type Fnor enzymes in cluster 2.

### *Role of NfnB as a monofunctional enzyme*

The NFN reaction (Equation 3) couples the exergonic transfer of electrons from reduced ferredoxin to $NADP^+$ with the endergonic transfer of electrons from NADH to $NADP^+$ (*i.e.* transhydrogenation) (25). The Nfn protein complex with two subunits (NfnA and NfnB) mediates this reaction. However, NfnB binds both ferredoxin and $NADP^+$, which raises the question of whether the NfnB subunit can perform the exergonic FNOR reaction (Equation 1) by itself or if the full NfnAB complex is necessary due to a strict 1:1 stoichiometric linkage between the exergonic and endergonic reactions.

Previous work, including detailed structural analysis (29, 32, 48, 62, 63), suggests that for the WT NfnAB complex, NfnB-type enzymes have strict energetic coupling, and therefore 1:1 stoichiometry, between the two branches. However, our enzyme assay results indicate that cluster 1 NfnB proteins can function independently, catalyzing an NADPH-linked FNOR reaction (Equation 1) without requiring NfnA for the associated transhydrogenation. This observation is not unprecedented: Wise *et al.* reported NADPH-linked FNOR activity (which they termed "half-confurcation") for WT NfnL (equivalent to Pfu_06615) in the absence of NfnS (NfnA complement of Pfu_06615) (48). These results are consistent with the overall reduction of $NADP^+$ from reduced Fd being exergonic (25), and therefore this NADPH-linked FNOR reaction can be catalyzed by individual NfnB proteins which bind both substrates. In our assays, Pfu_06615 exhibited an average NADPH-linked FNOR activity of 1.39 U/mg. This activity is almost two orders of magnitude higher than that reported by Wise *et al.* (48), which we attribute to higher temperatures (55 °C), near-saturating $NADP^+$ concentrations, and continuous regeneration of reduced Fd with Pfor in our assay conditions. However, all NfnB proteins characterized in this study have *in vitro* NADPH-linked FNOR activities at least an order of magnitude lower than those of the eight most active cluster 2 and cluster 3 proteins (Fig. 4*D*), suggesting that the NfnB proteins are not very efficient at catalyzing only one of their native reactions. Indeed, Wise *et al.* reported full-confurcation activity of *Pf*NfnI–AB complex with all substrates present being 1.5 times higher than NADPH-linked FNOR activity (48), indicating that the enzyme is more efficient when it can carry out its native activity of coupled FNOR and transhydrogenation reactions.

*P. furiosus* NfnII presents an enigma because, despite possessing both NfnA and NfnB (Pfu_09610) subunits, the complete complex is unable to carry out the NFN electron bifurcation reaction (31) (Equation 3). Structural analysis indicates that this is due to the occlusion of NAD(H)-binding site in *P. furiosus* NfnII-A. However, unlike *P. furiosus* NfnI, the complete NfnII–AB complex exhibits NADPH-linked monofunctional FNOR reactivity (Equation 1). It is hypothesized that, *in vivo*, NfnII functions either as an Fnor or "Xfn" (where X is an unknown substrate taking the place of NAD(H) in the electron bifurcation reaction). Our work demonstrates that *P. furiosus* NfnII-B (Pfu_09610) alone has 2.59 U/mg NADPH-linked FNOR activity (Fig. 4*D*), comparable to 5.0 U/mg FNOR activity reported by Nguyen *et al.* for the complete NfnII–AB complex (31). This raises the question of how important the NfnII-A subunit is to the FNOR activity of the complete NfnII complex. *In vitro* assays and functional

cluster 1 and cluster 2 genes, with protein products identified under the corresponding genes. Cluster 1 genes, exemplified by that of Pfu_06615, have an upstream *nfnA* gene, which is not present upstream or in the vicinity of cluster 2 genes, like that of Fmeta_03495.

expression in *C. thermocellum* of the complete *P. furiosus* NfnII–AB complex could be future targets for assessing its FNOR activity, albeit in a non-native context.

### Explaining the activities of other potential Fnors

Products of *nfnA* genes are commonly misannotated as class 1b dihydroorotate dehydrogenase electron transfer subunit, PyrK (64), to which these have homology. Nonetheless, all translated protein sequences from the gene upstream of cluster 1 genes are more homologous to the reference *Pf*NfnI-A sequence and each other rather than to the known PyrK from *Lactococcus lactis* (64, 65) or Tsac_1705, the putative PyrK from *T. saccharolyticum* (Figs. S4, 2). This homology includes the residues coordinating the [2Fe-2S] cluster: in NfnA, these are three cysteines and one aspartate (Asp217 in *Pyrococcus furiosus* NfnI-A (66)), while in a typical PyrK, four cysteines bind the [2Fe-2S] cluster.

Interestingly, Tsac_1705 is known to exhibit NADH-linked FNOR activity (37), which has been replicated in this work (Fig. 4). This FNOR activity could arise from nonspecific, collisional interactions between Tsac_1705 and Fd molecules in solution which results in thermodynamically favorable electron transfer from reduced Fd, *via* the surface-exposed [2Fe-2S] cluster and FAD, ultimately to the bound $NAD^+$. However, Tsac_1705 maintains its fidelity for the correct pyridine nucleotide, binding, and reducing $NAD^+$ but not $NADP^+$ and therefore not having any NADPH-linked FNOR activity. Tsac_2086, the NfnA protein from *T. saccharolyticum* and a homolog of Tsac_1705, does not have any FNOR activity, which could be due to its altered thermodynamics of electron transfer arising at least in part from the coordination state of its [2Fe-2S] cluster.

Despite its homology to both clusters 1 and 2, Sov_03740 (StnC) had the lowest detectable Fd oxidation activity with $NADP^+$ and higher BV reduction with NADPH than the control. These activities are consistent with its binding NADP(H) in the part of its structure most homologous to cluster 1 and 2 proteins (33) and engaging in electron transfer to/from it. However, StnC does not interact with Fd in the flavin-based electron bifurcation reaction carried out by the Stn complex, which indicates that its FNOR activity here is likely due to a nonspecific interaction with Fd (25).

Cluster 3 proteins are canonical plant-type Fnor enzymes (*see below*) which rely on FAD or FMN as their only electron-transferring cofactor. Consistent with this classification, both proteins (Telo_06135 and Hth_1218) from this cluster showed much higher FNOR reactivity with $NADP^+$ than with $NAD^+$ (Fig. 4). Since these originate from thermophilic bacteria, their thermostability is also evident by the retention of their activity after incubation at 55 °C (Fig. 5). Once the functional expression was achieved, Hth_1218 proved to be the best Fnor candidate *in vivo* in *C. thermocellum* strain LL1784, enabling a 2.2-fold increase in ethanol production under fermentation conditions. In addition to its intrinsic NADPH-linked FNOR activity, the success of Hth_1218 *in vivo* is likely also due to its structural simplicity: it is smaller than other proteins with NADPH-linked FNOR activity and relies only on a single flavin cofactor for its activity. These features differentiate it from the larger cluster 1 and 2 proteins—which require both FAD and (except for Caau_1998) [4Fe-4S] clusters for activity—and from Telo_06135, which may need to associate with phycobilisome proteins *via* its CpcD domain (34, 67) for full functionality. The structural and functional simplicity of Hth_1218 could enable its successful heterologous expression and cofactor incorporation in *C. thermocellum* strain LL1784, resulting in its high FNOR activity and therefore high ethanol titers under fermentation conditions.

