

15 February 2026

Beyond Canonical CO Oxidation: First Characterisation of a Clade B CODH using a New Anaerobic Cryo-EM Workflow

Maximilian Böhm¹, Vivek Srinivas², Benjamin Wiseman², Ping Huang¹, Moritz Senger³, Martin Högbom², Henrik Land¹

1. Department of Chemistry -Ångström Laboratory Uppsala University

2. Department of Biochemistry and Biophysics and Science for Life Laboratory Stockholm University

3. Department of Chemistry for Life Sciences Uppsala University

Abstract

Carbon monoxide dehydrogenases (CODHs) catalyse the reversible oxidation of CO to CO₂ and play central roles in microbial carbon metabolism. While well-characterised CODHs from different phylogenetic backgrounds exhibit high bidirectional activity, the enigmatic clade B remains functionally uncharacterised. Here we present the first structural and functional characterisation of a clade B CODH from *Ruminococcus flavefaciens* (RfCODH). It reveals striking divergence from canonical enzymes. A new anaerobic cryo-EM workflow was developed, carried out entirely under anoxic conditions by manual blotting and plunge freezing. It resulted in a 2.53 Å RfCODH structure. The structure adopts the typical CODH fold, but exhibits blocked gas channels, a compromised proton transfer pathway and disrupted cofactor coordination. This provides a structural rationale for RfCODH's severely attenuated CO oxidation activity (13 mU/mg vs. 900 U/mg for the well-studied ChCODH-II). EPR spectroscopy reveals unique oxidised C-cluster states not previously characterised in CODHs. Mirror tree analysis hints to co-evolution between clade B CODHs and associated ABC transporter substrate-binding proteins, suggesting these enzymes function in metabolism of substrates imported via the ABC transporter module. All findings indicate evolutionary repurposing of the CODH scaffold for alternative physiological functions.

Beyond Canonical CO Oxidation: First Characterisation of a Clade B CODH using a New Anaerobic Cryo-EM Workflow

Maximilian Böhm^[a], Vivek Srinivas^[b], Benjamin Wiseman^[b], Ping Huang^[a], Moritz Senger^[c], Martin Högbom^[b], and Henrik Land^{[a],*}

[a] Maximilian Böhm, Dr. Ping Huang, Dr. Henrik Land
Department of Chemistry – Ångström Laboratory
Uppsala University
Regementsvägen 10, 752 37 Uppsala, Sweden
E-mail: henrik.land@kemi.uu.se

[b] Dr. Vivek Srinivas, Prof. Dr. Martin Högbom
Department of Biochemistry and Biophysics and Science for Life Laboratory
Stockholm University
Svante Arrhenius väg 16, 106 91 Stockholm, Sweden

[c] Dr. Moritz Senger
Department of Chemistry for Life Sciences
Uppsala University
Husargatan 3, 752 37 Uppsala, Sweden

*Corresponding Author: henrik.land@kemi.uu.se

Supporting information is given in a separate file.

Abstract: Carbon monoxide dehydrogenases (CODHs) catalyse the reversible oxidation of CO to CO₂ and play central roles in microbial carbon metabolism. While well-characterised CODHs from different phylogenetic backgrounds exhibit high bidirectional activity, the enigmatic clade B remains functionally uncharacterised. Here we present the first structural and functional characterisation of a clade B CODH from *Ruminococcus flavefaciens* (RfCODH). It reveals striking divergence from canonical enzymes. A new anaerobic cryo-EM workflow was developed, carried out entirely under anoxic conditions by manual blotting and plunge freezing. It resulted in a 2.53 Å RfCODH structure. The structure adopts the typical CODH fold, but exhibits blocked gas channels, a compromised proton transfer pathway and disrupted cofactor coordination. This provides a structural rationale for RfCODH's severely attenuated CO oxidation activity (13 mU/mg vs. 900 U/mg for the well-studied ChCODH-II). EPR spectroscopy reveals unique oxidised C-cluster states not previously characterised in CODHs. Mirror tree analysis hints to co-evolution between clade B CODHs and associated ABC transporter substrate-binding proteins, suggesting these enzymes function in metabolism of substrates imported via the ABC transporter module. All findings indicate evolutionary repurposing of the CODH scaffold for alternative physiological functions.

Introduction

Nickel-iron carbon monoxide dehydrogenases (CODHs) catalyse the interconversion of CO and CO₂, playing crucial roles in microbial carbon metabolism and energy conservation.^[1] These enzymes are homodimeric and carry five metal clusters: two buried [4Fe4S] B clusters and one [4Fe4S] D cluster on the interface of the homodimer; substrate conversion takes place at the two biologically unique [Ni3Fe4S] C clusters with an additional Fe coordinated in proximity. These clusters are buried deep within the protein.^[1,2] Substrate access is regulated by proton transfer pathways mainly consisting of charged histidine residues^[3], and hydrophobic gas channels^[4].

Much of the biochemical diversity within the CODH family remains largely unexplored. While substantial progress has been made in understanding important features such as structure, catalytic mechanism and oxygen sensitivity, these studies only used CODH from phylogenetic clades A^[5], E^[6] and F^[7,8]. Meanwhile, clades C^[9] and D^[10] have only one biochemically characterised representative each, and clades B, G, and H lack any functional characterisation despite their widespread distribution in microbial genomes^[11,12]. Particularly eye opening was the study of the first clade D CODH from *Carboxydotherrmus hydrogenoformans*, ChCODH-V^[10]. It showed an altered nickel-free C cluster and no CO₂ reduction, nor CO oxidation activity.^[10] Only hydroxylamine could be identified as a substrate^[10], suggesting clade D “CODH” possess a different role in the organism than canonical CODHs. This supported the claims made by bioinformatic studies, that different clades serve different functions^[12,13] and hence perform different chemical reactions.

Clade B CODHs are particularly enigmatic for a different reason. Firstly, they are primarily found in symbiotic host-associated environments^[11]. Secondly their genomic context shows striking differences from canonical CODHs: the majority of clade B CODH are encoded within operons containing complete ABC transporter systems, comprising transmembrane domains (TMs), nucleotide-binding domains (NBDs), and substrate-binding proteins (SBPs)^[12]. Additionally, many clade B operons contain two conserved hypothetical proteins whose functions remain unknown but appear across diverse organisms including bacteria from the phyla *Bacillota* and *Pseudomonadota*.