### Cluster 2 proteins comprise a novel family of Fnors

Nicotinamide cofactors NAD(P) (H) are involved in the transfer of electron pairs (68), while ferredoxins are single-electron donors/acceptors (69). These two types of reductants are typically noninterchangeable and interact with different redox-active partners (70, 71). Depending on metabolic needs, one type of reductant could be required more than the other (69, 72), necessitating a way of converting between the two. Direct reduction of $NAD(P)^+$ with $Fd_{red}$ (or the reverse reaction) is not possible. Fnor enzymes are versatile metabolic "brokers" which enable this electron transfer between one- and two-electron physiological reductants (Equation 1) (69, 72). The FNOR activity is dependent on a flavin cofactor (FMN or FAD) in these enzymes, which can accept and donate both single-electrons and electron pairs (68). This makes the flavin cofactor an essential part of Fnor enzymes, although other cofactors such as iron-sulfur clusters can also be found in more complex systems (*see below*) to facilitate electron transfer.

The reduction of $NAD(P)^+$ by reduced ferredoxin (Equation 1) is typically highly exergonic (25, 73). Many microorganisms—especially extremophiles living at thermodynamic limits (74–77)—utilize this favorable reaction to drive other, unfavorable reactions by means of specialized proteins. One way to functionally categorize proteins with FNOR reactivity is based on the presence or absence of energy conservation (Fig. 9). The first category comprises energy-conserving Fnors: these proteins couple the exergonic FNOR reaction to another, endergonic reaction, resulting in energy conservation. The membrane-bound Rnf protein complex is an example of an energy-conserving Fnor, which couples $Fd_{red}$ to $NAD^+$ electron transfer with export of protons or $Na^+$ (Equation 2). This generates a transmembrane electrochemical ion gradient, the dissipation of which is coupled to ATP synthesis by membrane-bound ATP synthase (78, 79). Enzymes that engage in flavin-based electron bifurcation can also be classified as energy-conserving Fnors. The best example of such an enzyme is Nfn (25, 32, 63), where the NADPH-linked Fnor reaction is coupled to transhydrogenation of $NADP^+$ from NADH (Equation 3), although other electron-bifurcating flavoproteins such as Etf (80) and Stn (33) could be considered energy-conserving Fnors as well.

The second category (Fig. 9) consists of nonenergy-conserving Fnors: these proteins *only* catalyze $Fd_{red}$ to

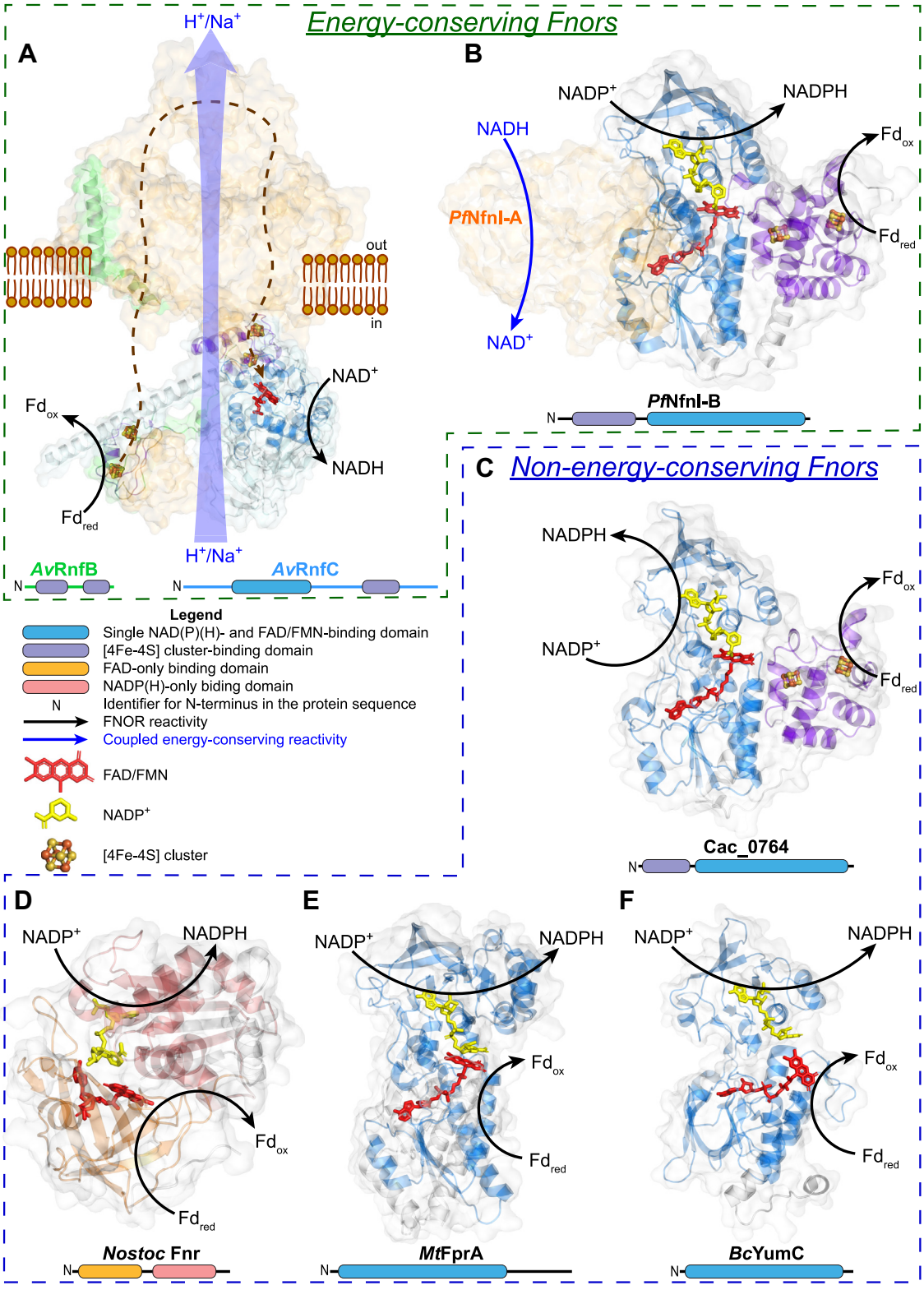



**Figure 9. A functional characterization of Fnors based on energy conservation.** Structure, domain organization, and electron transfer reactivity of different types of Fnors is shown. See Equations 1–3 for the complete reactions and accurate stoichiometries. Domain organization in the protein sequence is depicted below the structures, and domains are identified in the structure based on the same color coding. Subunits not involved in Fnor reactivity are shown by *orange* surfaces. **Energy-conserving Fnors**: (*A*) *Azotobacter vinelandii* Rnf1 complex (PDB ID: 8AHX (28)), where Fd oxidation takes place at RnfB, leading to electron transfer (*brown dashed arrow*) spanning the rest of the complex to RnfC, where $NAD^+$ reduction takes place. Electron transfer within Rnf is coupled to transmembrane $H^+/Na^+$ translocation. *B*, *Pyrocccus furiosus* NfnI (PDB ID: 5JCA (32)), where Fnor activity is coupled to transhydrogenation of $NADP^+$ by NADH. **Nonenergy-conserving Fnors**: (*C*) *Clostridium acetobutylicum* Cac_0764 (30) (AlphaFold model accession number: Q97L02, with $NADP^+$ and cofactors superimposed from *P. furiosus* NfnI structure (32)), a novel type of monofunctional Fnor homologous to NfnB; (*D*) *Nostoc* sp. PCC 7119 Fnr (PDB ID: 1GJR (81)), representative of plant-type Fnor; (*E*) *Mycobacterium tuberculosis* FprA (PDB ID: 1LQU (82)) representative of glutathione reductase-type Fnor; and (*F*) *Bacillus subtilis* YumC (PDB ID: 3LZW (85)) representative of thioredoxin reductase–type Fnor.

$NAD(P)^+$ electron transfer without any coupled reactivity (Equation 1). There are at least three distinct phylogenetic/structural families of known nonenergy-conserving Fnor (better known as Fnr) enzymes (44, 69, 72). The best known family is plant-type Fnor enzymes (81), further subdivided into plastidic- (found in plants and cyanobacteria) and bacterial-type Fnors. Sole_110801208 and Telo_06135 in cluster 3 from our phylogenetic analysis (Fig. 2) are plastidic-type Fnors, while Hth_1218 is a bacterial-type Fnor. The second family is glutathione reductase–type Fnors, typically found in eukaryotic mitochondria as well as some bacteria (82). A third family, thioredoxin reductase–type Fnors (83–85), has been more recently characterized. While these "true" Fnor families diverge structurally, all of these proteins bind and rely only on one flavin cofactor (FMN or FAD) to catalyze electron transfer between Fd and NADP(H).

Cac_0764-type (30) cluster 2 proteins therefore represent a novel phylogenetic and structural family of nonenergy-conserving Fnor proteins. In terms of reactivity, these are like the other three families of "true" nonenergy-conserving Fnors, linking only Fd and $NADP^+$. Structurally, however, these cluster 2 proteins are very different from the three known Fnor families and closer to NfnB proteins from cluster 1, relying on two [4Fe-4S] clusters in addition to FAD for their electron transfer activity. Nonetheless, despite their overall structural similarity to NfnB proteins, cluster 2 proteins are different enough (*see above*) that they are highly unlikely to be involved in electron bifurcation–type energy conservation. Therefore, we have classified Cac_0764-type cluster 2 proteins as a novel family of FAD- and [4Fe-4S] cluster-dependent nonenergy-conserving Fnor, for the first time to the best of our knowledge, which is distinct from other flavin-only Fnor and NfnB protein families.

## Conclusions

The original goal of this work was to identify Fnor enzymes that could be functionally expressed in *C. thermocellum* to increase ethanol production. The enzymes Hth_1218 and Claur_019100 exhibit exceptionally high NADPH-linked FNOR activity and thermostability *in vitro* and also enabled increased ethanol titer when expressed in *C. thermocellum*. Interestingly, Hth_1218 is a plant-type Fnor, while Claur_019100 is a novel Fnor similar to Cac_0764. These improvements in ethanol production demonstrate progress toward addressing one of the challenges in developing consolidated bioprocessing for lignocellulosic biofuel production.

In the course of this work, we determined that the original Cac_0764 enzyme (30) is a member of a new class of Fnor enzymes that are homologous to NfnB proteins and distinct from other, known Fnor families. We have identified three key characteristics of these enzymes that allow them to be identified and differentiated from NfnB proteins based on their amino acid sequence and genomic context. Finally, we characterized several examples of these enzymes for activity, nicotinamide cofactor specificity, thermotolerance, and functional expression in *C. thermocellum*.

Beyond their immediate application in ethanol production, the identification of this new class of Fnor enzymes has broader implications for metabolic engineering and our understanding of microbial electron transfer systems. These enzymes could potentially be applied to the production of other reduced products requiring NADPH, offering new tools for redox balance engineering in industrial microorganisms. Furthermore, the discovery and characterization of this distinct class of Fnor enzymes provides insights into the evolution of electron transfer systems across different organisms, contributing to our understanding of how these critical metabolic capabilities have developed and diversified.

## Experimental procedures

### Strains

*C. thermocellum* strain LL1784 (*Δhpt ΔreIII P2638::adhA(Tsc) nfnAB(Tsc) adhEG544D(Tsc) Clo1313_2637::Ptsc0046-pforA(Tsc)-ferredoxin(Tsc) Δpfor1 Δpfor4 Δpfor3 Δpfor2 Δpfor5 ΔadhE Δldh ΔnfnAB Δrnf*) was used as Fnor expression host. Its native genes *rnf* and *nfnAB* have been deleted to reduce background levels of FNOR activity and direct electron flux from Fd to Fnor candidates. This strain has NADH-linked ALDH activity and NADPH-linked ADH activity. *E. coli* T7 Express *lysY/I$^q$* (NEB catalog #C3013I) was used for the overexpression of Fnor enzymes for purification. *E. coli* T7 Express (NEB catalog #C2566H) was used for isolating plasmid DNA for *C. thermocellum* transformation. *E. coli* G8EB7 (MG1655 *ΔiscR::kana pthl-Fd-LL-C-Tag*) is an engineered strain used for improved expression of *T. saccharolyticum* ferredoxin.

### Plasmids

Information on the plasmids used for this work is listed in Table 1.

### Growth media

*C. thermocellum* cells were grown in either CTFUD-rich medium (86) or MTC (87) chemically defined medium. *E. coli* cells were grown in LB medium. Ferredoxin expression media contains 20 g/L glucose (Sigma Aldrich catalog #G7528), 10 g/L tryptone (Thermo Fisher Scientific catalog #BP1421), 5 g/L yeast extract (Sigma Aldrich catalog #Y1625), 23 g/L Hepes (Sigma Aldrich catalog #H4034), 50 mg/L $FeSO_4$ (Sigma Aldrich catalog #215422), 200 mg/L nitrilotriacetic acid (Sigma Aldrich catalog #72559), 0.5 g/L $K_2HPO_4$ (Thermo Fisher Scientific catalog #P288), 2 g/L NaCl (Thermo Fisher Scientific catalog #S271), 0.0811 g/L $FeCl_3$ (Sigma Aldrich catalog #236489), and 85 mg/L $NaNO_3$ (Sigma Aldrich catalog #S8170) with 100 μg/ml carbenicillin (Sigma Aldrich catalog #C1389).

### Cell extract preparation

Cells of *C. thermocellum* were grown to mid-log phase (absorbance at 600 nm of 0.6) collected by centrifugation and resuspended in B-PER lysis buffer (Thermo Fisher Scientific