The presence of a CODH in *Ruminococcus flavefaciens* was noted as puzzling as early as 2004, when Antonopoulos and colleagues observed its gene^[14]. Even though it was known that *R. flavefaciens* shows increased growth on high carbonate^[15], the presence of a CODH gene does not align with what is known about the overall metabolism of this gut microbe. This observation hints

at a broader question: do clade B CODHs actually function as carbon monoxide dehydrogenases, or has this enzyme scaffold been repurposed for alternative chemistry? Much like the clade D CODH, *ChCODH-V*, with an until now unknown function and an altered C cluster.^[10]

Here we present the first biochemical and structural characterisation of a clade B CODH from *Ruminococcus flavefaciens* (*RfCODH*). It is encoded in an operon encompassing the canonical clade B architecture: the CODH gene flanked by ABC transporter components and two conserved hypothetical proteins. Through a combination of solution-based activity assays, electron paramagnetic resonance (EPR) and FTIR spectroscopy, cryo-electron microscopy (cryo-EM), and phylogenetic co-evolution analysis, we demonstrate that *RfCODH* exhibits severely attenuated CO oxidation activity, which is several orders of magnitude lower than characterised clade F enzymes. *RfCODH* furthermore lacks detectable CO₂ reduction capability. For structure determination, we developed a new anaerobic grid preparation workflow using simple cross-locking tweezers for manual blotting and manual plunge freezing into liquid ethane, with grids stored submerged in the ethane cup, avoiding any introduction of liquid nitrogen into the glovebox. The 2.53 Å cryo-EM structure reveals that while *RfCODH* maintains the overall CODH fold and accessory metal cluster architecture, its substrate access pathways are systematically occluded by bulky hydrophobic residues and its proton transfer pathway is altered. We were also not able to fully resolve the active site C-cluster, likely due to disorder or alternate states, but the nature of the coordinating residues suggests a non-canonical conformation. Bioinformatic mirror tree analysis hints towards co-evolution between clade B CODHs and their associated SBPs, suggesting functional coupling. Together, these findings indicate that clade B CODHs represent a functionally divergent lineage that has retained the versatile CODH/hybrid cluster protein scaffold while evolving toward alternative physiological roles.

Results and Discussion

Activity of *RfCODH*

RfCODH, tagged with a C-terminal Strep-tag II, expresses well and is readily purified from *E. coli* cell lysate. Figure 1A shows an SDS-PAGE analysis with a clear band at the expected molecular weight of 79 kDa for *RfCODH*. The metal content after purification was 6–8 Fe/monomer and 0.2–0.5 Ni/monomer, determined via ICP-OES and TXRF; for full metalation of a canonical CODH we expect 10 Fe/monomer and 1 Ni/monomer. The protein exhibits a low activity towards CO oxidation, but only after preincubation in CO saturated buffer. CO oxidation was tested over a range of pH-values (Figure 1B) and *RfCODH*'s activity increases with higher pH, which has been observed for other CODHs as well.^[16] When *RfCODH* is incubated with CO alone we observed a decrease in absorbance between 350 and 400 nm (Figure 1D) which is commonly seen in the reduction of [FeS] cluster proteins, potentially suggesting a reductive activation of the enzyme. However, preincubation with sodium dithionite (NaDT) did not activate the enzyme. We could not observe any CO₂ reduction activity under standard CO evolution assay conditions, using methyl viologen and NaDT as electron mediator and donor, respectively.^[6] Substituting electron mediator and donor to dmDQ-H (3,11-dimethyl-7,8-dihydro-6*H*-dipyrido[1,2-*a*:2',1'-*c*][1,4]diazepine-5,9-diium) and europium EGTA (ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid) to increase overall driving force^[17], also did not result in any observable activity. Auxiliary activities towards nitrate and hydroxylamine (NH₂OH) were also tested using a methyl viologen oxidation assay. Only NH₂OH seems to be a substrate for *RfCODH* as it yielded significant methyl viologen oxidation (Figure 1C). We tested for CO₂ and carbonate reduction that does not yield CO in the process, with no observable activity (Figure 1C). Compared to other CODHs, CO oxidation activity is three to five orders of magnitude lower (Table 1). One reason could be that *RfCODH* was not fully matured in the expression host, which would explain the low Ni/monomer ratio. However, an alternative reason could be that CO is not the primary substrate for clade B CODHs. The NH₂OH reduction activities are higher compared to the prototypical clade F CODH, *ChCODH-II*, at the same temperature, but low when compared to the reported values of prototypical hybrid cluster protein from *Pyrococcus furiosus* (*PfHCP*) and clade F and D CODH, *ChCODH-II* and *ChCODH-V*.^[10,18–21] Even though the reported values reflect the activity at elevated temperatures, this makes NH₂OH a rather questionable substrate. Therefore, the optimum conditions for this enzyme, and its primary substrate are yet to be identified.

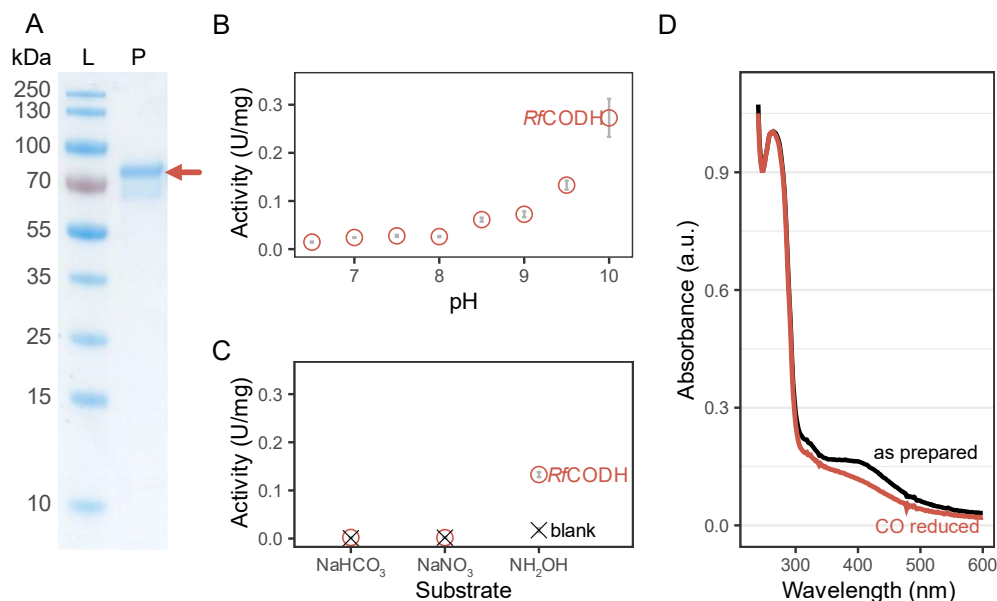


Figure 1. Biochemical characterisation of *RfCODH*. (A) SDS-PAGE analysis of purified *RfCODH* (P) showing a clear band for *RfCODH* at 79 kDa with PageRuler Plus Protein Ladder (L, ThermoFisher 26619). For uncut gel see supplementary information Figure S1. (B) pH dependency of *RfCODH* CO oxidation activity. One unit (U) is defined as the reduction of one μmol of methyl viologen (MV) per minute. (C) Auxiliary activities of *RfCODH* towards NaNO_3 , NaHCO_3 , and NH_2OH , measured in 50 mM TAPS-NaOH, 20 mM NaCl, pH 8. One U is defined as one μmol of MV oxidised per minute. (D) UV-vis spectra of *RfCODH* with buffer (50 mM Tris-HCl, 20 mM NaCl, pH 8) (black) and CO-saturated buffer (red).

Table 1. Activity of CODH towards selected substrates. One U is defined as the conversion of one μmol of substrate per minute at pH 8, room temperature, and MV as electron shuttle.

Name	Clade	Activity in U/mg			Ref.
		CO oxidation	CO ₂ reduction	NH ₂ OH reduction	
<i>RfCODH</i> ^[a]	B	0.013 ± 0.007	n.a.	0.066 ± 0.003	This study
<i>TkCODH-II</i> ^[b]	C	13.9 ± 4.1	n.r.	n.r.	[9]
<i>ChCODH-II</i> ^[a]	F	900 ± 28.2	16.9 ± n.r.	0.7 ^[c]	[18,19,21]
		280 ± 75.9	1.18 ± 0.026	0.009 ± 0.002	This study
<i>ChCODH-IV</i> ^[a]	F	85 ± 1.2	n.r.	n.r.	[19]
<i>ChCODH-V</i> ^{[a][c]}	D	n.a.	n.a.	0.11 ± 0.01	[10]
<i>PfHCP</i> ^[d]		n.r.	n.r.	4.5 ± 0.2	[20]

Not reported (n.r.), not active (n.a.).

[a] Recombinantly produced protein; [b] activity measured at 66 °C; [c] measured at 40 °C; [d] activity measured at 70 °C and pH 9.

Electron Paramagnetic Resonance Spectroscopy reports on functional *RfCODH* cofactors

Electron paramagnetic resonance (EPR) spectroscopy was employed to study the electronic state of *RfCODH*'s metal clusters. *RfCODH* was incubated under different redox conditions to yield the following EPR spectra (Figure S2 and Figure 2). Samples reduced with dithiothreitol (DTT), incubated with carbonate buffer or buffer with no additional substrates/redox substances were in a reduced state and show no differences in redox states of the clusters (compare spectra "As isolated" and "DTT reduced", Figure S2). From this we conclude that our reductive purification procedure keeps *RfCODH* in a reduced state, that cannot be further enriched with prolonged DTT exposure. Furthermore, carbonate (and hence CO₂) was not able to oxidise the enzyme (compare spectra "As isolated" and "CO₂ exposed", Figure S2), which is in line with the lack of activity towards CO₂ reduction we demonstrated above. All EPR signals from various species exhibit rhombic character. EPR-active redox states are referred to by their associated metal cluster.

When *RfCODH* is oxidised with thionine, we detected a spectrum which has not previously been characterised for CODH (Figure 2A and "Thionine oxidised" in Figure S2). Simulation suggests that this spectrum comprises two rhombic components that most likely stem from an alternated oxidised C cluster, and are therefore annotated as C_{alt-ox} (C_{alt-ox1}, $g_{av} = 1.984$ [2.0216, 2.0105, 1.9187] and C_{alt-ox2}, $g_{av} = 1.983$ [2.0107, 1.9743, 1.9626]). Svetlitchnyi *et al.* (2001) observed a similar feature for *ChCODH-I* and *ChCODH-II* after exposure to oxygen.^[8] But, the resolution of their spectra is too low to make meaningful comparisons. Interestingly, even though we determined a lower iron content than would be expected for a canonical CODH, no obvious indication of damaged FeS clusters could

be observed in EPR, e.g., [3Fe4S] from degraded [4Fe4S]. The B and D clusters are in an oxidised EPR-silent state under these conditions and do not contribute to the observed spectra.

However, when *RfCODH* is reduced with NaDT, EPR spectroscopy shows features of two overlapping signals in the $g \sim 2.0$ region that match with previously described signals from the B cluster (B_{red1} , $g_{av} = 1.974$ [2.0573, 1.9561, 1.9080] and B_{red2} , $g_{av} = 1.976$ [2.0617, 1.9543, 1.9113])^[8,22–27] (Figure 2B and S2). A signal with $g_{av} = 2.232$ [2.282, 2.213, 2.200], typical for $d^9 Ni^+$, appears alongside with other signals from B or C-clusters (Figure 2B “difference”). This signal closely resembles the one previously reported by Basak *et al.*^[22], though the signal reported here shows less rhombicity. It probably stems from a smaller population of Ni^+ from an independently bound Ni^+ or an alternated reduced C-cluster. We therefore denote it as $C_{alt-red}$. Even though it is likely that a Ni^+ gives rise to this $g_{av} \sim 2.2$ signal, we could not find a convincing fit, assuming a single spin $\pm 1/2$ transition. To further investigate this issue, the signal amplitudes (i.e., $d\chi''/dB$, as an indicator of its susceptibility, χ), were measured under varied temperature and microwave power to yield Curie plots of two variations (Figure S3C, $\chi \sim 1/T$ and S3D, $\chi T \sim T$). The non-linear correlation in both plots demonstrates spin transitions from a non-“ground state only” behaviour even at temperature as low as 5 K. This result reinforces the assignment of an alternative C-cluster - C_{alt} .

After incubation with CO, *RfCODH*'s EPR signal changes to also include a new feature in the lower $g \sim 1.7$ region (Figure 2C and S2) and additional intensity in the $g \sim 2.0$ region. Simulation indicates that the signal is composed of three species, denoted as C_{red1} , B_{red1} and B_{red2} . The C_{red1} spectrum matches closely with the previously described reduced C cluster with a coordinated CO (C_{red1} , $g_{av} = 1.877$ [2.0300, 1.8970, 1.7054]).^[22,25–27] Power temperature studies on this signal can be seen in Figure S4. Simulations also show that the B_{red} signals shift slightly which is indicated with an * (B_{red*1} , $g_{av} = 1.973$ [2.0551, 1.9556, 1.9091] and B_{red*2} , $g_{av} = 1.977$ [2.0612, 1.9553, 1.9148]), which can be rationalised as interaction with the now reduced C cluster. The g -values for all reported states are summarised in Table S1.

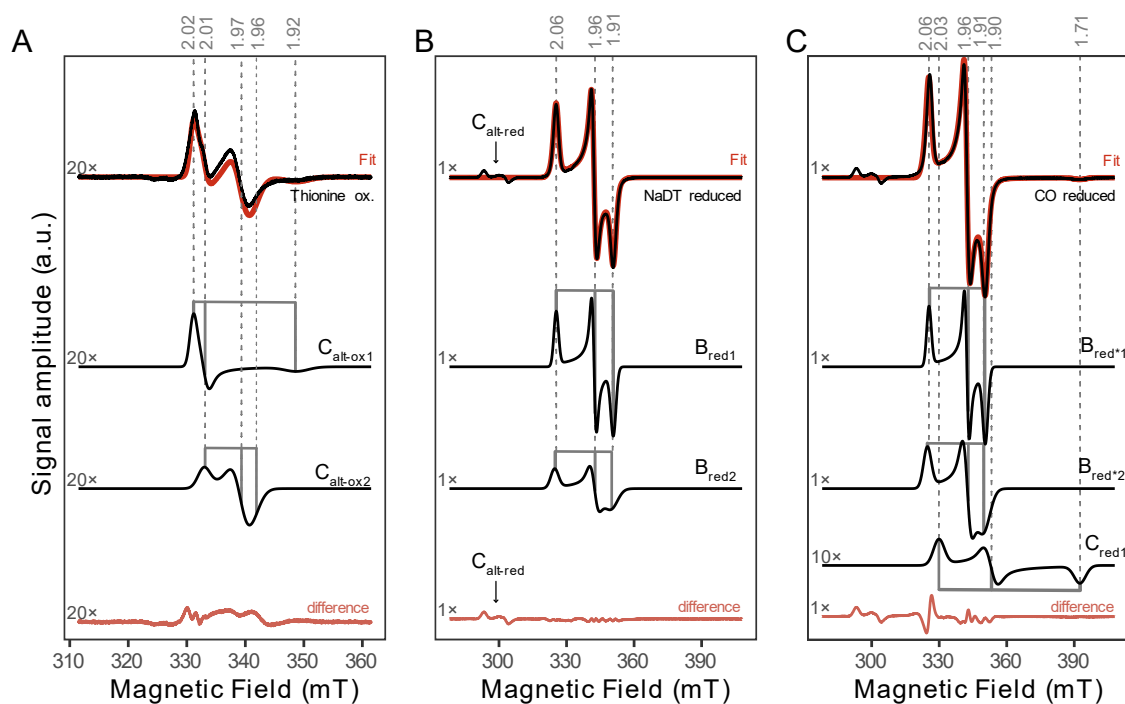


Figure 2. EPR spectra and simulations of *RfCODH*. Experimental and simulated EPR spectra for (A) Thionine oxidised *RfCODH*. (B) NaDT reduced *RfCODH*. (C) CO reduced *RfCODH*. Measured at 15 K, 15 G, 100 kHz, 9.37 GHz, and 2.5 mW.