**Table 1**
**Plasmids used for this work**

| Plasmid | Function | Genbank accession number |
|---|---|---|
| pLL1508 | Overexpression of strep-tagged *T. saccharolyticum* ferredoxin (Tsac_2084) in *E. coli* | PQ321886 |
| pLL1509 | Overexpression of Hth_1218 in *C. thermocellum* | PQ321887 |
| pLL1510 | Overexpression of Clth_22750 in *C. thermocellum* | PQ321888 |
| pLL1511 | Overexpression of Claur_019100 in *C. thermocellum* | PQ321889 |
| pLL1512 | Overexpression of Tform_00461 in *C. thermocellum* | PQ321890 |
| pLL1513 | Overexpression of Tthal_01343 in *C. thermocellum* | PQ321891 |
| pLL1514 | Overexpression of Fmeta_03495 in *C. thermocellum* | PQ321892 |
| pLL1515 | Overexpression of Cazo_06885 in *C. thermocellum* | PQ321893 |
| pLL1516 | Overexpression of Cmit_267 in *C. thermocellum* | PQ321894 |
| pLL1517 | Overexpression of Telo_06135 in *C. thermocellum* | PQ321895 |
| pLL1518 | Overexpression of Tsac_2085 in *C. thermocellum* | PQ321896 |
| pLL1519 | Overexpression of C-terminal 6x his tagged Tsac_2085 in *E. coli* | PQ321897 |
| pLL1520 | Overexpression of C-terminal 6x his tagged Ctx_1849 in *E. coli* | PQ321898 |
| pLL1521 | Overexpression of C-terminal 6x his tagged Pfu_06615 in *E. coli* | PQ321899 |
| pLL1522 | Overexpression of C-terminal 6x his tagged Pfu_09610 in *E. coli* | PQ321900 |
| pLL1523 | Overexpression of C-terminal 6x his tagged Sov_03740 in *E. coli* | PQ321901 |
| pLL1524 | Overexpression of C-terminal 6x his tagged Telo_06135 in *E. coli* | PQ321902 |
| pLL1525 | Overexpression of C-terminal 6x his tagged Tma_1681 in *E. coli* | PQ321903 |
| pLL1526 | Overexpression of N-terminal 6x his tagged Tsac_1705 in *E. coli* | PQ321904 |
| pLL1527 | Overexpression of N-terminal 6x his tagged Cac_0764 in *E. coli* | PQ321905 |
| pLL1528 | Overexpression of N-terminal 6x his tagged Tsac_2086 in *E. coli* | PQ321906 |
| pLL1529 | Overexpression of N-terminal 6x his tagged Agott_100192 in *E. coli* | PQ321907 |
| pLL1530 | Overexpression of N-terminal 6x his tagged Claur_019100 in *E. coli* | PQ321908 |
| pLL1531 | Overexpression of N-terminal 6x his tagged Caau_1998 in *E. coli* | PQ321909 |
| pLL1532 | Overexpression of N-terminal 6x his tagged Cazo_06885 in *E. coli* | PQ321910 |
| pLL1533 | Overexpression of N-terminal 6x his tagged Cbei_13145 in *E. coli* | PQ321911 |
| pLL1534 | Overexpression of N-terminal 6x his tagged Cbei_00860 in *E. coli* | PQ321912 |
| pLL1535 | Overexpression of N-terminal 6x his tagged Cmit_267 in *E. coli* | PQ321913 |
| pLL1536 | Overexpression of N-terminal 6x his tagged Clth_22750 in *E. coli* | PQ321914 |
| pLL1537 | Overexpression of N-terminal 6x his tagged Fmeta_03495 in *E. coli* | PQ321915 |
| pLL1538 | Overexpression of N-terminal 6x his tagged Hth_1218 in *E. coli* | PQ321916 |
| pLL1539 | Overexpression of N-terminal 6x his tagged Kpar_06265 in *E. coli* | PQ321917 |
| pLL1540 | Overexpression of N-terminal 6x his tagged Tcaen_100634 in *E. coli* | PQ321918 |
| pLL1541 | Overexpression of N-terminal 6x his tagged Tform_00461 in *E. coli* | PQ321919 |
| pLL1542 | Overexpression of N-terminal 6x his tagged Tmt_2096 in *E. coli* | PQ321920 |
| pLL1543 | Overexpression of N-terminal 6x his tagged Tsac_1891 in *E. coli* | PQ321921 |
| pLL1544 | Overexpression of N-terminal 6x his tagged Tthal_01343 in *E. coli* | PQ321922 |
| pLL1545 | Overexpression of N-terminal 6x his tagged Tthe_1158 in *E. coli* | PQ321923 |
| pDGO143 | Empty expression vector for *C. thermocellum* | KX259110 |
| pRKisc | Overexpression of *isc* gene cluster | N/A (88) |

The table lists the function of each plasmid, and their sequences can be accessed through the GenBank accession numbers provided.

catalog #78243). Ready-Lyse lysozyme (Lucigen catalog #R1804M) and OmniCleave endonuclease (Lucigen catalog #OC7850K) were added to the suspension, which was then incubated on ice for 30 min. After incubation, the mixture was centrifuged, and the supernatant was collected for further analysis.

### *Recombinant expression and purification of Fnor and Pfor*

Expression plasmids were transformed into the *E. coli* T7 Express lysY/Iq strain along with plasmid pRKisc (88). The cells were aerobically cultivated in LB medium supplemented with 10 μg/ml tetracycline and 35 μg/ml kanamycin at 37 °C with shaking at 225 RPM. At mid-log phase (absorbance at 600 nm of 0.5), expression of both the target protein and the *isc* operon was induced by the addition of 0.4 mM IPTG (IPTG, Promega catalog #V395D). To enhance iron-sulfur cluster synthesis, the LB growth medium was supplemented with 0.12 g/L cysteine (Sigma Aldrich catalog #C6852), 0.1 g/L ferrous sulfate (Sigma Aldrich catalog #215422), 0.1 g/L ferric citrate (Sigma Aldrich catalog #F3388), and 0.1 g/L ferric ammonium citrate (Sigma Aldrich catalog #F5879) as suggested by Tian *et al.* (37). The cells were then grown anaerobically at 30 °C overnight and harvested by centrifugation. The cell pellets were washed with 50 mM Tris-HCl–0.2 mM DTT (Sigma Aldrich catalog #43816) (pH 8) and stored at −80 °C.

Protein purification was performed under anaerobic conditions. Cytiva His SpinTrap columns (Cytiva catalog #28401353) were used for Fnor enzyme purification. The columns were equilibrated with 600 μl of binding buffer (20 mM Tris–HCl, 500 mM NaCl, 40 mM imidazole, pH 8.0). The cell lysates were then applied to the columns, followed by three washes with 600 μl of binding buffer each time. The proteins were subsequently eluted twice with 200 μl of elution buffer (20 mM Tris–HCl, 500 mM NaCl, 300 mM imidazole, pH 8.0) each time.

### *Recombinant expression and purification of Fd*

Ferredoxin was purified as previously described (89). The recombinant *E. coli* strain G8EB7 (89) was grown anaerobically at 30 °C with 200 rpm shaking in ferredoxin expression media. Cells were harvested by centrifugation when the OD reached ∼2. The cell pellet was resuspended in the binding buffer (100 mM Tris–HCl, 150 mM NaCl, 1 mM EDTA, pH 8) and was lysed *via* sonication. After sonication, the lysate was clarified by centrifugation to separate it from the cell debris. A Cytiva StrepTrap XT column (Cytiva catalog #29401320) was used for the purification. The column was first equilibrated


ASBMB

with 5 column volumes (CV) of binding buffer before the sample was loaded onto the column. It was then washed with 10 CV of binding buffer. Ferredoxin was eluted with six CV of elution buffer (50 mM biotin in binding buffer), and the elution was fractionated to collect the most concentrated fractions. The eluted ferredoxin has a brown color.

### Protein quantification

Protein concentrations were determined *via* the Bradford assay, which relies on the color change of Coomassie G-250 dye (Thermo Fisher Scientific catalog #23238) measured at 595 nm. Bovine serum albumin (Thermo Fisher Scientific catalog #23209) was used as the standard.

### Enzyme assays

Enzyme assays were performed using Agilent 8453 UV-visible Spectroscopy System under anaerobic conditions at 55 °C unless otherwise noted. Each reaction had a total volume of 1200 µl, with 1 to 15 µg of the protein of interest added. Pfor was included in the FNOR reaction assays (along with substrates pyruvate and coenzyme A) to continuously regenerate $Fd_{red}$ from $Fd_{ox}$ as a result of the PFOR reaction (63), thereby maintaining a constant concentration of almost fully reduced ferredoxin for the FNOR reaction (Equation 1). The concentrations of NAD(P) (H) cofactors used were consistent with those reported for similar assays (29, 30–32, 37, 62, 63). While affinity data are scarce, the concentrations used are 1 to 3 orders of magnitude higher than the five reported $K_D$ (47) or $K_M$ (29, 36, 63, 90) values. It is reasonable to extrapolate that the cofactor concentrations used were saturating given that most (24 of 27) proteins are homologous to a protein with a measured affinity value. The remaining imidazole in enzymatic assays is in the range of 0.2 to 8.3 mM.

### BV reduction with NAD(P)H (FNOR)

The reaction mixture contained 100 mM Tris–HCl (pH 7.5), 0.3 mM NADH (Sigma Aldrich catalog #N8129) or NADPH (Sigma Aldrich catalog #N7505), and 1 mM benzyl viologen (Sigma Aldrich catalog #271845). The reaction was started with NAD(P)H, and the BV reduction was followed by photometric observations at 578 nm (ε = 7.8 mM − 1 cm−1).