Cyanide Probing Using IR Spectroscopy

To better understand the nature of the C cluster in *Rf*CODH and to compare it to canonical CODHs, such as *Ch*CODH-II, we used cyanide (CN⁻) as an infra-red (IR) sensitive probe. This has previously been shown to strongly inhibit the C-cluster by binding to the nickel, leading to a distinct signal that can be used to probe the presence of a canonical C-cluster.^[28] We incubated each, *Rf*CODH and *Ch*CODH-II, with potassium cyanide to generate cyanide bound forms following an established protocol.^[28] The resulting IR spectra in the CN⁻ wavenumber region can be seen in Figure 3. *Ch*CODH-II showed an absorption at 2110 cm⁻¹ as described previously for a C cluster with CN⁻ ligand bound to the nickel.^[28] An identical, but less intense signal emerged after *Rf*CODH was treated with potassium cyanide even though the enzyme concentration was approximately double (Figure 3). The lower intensity can have different reasons. Firstly, that only a fraction of *Rf*CODH has a typical C cluster. Secondly, that only a limited amount of CN⁻ reaches the active site. Thirdly, that the extinction coefficient for this C≡N vibration is lower in *Rf*CODH due to differences in the active site coordination sphere. Lastly, our *Rf*CODH preparation only incorporated 0.2 - 0.5 Ni/monomer, providing limited sites for cyanide to bind, which in turn leads to a weaker signal. Overall, we can conclude that some *Rf*CODH in our sample contains a classical C cluster, containing a nickel with an open coordination site accessible to cyanide as indicated by the identical cyanide band position. when compared to *Ch*CODH-II.

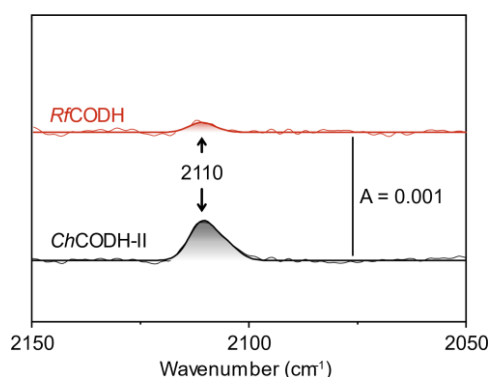


Figure 3. FTIR spectra of CN⁻-bound CODHs. *Rf*CODH (red, above) and *Ch*CODH-II (black, below) were incubated with CN⁻ following Ciaccavava *et al.*^[28] Enzymes were treated with potassium cyanide in Tris-HCl buffer containing NaCl and NaDT, concentrated, and measured under anoxic conditions. Reported CN⁻ vibrational frequency (2110 cm⁻¹) bound to C clusters is indicated with arrows.^[28] Note the lower absorbance signal for *Rf*CODH present in approx. double the concentration of *Ch*CODH-II. Bold traces show fits, thin traces background-subtracted data.

Structural Analysis of *Rf*CODH

To study the structural basis for *Rf*CODH's unusual catalytic properties, we determined its structure using cryo-electron microscopy to an overall resolution of 2.53 Å (Figure 4). Sample application and vitrification of the grids was performed in a glovebox. We developed this new anaerobic grid preparation workflow using solidified ethane to manage transfers through the glovebox antechamber and a deep well to maintain ethane at cryogenic temperature. This enables preparation of grids suitable for cryo-EM studies of oxygen sensitive metalloproteins requiring minimum modifications to existing anaerobic gloveboxes. This method does not require dedicated plunge freezing robots inside anaerobic chambers^[29,30] or construction of custom anoxic chambers around such robots^[31]. It maintains strict anaerobicity by transferring only solidified ethane into the glovebox while keeping liquid nitrogen external. This approach simultaneously reduces the risk of any gaseous expansion of liquid nitrogen inside the enclosure and removes the risk of oxygen contamination from foam dewars. Also, in contrast to blot-free or chemically protected workflows^[32], this protocol does not rely on oxygen scavengers such as dithionite to preserve metalloprotein integrity, thus avoiding unintended chemical reduction of other oxidised cofactor sites. The main limitation of this setup is manual blotting, which reduces reproducibility and may require additional grid screening compared to the protocols discussed above. The detailed setup (Figure S5) and workflow description are presented in the SI. Data collection and refinement workflow, as well as statistics are summarised in Figure S6 and Table S2.

*Rf*CODH shows a typical CODH fold and homodimeric structure (RMSD of 1.209 compared to *Ch*CODH-II, PDB ID: 4UDX) with well-resolved D and B clusters. Notably, *Rf*CODH contains an additional N-terminal extension consisting of a histidine rich 35 amino acid long region (not resolved in the structure, due to its unordered nature), a small beta sheet, and an alpha helix (orange in Figure 4A). The helix is connected to the sheet via a structurally unresolved loop of 12 amino acids (L2). These elements are not present in other CODHs, though the functional significance of them remains unclear. An additional loop (L1) could be identified, not present in other characterised CODHs, which will be discussed in more detail later. The current structural resolution within the active site was not resolved enough to model a C cluster confidently (Figure 4B). This could either be due to a damaged protein during preparation, sample heterogeneity or due to alternate conformations of the C cluster. Figure 4C shows the sequence logo of cluster coordinating residues for clades B and F to compare sequence conservation between clades. As can be seen, C cluster coordination differs mainly at position 415, where in canonical clade F CODHs a thiol cysteine coordinates the [Ni3Fe4S] cluster (Cys333 in *Ch*CODH-II). In almost all clade B enzymes, this is exchanged to a carboxylic acid residue (glutamate or aspartate, Asp415 in *Rf*CODH). B and D cluster coordination is conserved between the two clades.