### $NAD(P)^+$ reduction with reduced ferredoxin (FNOR)

The reaction mixture contained 100 mM Tris–HCl (pH 7.5), 0.012 mM FAD (Sigma Aldrich catalog #F6625), 2 mM $MgCl_2$ (Sigma Aldrich catalog #M2670), 1 mM coenzyme A (Sigma Aldrich catalog #C3144), 0.4 mM thiamine pyrophosphate (Sigma Aldrich catalog #C8754), 2 mM $NAD(P)^+$ (Sigma Aldrich catalog #N6522 and N5755), 150 mg/L ferredoxin, 5 U/ml Pfor (activity based on BV activity), 2 mM sodium pyruvate (Sigma Aldrich catalog #P2256). The reaction was started with pyruvate, and the $NAD(P)^+$ reduction was followed by photometric observations at 340 nm (ε = 6.3 mM − 1 cm−1).

### BV reduction with pyruvate (PFOR)

The reaction mixture contained 100 mM Tris–HCl (pH 7.5), 2 mM $MgCl_2$, 1 mM coenzyme A, 0.4 mM thiamine pyrophosphate, 1 mM benzyl viologen, 2 mM sodium pyruvate. The reaction was started with pyruvate, BV reduction was followed by photometric observations at 578 nm (ε = 7.8 mM−1 cm−1).

### C. thermocellum *transformation*

*C. thermocellum* transformation was performed as previously described (86). To prepare the plasmid DNA for transformation, we introduced the plasmids into the *E. coli* T7 Express strain and subsequently performed plasmid purification (note that we use *E. coli* B strains rather than the more commonly used cloning strains due to differences in DNA methylation that increase transformation efficiency in *C. thermocellum* (91)). *C. thermocellum* cells were cultured anaerobically in 50 ml of CTFUD media until reaching mid-log phase. The cells were then harvested by centrifugation and washed three times with water. The resulting pellet was resuspended in 100 µl of water. A 20 µl aliquot of the cell suspension was mixed with 5 µl of plasmid DNA and transferred to an electroporation cuvette. The mixture was subjected to a square wave electroporation at 1500 V for 1.5 ms. Following electroporation, the mixture was transferred to 5 ml of CTFUD and incubated at 51 °C overnight. Several dilutions of the recovered culture were plated on CTFUD-agar plates supplemented with 6 µg/ml thiamphenicol (Sigma Aldrich catalog #T0261). Colony PCR was used to confirm the presence of the correct plasmid.

### C. thermocellum *fermentation*

High substrate (50 g/L cellobiose) fermentations were performed in 14 ml culture tubes with 1 ml working volume at 55 °C in Coy anaerobic glove bag (Coy Laboratory Products).

### Analytical methods

Cellobiose, glucose, pyruvate, lactate, formate, acetate, malate, and ethanol were quantified by HPLC (Waters or LC-2030, Shimadzu) using an Aminex HPX-87H column (Bio-Rad) with refractive index and UV detection, and a 5 mM sulfuric acid solution as the eluent using a 0.6 ml/minute flow rate, and a column temperature of 60 °C (92).

### Bioinformatic analyses

The phylogenetic tree was constructed with COBALT using the Cobalt tree method. Further sequence alignments were performed with COBALT (93) and visualized with SnapGene Viewer software (www.snapgene.com). Logograms were generated with WebLogo 3 (94).

### Pathway thermodynamic analysis

Pathway thermodynamic analysis was performed using the equilibrator_api Python package version 0.5.0 (https://gitlab.com/equilibrator/equilibrator-api) (95). This type of analysis

takes as input a list of stoichiometric reactions, a list of compounds, and the maximum and minimum concentration ranges for each compound. With these constraints, the MDF of a pathway can be determined by identifying the concentrations for each compound that maximize the minimum thermodynamic driving force for the least favorable reaction (57). Key constraints include a glucose concentration of 10 mM and an ethanol concentration of 2000 mM. Ratios for redox cofactor pairs (NADH/$NAD^+$, NADPH/$NADP^+$, and $Fd_{red}/Fd_{ox}$) were allowed to vary between 0.01 and 100.

Detailed information for replicating this analysis is provided in Data S1. This includes an envirnoment.yml file that lists the dependencies for the Python software environment, a Jupyter Notebook file with the code used to perform the analysis and generate the figures, and an Excel file with the input parameters (reactions, compounds, and compound concentration limits) used to constrain the model.

## Data availability

All data are available in the main text or the supporting materials.

*Supporting information*—This article contains supporting information (96).

*Acknowledgments*—Synthesis of DNA constructs was conducted by the U.S. Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231.
We thank Professor Isabelle Meynial-Salles and her team at National Institute of Applied Sciences of Toulouse for the *E. coli* G8EB7 strain. This work was authored in part by the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

*Author contributions*—S. H., S. M. S. I., C. E. L., and D. G. O. writing–review and editing; S. H., S. M. S. I., and D. G. O. writing–original draft; S. H. and S. M. S. I. visualization; S. H. methodology; S. H., S. M. S. I., A. A. L., and S. K. H. investigation; S. H. and D. G. O. conceptualization; C. E. L., L. R. L., and D. G. O. supervision; C. E. L., L. R. L., and D. G. O. funding acquisition.

*Funding and additional information*—This work was supported by the Center for Bioenergy Innovation (CBI), U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number ERKP886 (S. H., S. K. H., A. A. L., D. G. O., and L. R. L.).
Funding was provided to C. E. L. (data interpretation and writing/editing of manuscript) by the U.S. DOE Office of Science Early Career Program and supported by the Laboratory Directed Research and Development (LDRD) Program at NREL to C. E. L. and S. M. S. I. (phylogenetic analysis of FNOR proteins, data interpretation, and writing/editing of manuscript).

*Conflicts of interest*—L. R. L. is the co-founder and CTO of the Terragia corporation (https://terragiabiofuel.com/). Terragia has a financial interest in the commercialization of *C. thermocellum*. The other authors declare that they have no conflicts of interests with the contents of this article.

*Abbreviations*—The abbreviations used are: BV, benzyl viologen; CBP, consolidated bioprocessing; CV, column volume; Fnor, ferredoxin:NAD(P)+ oxidoreductase; Fd, ferredoxin; MDF, max-min driving force; Pfor, pyruvate:ferredoxin oxidoreductase.