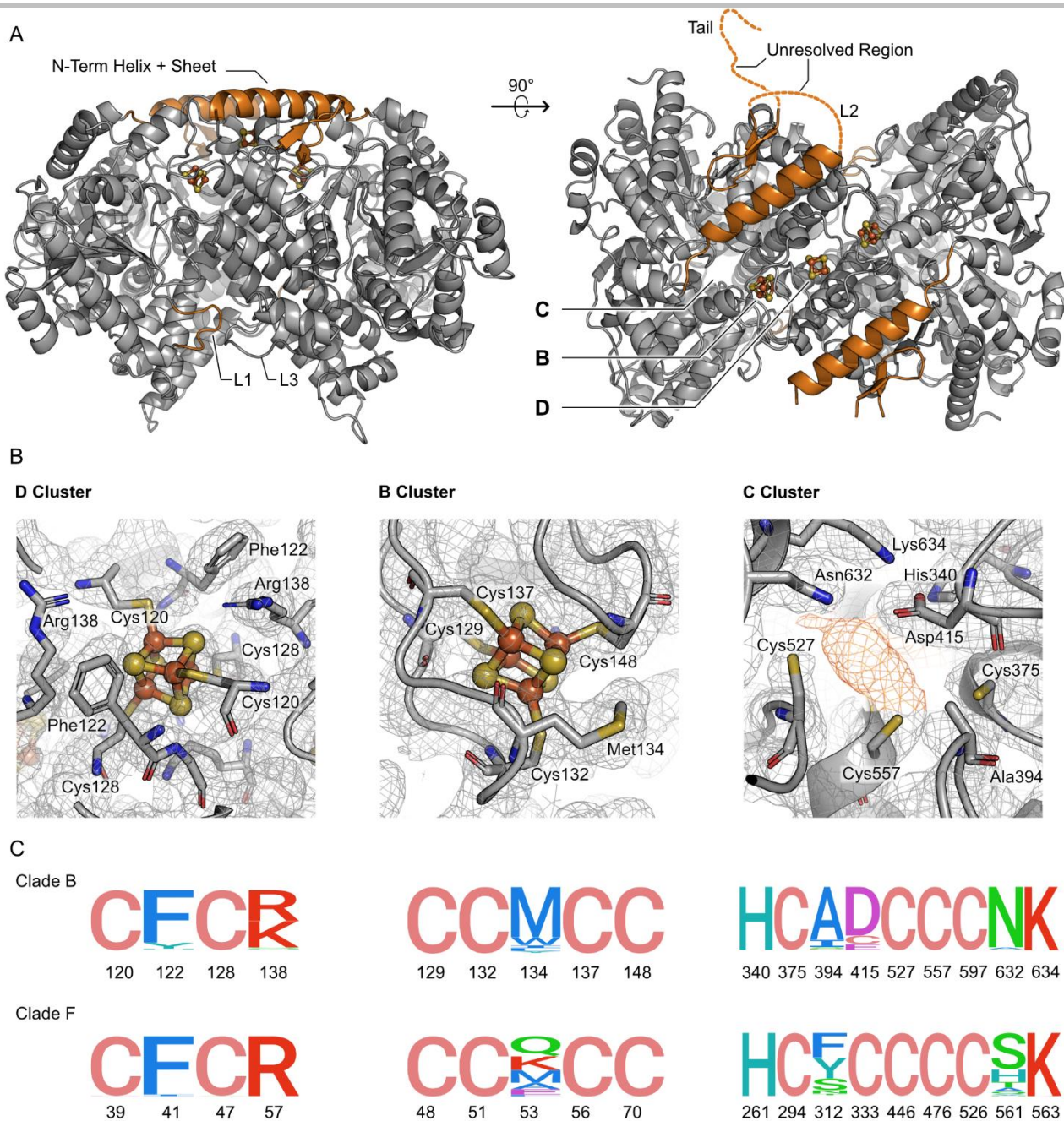


Figure 4. Structural overview of cryo-EM structure of *Rf*CODH. (A) Top and side view of *Rf*CODH, with non-canonical structural elements marked in orange, compared to clade F CODHs. (B) D, B and C cluster regions in *Rf*CODH. (C) Sequence logos of selected residues from cluster regions of clade B and clade F CODH, with *Rf*CODH and *Ch*CODH-II numbering, respectively. The resolution distribution of the final *Rf*CODH volume can be seen in Figure S7. More detailed predicted volumes of ligand binding sites and chain differences are presented in Figure S8. DeepEMhancer^[33] processed map was used for figures, a non-uniform refinement map was used for model building.

Despite the overall tertiary and quaternary structural conservation, closer examination reveals critical differences in substrate access pathways that rationalise *Rf*CODH's severely attenuated CO oxidation activity. We analysed the potential gas channels with CAVER 3.0^[34] and two different probe radii of 1.1 Å and 0.9 Å. We used Biester *et al.*'s structure from *Nitratidesulfovibrio vulgaris* CODH^[4] (*Nv*CODH, formerly known as *Dv*CODH) to determine a probe radius that could replicate the gas channels observed in their structure based on xenon soaked crystals (PDB ID: 7TSJ). We found that a probe radius of 1.1 Å replicates the data, which we used for the primary analysis (Figure 5). We also analysed *Rf*CODH with the commonly used 0.9 Å probe radius (Figure S9). The results demonstrate that all but one of the canonical CO entry routes identified in *Nv*CODH are blocked in *Rf*CODH by bulky hydrophobic residues (see Figure S9 and S10). The one non-blocked canonical entry point is slightly shifted and smaller compared to *Nv*CODH, due to steric hindrance by Phe666 (Figure S9). We identified an alternative tunnel in *Rf*CODH leading directly to the presumed C-cluster active site (Figure 5), which has not been seen in other CODHs before. This tunnel also connects to a cavity beneath the D-cluster, similar to what was predicted by others for different CODHs using CAVER^[4]. The architectural rearrangement of the gas channel suggests that if *Rf*CODH catalyses CO oxidation, substrate access must occur through a significantly more restricted pathway than in

canonical CODHs, which usually have multiple entry and exit points. *Rf*CODH has some small gas channels predicted with small probe radius (see Figure S9), however, it contains only one bigger entry and exit point for substrate access to the active site (Figure 5).

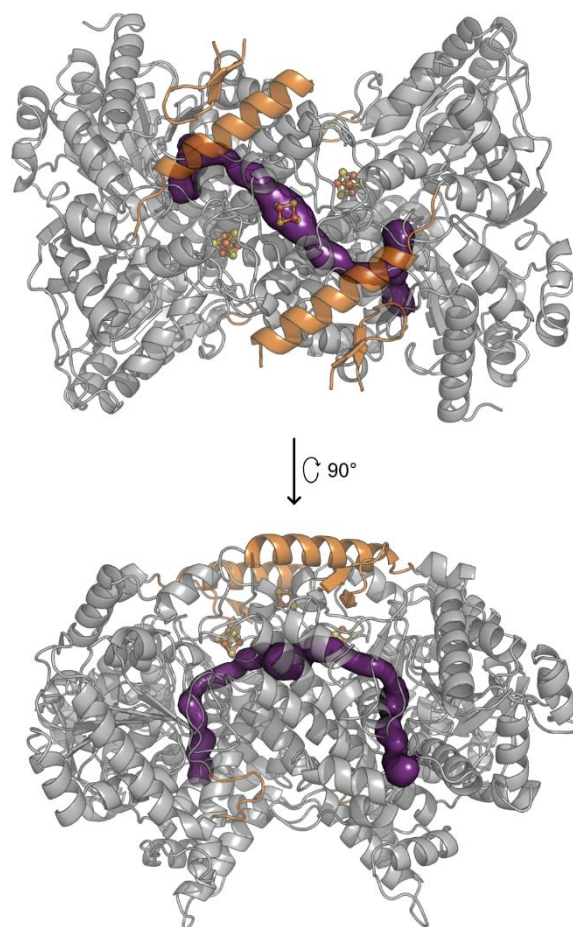


Figure 5. Gas channel of *Rf*CODH. The gas channel of *Rf*CODH (purple) was predicted with CAVER 3.0^[34] with a probe radius of 1.1 Å.