## References

1. Reid, W. V., Ali, M. K., and Field, C. B. (2020) The future of bioenergy. *Glob. Change Biol.* **26**, 274–286
2. Fulton, L. M., Lynd, L. R., Körner, A., Greene, N., and Tonachel, L. R. (2015) The need for biofuels as part of a low carbon energy future. *Biofuels Bioprod. Biorefining.* **9**, 476–483
3. Lynd, L. R., Guss, A. M., Himmel, M. E., Beri, D., Herring, C., Holwerda, E. K., *et al.* (2017) Advances in consolidated bioprocessing using Clostridium thermocellum and thermoanaerobacter saccharolyticum. In *Industrial Biotechnology*, John Wiley & Sons, Ltd: 365–394. https://doi.org/10.1002/9783527807796.ch10
4. Demain, A. L., Newcomb, M., and Wu, J. H. D. (2005) Cellulase, clostridia, and ethanol. *Microbiol. Mol. Biol. Rev. MMBR.* **69**, 124–154
5. Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I. S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66**, 506–577
6. Tian, L., Papanek, B., Olson, D. G., Rydzak, T., Holwerda, E. K., Zheng, T., *et al.* (2016) Simultaneous achievement of high ethanol yield and titer in Clostridium thermocellum. *Biotechnol. Biofuels.* **9**, 116
7. Mazzoli, R., and Olson, D. G. (2020) Clostridium thermocellum: a microbial platform for high-value chemical production from lignocellulose. *Adv. Appl. Microbiol.* **113**, 111–161
8. Currie, D. H., Raman, B., Gowen, C. M., Tschaplinski, T. J., Land, M. L., Brown, S. D., *et al.* (2015) Genome-scale resources for Thermoanaerobacterium saccharolyticum. *BMC Syst. Biol.* **9**, 30
9. Shaw, A. J., Podkaminer, K. K., Desai, S. G., Bardsley, J. S., Rogers, S. R., Thorne, P. G., *et al.* (2008) Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. *Proc. Natl. Acad. Sci.* **105**, 13769–13774
10. Herring, C. D., Kenealy, W. R., Joe Shaw, A., Covalla, S. F., Olson, D. G., Zhang, J., *et al.* (2016) Strain and bioprocess improvement of a thermophilic anaerobe for the production of ethanol from wood. *Biotechnol. Biofuels.* **9**, 125
11. Hon, S., Holwerda, E. K., Worthen, R. S., Maloney, M. I., Tian, L., Cui, J., *et al.* (2018) Expressing the Thermoanaerobacterium saccharolyticum pforA in engineered Clostridium thermocellum improves ethanol production. *Biotechnol. Biofuels.* **11**, 242
12. Holwerda, E. K., Olson, D. G., Ruppertsberger, N. M., Stevenson, D. M., Murphy, S. J. L., Maloney, M. I., *et al.* (2020) Metabolic and evolutionary responses of clostridium thermocellum to genetic interventions aimed at improving ethanol production. *Biotechnol. Biofuels.* **13**, 40
13. Lynd, R., Beckham, L. T., Guss, G. M., Jayakody, A. N., Karp, L. M., Maranas, E., *et al.* (2022) Toward low-cost biological and hybrid biological/catalytic conversion of cellulosic biomass to fuels. *Energy Environ. Sci.* **15**, 938–990
14. Olson, D. G., Sparling, R., and Lynd, L. R. (2015) Ethanol production by engineered thermophiles. *Curr. Opin. Biotechnol.* **33**, 130–141
15. Lo, J., Olson, D. G., Murphy, S. J.-L., Tian, L., Hon, S., Lanahan, A., *et al.* (2017) Engineering electron metabolism to increase ethanol production in *Clostridium thermocellum*. *Metab. Eng.* **39**, 71–79
16. Brown, S. D., Guss, A. M., Karpinets, T. V., Parks, J. M., Smolin, N., Yang, S., *et al.* (2011) Mutant alcohol dehydrogenase leads to improved ethanol

tolerance in Clostridium thermocellum. *Proc. Natl. Acad. Sci.* **108**, 13752–13757
17. Shao, X., Raman, B., Zhu, M., Mielenz, J. R., Brown, S. D., Guss, A. M., *et al.* (2011) Mutant selection and phenotypic and genetic characterization of ethanol-tolerant strains of clostridium thermocellum. *Appl. Microbiol. Biotechnol.* **92**, 641–652
18. Olson, D. G., Maloney, M. I., Lanahan, A. A., Cervenka, N. D., Xia, Y., Pech-Canul, A., *et al.* (2023) Ethanol tolerance in engineered strains of Clostridium thermocellum. *Biotechnol. Biofuels Bioprod.* **16**, 137
19. Pech-Canul, A., Hammer, S. K., Ziegler, S. J., Richardson, I. D., Sharma, B. D., Maloney, M. I., *et al.* (2024) The role of AdhE on ethanol tolerance and production in *Clostridium thermocellum*. *J. Biol. Chem.* **300**, 107559
20. Chiarelli, D. P., Sharma, B. D., Hon, S., Bergamo, L. W., Lynd, L. R., and Olson, D. G. (2024) Expression and characterization of monofunctional alcohol dehydrogenase enzymes in *Clostridium thermocellum*. *Metab. Eng. Commun.* **19**, e00243
21. Tian, L., Perot, S. J., Stevenson, D., Jacobson, T., Lanahan, A. A., Amador-Noguez, D., *et al.* (2017) Metabolome analysis reveals a role for glyceraldehyde 3-phosphate dehydrogenase in the inhibition of C. thermocellum by ethanol. *Biotechnol. Biofuels.* **10**, 276
22. Dash, S., Olson, D. G., Joshua Chan, S. H., Amador-Noguez, D., Lynd, L. R., and Maranas, C. D. (2019) Thermodynamic analysis of the pathway for ethanol production from cellobiose in clostridium thermocellum. *Metab. Eng.* **55**, 161–169
23. Hon, S., Olson, D. G., Holwerda, E. K., Lanahan, A. A., Murphy, S. J. L., Maloney, M. I., *et al.* (2017) The ethanol pathway from *Thermoanaerobacterium saccharolyticum* improves ethanol production in *Clostridium thermocellum*. *Metab. Eng.* **42**, 175–184
24. Schmehl, M., Jahn, A., Meyer zu Vilsendorf, A., Hennecke, S., Masepohl, B., Schuppler, M., *et al.* (1993) Identification of a new class of nitrogen fixation genes in Rhodobacter capsulatus: a putative membrane complex involved in electron transport to nitrogenase. *Mol. Gen. Genet. MGG.* **241**, 602–615
25. Buckel, W., and Thauer, R. K. (2013) Energy conservation via electron bifurcating ferredoxin reduction and proton/Na+ translocating ferredoxin oxidation. *Biochim. Biophys. Acta BBA - Bioenerg.* **1827**, 94–113
26. Kuhns, M., Trifunović, D., Huber, H., and Müller, V. (2020) The Rnf complex is a Na+ coupled respiratory enzyme in a fermenting bacterium, Thermotoga maritima. *Commun. Biol.* **3**, 1–10
27. Vitt, S., Prinz, S., Eisinger, M., Ermler, U., and Buckel, W. (2022) Purification and structural characterization of the Na+-translocating ferredoxin: NAD+ reductase (Rnf) complex of clostridium tetanomorphum. *Nat. Commun.* **13**, 6315
28. Zhang, L., and Einsle, O. (2024) Architecture of the RNF1 complex that drives biological nitrogen fixation. *Nat. Chem. Biol.* **20**, 1078–1085
29. Wang, S., Huang, H., Moll, J., and Thauer, R. K. (2010) NADP+ reduction with reduced ferredoxin and NADP+ reduction with NADH are coupled via an electron-bifurcating enzyme complex in Clostridium kluyveri. *J. Bacteriol.* **192**, 5115–5123
30. Foulquier, C., Rivière, A., Heulot, M., Dos Reis, S., Perdu, C., Girbal, L., *et al.* (2022) Molecular characterization of the missing electron pathways for butanol synthesis in clostridium acetobutylicum. *Nat. Commun.* **13**, 4691
31. Nguyen, D. M. N., Schut, G. J., Zadvornyy, O. A., Tokmina-Lukaszewska, M., Poudel, S., Lipscomb, G. L., *et al.* (2017) Two functionally distinct NADP+-dependent ferredoxin oxidoreductases maintain the primary redox balance of Pyrococcus furiosus. *J. Biol. Chem.* **292**, 14603–14616
32. Lubner, C. E., Jennings, D. P., Mulder, D. W., Schut, G. J., Zadvornyy, O. A., Hoben, J. P., *et al.* (2017) Mechanistic insights into energy conservation by flavin-based electron bifurcation. *Nat. Chem. Biol.* **13**, 655–659
33. Kumar, A., Kremp, F., Roth, J., Freibert, S. A., Müller, V., and Schuller, J. M. (2023) Molecular architecture and electron transfer pathway of the Stn family transhydrogenase. *Nat. Commun.* **14**, 5484
34. Nakajima, M., Sakamoto, T., and Wada, K. (2002) The complete purification and characterization of three forms of ferredoxin-NADP+ oxidoreductase from a thermophilic cyanobacterium synechococcuselongatus. *Plant Cell Physiol.* **43**, 484–493
35. Liauw, P., Mashiba, T., Kopczak, M., Wiegand, K., Muraki, N., Kubota, H., *et al.* (2012) Cloning, expression, crystallization and preliminary X-ray studies of the ferredoxin-NAD(P)+ reductase from the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* **68**, 1048–1051
36. Ikeda, T., Nakamura, M., Arai, H., Ishii, M., and Igarashi, Y. (2009) Ferredoxin-NADP+ reductase from the thermophilic hydrogen-oxidizing bacterium, Hydrogenobacter thermophilus TK-6. *FEMS Microbiol. Lett.* **297**, 124–130
37. Tian, L., Lo, J., Shao, X., Zheng, T., Olson, D. G., and Lynd, L. R. (2016) Ferredoxin:NAD+ oxidoreductase of Thermoanaerobacterium saccharolyticum and its role in ethanol formation. *Appl. Environ. Microbiol.* **82**, 7134–7141
38. Markowitz, V. M., Chen, I.-M. A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., *et al.* (2012) IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res.* **40**, D115–D122
39. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410
40. Oberortner, E., Cheng, J.-F., Hillson, N. J., and Deutsch, S. (2017) Streamlining the design-to-build transition with build-optimization software tools. *ACS Synth. Biol.* **6**, 485–496
41. Wang, J., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, S., *et al.* (2023) The conserved domain database in 2023. *Nucleic Acids Res.* **51**, D384–D388
42. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., *et al.* (2023) InterPro in 2022. *Nucleic Acids Res.* **51**, D418–D427
43. Dobritzsch, D., Ricagno, S., Schneider, G., Schnackerz, K. D., and Lindqvist, Y. (2002) Crystal structure of the productive ternary complex of dihydropyrimidine dehydrogenase with NADPH and 5-iodouracil. Implications for mechanism of inhibition and electron transfer. *J. Biol. Chem.* **277**, 13155–13166
44. Aliverti, A., Pandini, V., Pennati, A., de Rosa, M., and Zanetti, G. (2008) Structural and functional diversity of ferredoxin-NADP+ reductases. *Arch. Biochem. Biophys.* **474**, 283–291
45. Shin, M. (1971) Ferredoxin-NADP reductase from spinach. In *Methods in Enzymology, Photosynthesis and Nitrogen Part A* **23**. Academic Press: 440–447
46. Bruns, C. M., and Karplus, P. A. (1995) Refined crystal structure of spinach ferredoxin reductase at 1.7 A resolution: oxidized, reduced and 2'-phospho-5'-AMP bound states. *J. Mol. Biol.* **247**, 125–145
47. Wise, C. E., Ledinina, A. E., Mulder, D. W., Chou, K. J., Peters, J. W., King, P. W., *et al.* (2022) An uncharacteristically low-potential flavin governs the energy landscape of electron bifurcation. *Proc. Natl. Acad. Sci.* **119**, e2117882119
48. Wise, C. E., Ledinina, A. E., and Lubner, C. E. (2022) Site-differentiated iron–sulfur cluster ligation affects flavin-based electron bifurcation activity. *Metabolites* **12**, 823
49. Yang, J., Lee, S. H., Ryu, J.-Y., Lee, H. S., and Kang, S. G. (2022) A novel NADP-dependent formate dehydrogenase from the hyperthermophilic archaeon thermococcus onnurineus NA1. *Front. Microbiol.* https://doi.org/10.3389/fmicb.2022.844735
50. Yang, X., and Ma, K. (2010) Characterization of a thioredoxin-thioredoxin reductase system from the hyperthermophilic bacterium Thermotoga maritima. *J. Bacteriol.* **192**, 1370–1376
51. Booth, W. T., Schlachter, C. R., Pote, S., Ussin, N., Mank, N. J., Klapper, V., *et al.* (2018) Impact of an N-terminal polyhistidine tag on protein thermal stability. *ACS Omega* **3**, 760–768
52. Hon, S., Jacobson, T., Stevenson, D. M., Maloney, M. I., Giannone, R. J., Hettich, R. L., *et al.* (2022) Increasing the thermodynamic driving force of the phosphofructokinase reaction in Clostridium thermocellum. *Appl. Environ. Microbiol.* **88**, e0125822
53. Ellis, L. D., Holwerda, E. K., Hogsett, D., Rogers, S., Shao, X., Tschaplinski, T., *et al.* (2012) Closing the carbon balance for fermentation by *Clostridium thermocellum* (ATCC 27405). *Bioresour. Technol.* **103**, 293–299