Equally striking are the differences in the proton transfer pathway. In clade F CODHs, a chain of histidine residues provides a proton relay from the protein surface to the active site, essential for coupling CO oxidation to proton release.^[3] In *Rf*CODH, this pathway appears altered. First, the aforementioned additional surface loop L1 close to the surface of *Rf*CODH harbours three histidine residues and represents a putative new entry point (Figure 6). Another loop (L3), present in *Ch*CODH-II but extended in *Rf*CODH, occludes the entry point known for *Ch*CODH-II. Second, sequence analysis reveals that clade B CODHs contain two conserved substitutions (Glu177, Leu180 in *Rf*CODH) at the terminal positions of the clade F histidine chain (His99, His102, *Ch*CODH-II numbering, Figure 6), that are described by Kim *et al.* to be non-essential, but modulate CO oxidation activity greatly.^[3] Thirdly, another amino acid mentioned by Kim *et al.* to be responsible for activity attenuation (Asn262, *Ch*CODH-II numbering) is consistently exchanged to Ser or Leu in clade B CODHs (Leu341 in *Rf*CODH). These structural variations of the proton transfer pathway in *Rf*CODH and all clade B CODHs are expected to alter proton transfer efficiency. It is conceivable that the three-histidine loop could adopt alternative conformations under specific conditions (pH, binding of metal ions, temperature, ...), enhancing or restricting proton transfer. However, our structure provides no evidence for such flexibility.

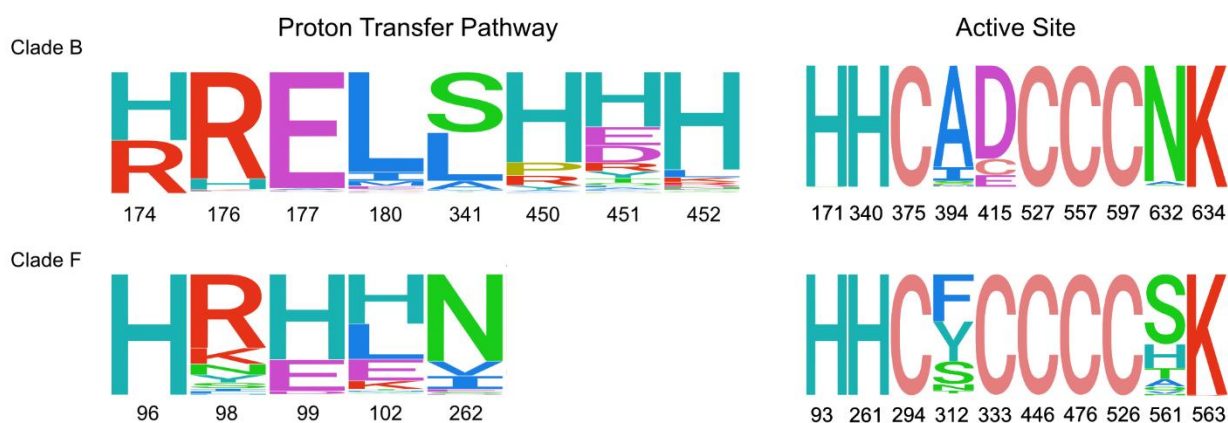
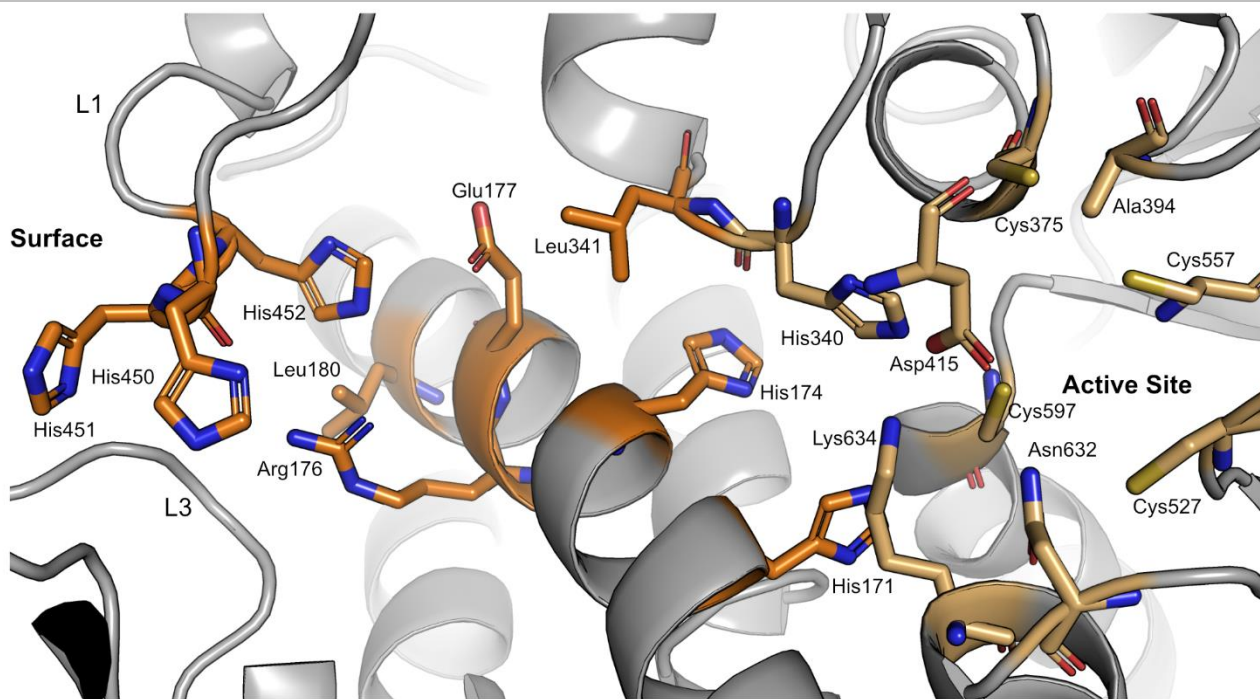


Figure 6. Putative proton transfer pathway. Top panel. Structure of putative proton transfer pathway (orange residues) and active site (gold residues) in *RfCODH* (with *RfCODH* numbering). Lower Panel. Sequence logo for proton transfer pathway and active site for clade B CODH and clade F CODH with *RfCODH* and *ChCODH-II* numbering, respectively. For detailed view of the putative proton transfer pathway of *ChCODH-II* see Figure S11.

The D-cluster, typically serving as the gateway for electron transfer, shows canonical [4Fe-4S] architecture with well-resolved density. Similarly, the B-cluster, which serves as an electron relay between the D and the C cluster, exhibits the expected [4Fe-4S] cubane structure, as supported by our EPR spectroscopy. However, the C-cluster region, despite reasonable density, shows features that may explain the unusual EPR signals we observe. As mentioned earlier, in canonical CODH the [Ni3Fe4S] cluster is coordinated by four cysteines. In clade B CODH this coordination is altered to exchange one cysteine to a carboxylic acid, most often an aspartate (Asp415 in *RfCODH*). We believe, that it is this exchange that drastically alters the C cluster. In our structure we find that the aspartate reaches too far into the space where we would expect the C cluster (Figure 4B), if it would have the canonical position and coordination. An alternative geometry of the cluster would therefore be necessary to accommodate Asp415.

The structural data provide a compelling rationale for *RfCODH*'s three-orders-of-magnitude lower CO oxidation activity compared to canonical CODHs. The enzyme appears to be a monodirectional catalyst, as evidenced by our inability to detect CO₂ reduction activity and the structural barriers. Whether CO oxidation represents a vestigial activity or whether *RfCODH* processes CO under specific physiological conditions that overcome these structural constraints remains an open question. The possibility that clade B CODHs utilise entirely different substrates, delivered through the ABC transporter system, represents an attractive alternative hypothesis that warrants further investigation.

Co-evolution Analysis of Clade B CODH Operons

Our previous work and that of others showed clearly that clade B CODH co-occur to a high probability with ABC transporter units.^[12,13] So is the case for *Rf*CODH, and is illustrated in Figure 7A. Its operon exists of three ABC transporter units: substrate binding protein (SBP), transmembrane protein (TM) and a nucleotide binding protein (NBP). Additionally, two hypothetical, but conserved, proteins (HP1 and HP2) are encoded in proximity to *Rf*CODH (CooS). We aim to learn more about the native chemistry of clade B CODH by understanding the connection with their neighbouring genes.

We used the mirror tree method by Pazos and Valencia^[35] to understand to what degree clade B CODH and ABC transporter units are co-evolved. For this, we generated phylogenetic trees for all ABC transporter subunits and clade B CODH. We extracted the distance matrices and vectorised these matrices to calculate Pearson correlation between clade B CODH and their associated ABC transporter units (Figure S12 and S13), to estimate their co-evolution.

To control for phylogenetic background signal, we compared CODHs with a phylogenetic marker such as the 16S ribosomal subunit (16S rRNA) and RNA polymerase beta-subunit (*rpoB*) from the organisms encoding clade B CODH. Specific co-evolution should exceed background phylogenetic similarity between CODH and those markers.

The highest correlation was determined between the ABC transporter subunits (0.81 SBP/NBD, 0.87 TM/NBD, 0.91 SBP/TM), all above the threshold of 0.8 for true protein-protein interaction as established by Goh *et al.*^[36] and Pazos *et al.*^[35] (Figure S12). Moderate, though below-threshold, correlations are observed between CODH and ABC transporter subunits (0.49 CODH/SBP, 0.31 CODH/TM, 0.24 CODH/NBP). These correlations are stronger than those observed between CODH and phylogenetic background markers such as 16S rRNA or *rpoB* (0.18 CODH/16s RNA, 0.14 CODH/*rpoB*).

The evolution of clade B CODH correlates best with the evolution of SBP subunits from their respective operons. We propose that this partly results from both proteins interacting with a shared substrate molecule. The below-threshold correlation ($0.49 < 0.80$) is explained by their differing cellular localizations: CODHs are likely cytoplasmic, whereas SBPs are located outside the cell (Figure 7B). Consequently, we expected no physical protein-protein interaction. We interpret the observed trend as an indication that both ABC transporter unit and CODH interact with the same molecule, but not with each other. Therefore, either the product or the substrate of CODH is transported by the ABC transporter unit and since ABC transporters with SBPs typically function as importers^[37,38], we infer that the substrate of CODH is being imported via this transporter system. Given its low activity toward CO and the fact that CO can diffuse freely across membranes, we conclude that CO is unlikely to be the primary substrate of clade B CODHs.

The substrate transported by these ABC transporter systems remains unclear. While nickel transport was initially proposed^[13], clade B CODH associated SBPs (around 300 amino acids) are significantly smaller than known ABC transporter nickel-binding proteins (around 500 amino acids)^[39]. Additionally, clade B CODH associated SBPs show great phylogenetic similarity to SBPs that bind small anions such as bicarbonate and nitrate^[37]. They most closely resemble a zinc-binding protein from *Staphylococcus aureus* (PDB ID: 3UN6) of unknown function^[40], which enabled good structural modelling, but this did not help substrate prediction. Examination of the AlphaFold 3 model of the *Rf*CODH-associated SBP suggests that it binds a zinc or another metal ion, coordinated by two cysteines and one or two histidines, resembling a zinc-finger motif (Figure S14). The metal binding motif similarity between SBP and bacterial β -carbonic anhydrases was also noted; however, structural alignments of known β -carbonic anhydrase structures (PDB ID: 1EKJ) to the AlphaFold model yielded poor alignments. In literature, we found close relatives to *S. aureus* SBP homologous to clade B CODH associated SBPs, that are known to bind bicarbonate and nitrate (phylogenetic cluster F-IV as described by Scheepers *et al.*)^[37] We tested *Rf*CODH towards the reduction of both of these ions and could not see any turnover (Figure 1C). Interestingly, our mirror tree analysis revealed that many clade B CODHs are encoded adjacent to multiple ABC transporter units (Figure S13). These cases were excluded from the primary correlation analysis due to their added genomic complexity but might help triangulate the true substrate of clade B CODHs in future studies.