54. Cui, J., Stevenson, D., Korosh, T., Amador-Noguez, D., Olson, D. G., and Lynd, L. R. (2020) Developing a cell-free extract reaction (CFER) system in Clostridium thermocellum to identify metabolic limitations to ethanol production. *Front. Energy Res.* https://doi.org/10.3389/fenrg.2020.00072
55. Joe Shaw, A., Jenney, F. E., Adams, M. W. W., and Lynd, L. R. (2008) End-product pathways in the xylose fermenting bacterium, *Thermoanaerobacterium saccharolyticum*. *Enzyme Microb. Technol.* **42**, 453–458
56. Lo, J., Zheng, T., Olson, D. G., Ruppertsberger, N., Tripathi, S. A., Guss, A. M., *et al.* (2015) Deletion of nfnAB in Thermoanaerobacterium saccharolyticum and its effect on metabolism. *J. Bacteriol.* **197**, 2920–2929
57. Noor, E., Bar-Even, A., Flamholz, A., Reznik, E., Liebermeister, W., and Milo, R. (2014) Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLoS Comput. Biol.* **10**, e1003483
58. Tepper, N., Noor, E., Amador-Noguez, D., Haraldsdóttir, H. S., Milo, R., Rabinowitz, J., *et al.* (2013) Steady-state metabolite concentrations reflect a balance between maximizing enzyme efficiency and minimizing total metabolite load. *PLoS One* **8**, e75370
59. Corless, E. I., Saad Imran, S. M., Watkins, M. B., Bacik, J.-P., Mattice, J. R., Patterson, A., *et al.* (2021) The flexible N-terminus of BchL autoinhibits activity through interaction with its [4Fe-4S] cluster and released upon ATP binding. *J. Biol. Chem.* **296**, 100107
60. Liang, J., Huang, H., and Wang, S. (2019) Distribution, evolution, catalytic mechanism, and physiological functions of the flavin-based electron-bifurcating NADH-dependent reduced ferredoxin: NADP+ oxidoreductase. *Front. Microbiol.* **10**, 373
61. Poudel, S., Dunham, E. C., Lindsay, M. R., Amenabar, M. J., Fones, E. M., Colman, D. R., *et al.* (2018) Origin and evolution of flavin-based electron bifurcating enzymes. *Front. Microbiol.* https://doi.org/10.3389/fmicb.2018.01762
62. Huang, H., Wang, S., Moll, J., and Thauer, R. K. (2012) Electron bifurcation involved in the energy metabolism of the acetogenic bacterium moorella thermoacetica growing on glucose or H2 plus CO2. *J. Bacteriol.* **194**, 3689–3699
63. Demmer, J. K., Huang, H., Wang, S., Demmer, U., Thauer, R. K., and Ermler, U. (2015) Insights into flavin-based electron bifurcation via the NADH-dependent reduced ferredoxin:NADP oxidoreductase structure. *J. Biol. Chem.* **290**, 21985–21995
64. Rowland, P., Nørager, S., Jensen, K. F., and Larsen, S. (2000) Structure of dihydroorotate dehydrogenase B: electron transfer between two flavin groups bridged by an iron-sulphur cluster. *Structure* **8**, 1227–1238
65. Mohsen, A.-W. A., Rigby, S. E. J., Jensen, K. F., Munro, A. W., and Scrutton, N. S. (2004) Thermodynamic basis of electron transfer in dihydroorotate dehydrogenase B from Lactococcus lactis: analysis by potentiometry, EPR spectroscopy, and ENDOR spectroscopy. *Biochemistry* **43**, 6498–6510
66. Hagen, W. R., Silva, P. J., Amorim, M. A., Hagedoorn, P.-L., Wassink, H., Haaker, H., *et al.* (2000) Novel structure and redox chemistry of the prosthetic groups of the iron-sulfur flavoprotein sulfide dehydrogenase from Pyrococcus furiosus; evidence for a [2Fe-2S] cluster with Asp(Cys)3 ligands. *JBIC J. Biol. Inorg. Chem.* **5**, 527–534
67. Liu, H., Weisz, D. A., Zhang, M. M., Cheng, M., Zhang, B., Zhang, H., *et al.* (2019) Phycobilisomes harbor FNRL in cyanobacteria. *mBio* **10**, e00669-19
68. Iyanagi, T. (2019) Molecular mechanism of metabolic NAD(P)H-dependent electron-transfer systems: the role of redox cofactors. *Biochim. Biophys. Acta BBA - Bioenerg.* **1860**, 233–258
69. Mulo, P., and Medina, M. (2017) Interaction and electron transfer between ferredoxin–NADP+ oxidoreductase and its partners: structural, functional, and physiological implications. *Photosynth. Res.* **134**, 265–280
70. Sellés Vidal, L., Kelly, C. L., Mordaka, P. M., and Heap, J. T. (2018) Review of NAD(P)H-dependent oxidoreductases: properties, engineering and application. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1866**, 327–347
71. Runda, M. E., de Kok, N. A. W., and Schmidt, S. (2023) Rieske oxygenases and other ferredoxin-dependent enzymes: electron transfer principles and catalytic capabilities. *ChemBioChem.* **24**, e202300078
72. Spaans, S. K., Weusthuis, R. A., van der Oost, J., and Kengen, S. W. M. (2015) NADPH-generating systems in bacteria and archaea. *Front. Microbiol.* **6**, 742
73. Imran, S. M. S., Wiley, S. A., and Lubner, C. E. (2024) Electrochemistry of flavin-based electron bifurcation: 'Current' past and 'potential' futures. *Curr. Opin. Electrochem.* **47**, 101536
74. McInerney, M. J., Rohlin, L., Mouttaki, H., Kim, U., Krupp, R. S., Rios-Hernandez, L., *et al.* (2007) The genome of Syntrophus aciditrophicus: life at the thermodynamic limit of microbial growth. *Proc. Natl. Acad. Sci.* **104**, 7600–7605
75. Schuchmann, K., and Müller, V. (2014) Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* **12**, 809–821
76. Prakash, D., Chauhan, S. S., and Ferry, J. G. (2019) Life on the thermodynamic edge: respiratory growth of an acetotrophic methanogen. *Sci. Adv.* **5**, eaaw9059
77. Frolov, E. N., Elcheninov, A. G., Gololobova, A. V., Toshchakov, S. V., Novikov, A. A., Lebedinsky, A. V., *et al.* (2023) Obligate autotrophy at the thermodynamic limit of life in a new acetogenic bacterium. *Front. Microbiol.* **14**, 1185739
78. Biegel, E., Schmidt, S., González, J. M., and Müller, V. (2011) Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. *Cell. Mol. Life Sci. CMLS* **68**, 613–634
79. Buckel, W., and Thauer, R. K. (2018) Flavin-based electron bifurcation, ferredoxin, flavodoxin, and anaerobic respiration with protons (Ech) or NAD+ (Rnf) as electron acceptors: a historical review. *Front. Microbiol.* https://doi.org/10.3389/fmicb.2018.00401
80. Demmer, J. K., Pal Chowdhury, N., Selmer, T., Ermler, U., and Buckel, W. (2017) The semiquinone swing in the bifurcating electron transferring flavoprotein/butyryl-CoA dehydrogenase complex from Clostridium difficile. *Nat. Commun.* **8**, 1577
81. Hermoso, J. A., Mayoral, T., Faro, M., Gómez-Moreno, C., Sanz-Aparicio, J., and Medina, M. (2002) Mechanism of coenzyme recognition and binding revealed by crystal structure analysis of ferredoxin–NADP+ reductase complexed with NADP+. *J. Mol. Biol.* **319**, 1133–1142
82. Bossi, R. T., Aliverti, A., Raimondi, D., Fischer, F., Zanetti, G., Ferrari, D., *et al.* (2002) A covalent modification of NADP+ revealed by the atomic resolution structure of FprA, a Mycobacterium tuberculosis oxidoreductase. *Biochemistry* **41**, 8807–8818
83. Hammerstad, M., and Hersleth, H.-P. (2021) Overview of structurally homologous flavoprotein oxidoreductases containing the low *Mr* thioredoxin reductase-like fold – a functionally diverse group. *Arch. Biochem. Biophys.* **702**, 108826
84. Lesanavičius, M., Seo, D., and Čėnas, N. (2022) Thioredoxin reductase-type ferredoxin: NADP+ oxidoreductase of rhodopseudomonas palustris: potentiometric characteristics and reactions with nonphysiological oxidants. *Antioxidants* **11**, 1000
85. Komori, H., Seo, D., Sakurai, T., and Higuchi, Y. (2010) Crystal structure analysis of Bacillus subtilis ferredoxin-NADP+ oxidoreductase and the structural basis for its substrate selectivity. *Protein Sci.* **19**, 2279–2290
86. Olson, D. G., and Lynd, L. R. (2012) Transformation of Clostridium thermocellum by electroporation. *Methods Enzymol.* **510**, 317–330
87. Hogsett, D. A. L. (1995) *Cellulose Hydrolysis and Fermentation by Clostridium Thermocellum for the Production of Ethanol*. Ph.D. thesis, Dartmouth College
88. Takahashi, Y., and Nakamura, M. (1999) Functional assignment of the ORF2-iscS-iscU-iscA-hscB-hscA-fdx-ORF3 gene cluster involved in the assembly of Fe-S clusters in Escherichia coli. *J. Biochem. (Tokyo)* **126**, 917–926
89. Gauquelin, C., Baffert, C., Richaud, P., Kamionka, E., Etienne, E., Guieysse, D., *et al.* (2018) Roles of the F-domain in [FeFe] hydrogenase. *Biochim. Biophys. Acta BBA - Bioenerg.* **1859**, 69–77

90. Zanetti, G. (1981) The reduction of iodonitrotetrazolium chloride by ferredoxin-NADP+ reductase: a new tool for the characterization of the spinach chloroplast flavoprotein. *Plant Sci. Lett.* **23**, 55–61
91. Guss, A. M., Olson, D. G., Caiazza, N. C., and Lynd, L. R. (2012) Dcm methylation is detrimental to plasmid transformation in Clostridium thermocellum. *Biotechnol. Biofuels.* **5**, 30
92. Sharma, B. D., Hon, S., Thusoo, E., Stevenson, D. M., Amador-Noguez, D., Guss, A. M., *et al*. (2024) Pyrophosphate-free glycolysis in Clostridium thermocellum increases both thermodynamic driving force and ethanol titers. *Biotechnol. Biofuels Bioprod.* **17**, 146
93. Papadopoulos, J. S., and Agarwala, R. (2007) COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* **23**, 1073–1079
94. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190
95. Noor, E., Bar-Even, A., Flamholz, A., Lubling, Y., Davidi, D., and Milo, R. (2012) An integrated open framework for thermodynamics of reactions that combines accuracy and coverage. *Bioinforma. Oxf. Engl.* **28**, 2037–2044
96. Cui, J., Olson, D. G., and Lynd, L. R. (2019) Characterization of the *Clostridium thermocellum* AdhE, NfnAB, ferredoxin and Pfor proteins for their ability to support high titer ethanol production in *Thermoanaerobacterium saccharolyticum*. *Metab. Eng.* **51**, 32–42
97. Beber, M. E., Gollub, M. G., Mozaffari, D., Shebek, K. M., Flamholz, A. I., Milo, R., *et al*. (2022) eQuilibrator 3.0: a database solution for thermodynamic constant estimation. *Nucleic Acids Res.* **50**, D603–D609
98. Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., *et al*. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285