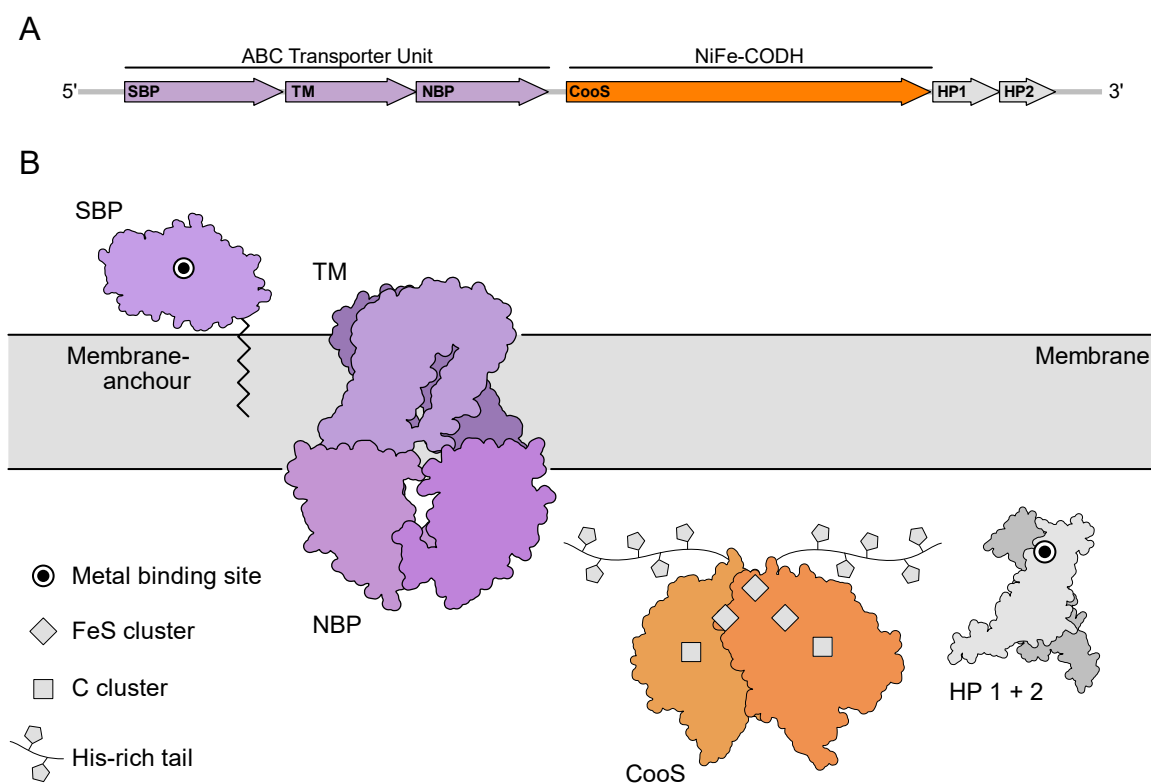


Figure 7. Genetic and structural organization of *Rf*CODH. (A) Operon structure of *Rf*CODH encompassing the canonical clade B architecture: the CODH gene (orange, CooS) flanked by ABC transporter units (purple, substrate binding protein (SBP), nucleotide binding protein (NBP) and trans membrane proteins (TM)) and two conserved hypothetical proteins (grey, HP1 and HP2). (B) Schematic view of operon-encoded proteins. The SBP is connected to the membrane via a lipid anchor. TM and NBP build a complex in the membrane, with NBP on the cytosolic side. Colour schemes same as in (A).

Conclusion

Our structural and functional characterisation of *Rf*CODH supports that clade B CODHs represent a distinct evolutionary lineage that has diverged from canonical CO-oxidising enzymes. While maintaining the core CODH scaffold, *Rf*CODH exhibits systematic modifications to substrate access pathways, including occluded gas channels and an alternated proton transfer pathway, that fundamentally change its catalytic profile. The reduction in CO oxidation activity, combined with the absence of CO₂ reduction capability, demonstrates that despite its name and structural homology, *Rf*CODH does not function as a traditional carbon monoxide dehydrogenase.

The co-evolution of clade B CODHs with ABC transporter substrate-binding proteins, points toward an alternative physiological role where substrate delivery is mediated through active transport rather than passive gas diffusion.

Our EPR analysis reveals unprecedented oxidised C-cluster states not previously characterised in CODHs, suggesting that the active site electronic structure has been modified. The additional presence of a confined nickel species and the unique rhombic signals upon thionine oxidation indicate either incomplete cluster assembly in the heterologous host or, more intriguingly, an intrinsically different C-cluster architecture optimised for alternative chemistry. The weak but detectable CN⁻ binding observed by FTIR spectroscopy confirms the presence of an open coordination site at nickel, preserving the fundamental capacity for substrate coordination while potentially altering the energetics of catalysis.

While our heterologous expression system has enabled the first structural and spectroscopic characterisation of a clade B CODH, the sub-stoichiometric nickel incorporation (0.2 – 0.5 Ni/monomer) indicates incomplete maturation, which undoubtedly contributes to the low observed activities.

The novel anaerobic grid preparation approach described here, enables preparation of grids suitable for cryo-EM studies of oxygen sensitive metalloproteins requiring minimum modifications to existing anaerobic workflows.

Future studies employing native expression systems or reconstitution with appropriate, but yet unknown, maturation factors will be essential to fully realise the catalytic potential of these enzymes. Nevertheless, the present work establishes that clade B CODHs represent functionally distinct enzymes whose physiological roles likely extend beyond CO metabolism.

Supporting Information

The sequences used in the bioinformatic analysis are supplied in an additional file. Table S1 to S2 and Figures S1 to S14 can also be found in the electronic supporting information. The authors have cited additional references within the Supporting Information.^[41–74] The atomic coordinates of RfCODH and cryo-EM volume map have been deposited in the Protein Data Bank (PDB) and Electron Microscopy Data Bank (EMDB) under the accession codes 28LW and 56604, respectively.

Acknowledgements

The Novo Nordisk Foundation (Grant ref. NNF21OC0066716, HL, MB, PH, NNF23OC0085682, MS), the Knut and Alice Wallenberg foundation (2023.0201, MH), the Swedish Research Council (2017-04018, MH), the Swedish Research Council (VR, grant no 2023-04593, MS) and FORMAS (2024-00586_3, MS) are gratefully acknowledged for funding. The cryo-EM data was collected at the Cryo-EM Swedish Infrastructure Unit, funded by SciLifeLab, the Knut and Alice Wallenberg Foundation, the Family Erling Persson Foundation, the Kempe Foundations, Stockholm University and Umeå University. We thank Dr. Xiaodong Zou's laboratory at Stockholm University for providing access to the PELCO easiGlow glow discharge system and ethane. We thank Dr. Leona Svecova for performing the TXRF experiment and analysis.

References

- [1] M. Can, F. A. Armstrong, S. W. Ragsdale, "Structure, function, and mechanism of the nickel metalloenzymes, CO dehydrogenase, and acetyl-CoA synthase" *Chem Rev* **2014**, *114*, 4149–74.
- [2] H. Dobbek, V. Svetlitchnyi, L. Gremer, R. Huber, O. Meyer, "Crystal Structure of a Carbon Monoxide Dehydrogenase Reveals a [Ni-4Fe-5S] Cluster" *Science* **2001**, *293*, DOI 10.1126/science.1061500.
- [3] E. J. Kim, J. Feng, M. R. Bramlett, P. A. Lindahl, "Evidence for a Proton Transfer Network and a Required Persulfide-Bond-Forming Cysteine Residue in Ni-Containing Carbon Monoxide Dehydrogenases" *Biochemistry* **2004**, *43*, 5728–5734.
- [4] A. Biester, S. Dementin, C. L. Drennan, "Visualizing the gas channel of a monofunctional carbon monoxide dehydrogenase" *Journal of Inorganic Biochemistry* **2022**, DOI 10.1016/j.jinorgbio.2022.111774.
- [5] W. Gong, B. Hao, Z. Wei, D. J. Ferguson, T. Tallant, J. A. Krzycki, M. K. Chan, "Structure of the $\alpha_2\beta_2$ Ni-dependent CO dehydrogenase component of the Methanosarcina barkeri acetyl-CoA decarbonylase/synthase complex" *Proceedings of the National Academy of Sciences* **2008**, *105*, 9558–9563.
- [6] M. Benvenuti, M. Meneghello, C. Guendon, A. Jacq-Bailly, J. H. Jeoung, H. Dobbek, C. Leger, V. Fourmond, S. Dementin, "The two CO-dehydrogenases of Thermococcus sp. AM4" *Biochim Biophys Acta Bioenerg* **2020**, *1861*, 148188.
- [7] J. Hadj-Said, M. E. Pandelia, C. Leger, V. Fourmond, S. Dementin, "The Carbon Monoxide Dehydrogenase from Desulfovibrio vulgaris" *Biochim Biophys Acta* **2015**, *1847*, 1574–83.
- [8] V. Svetlitchnyi, C. Peschel, G. Acker, O. Meyer, "Two membrane-associated NiFeS-carbon monoxide dehydrogenases from the anaerobic carbon-monoxide-utilizing eubacterium Carboxydothemus hydrogenoformans" *J Bacteriol* **2001**, *183*, 5134–44.
- [9] S. Jain, A. Katsyv, M. Basen, V. Muller, "The monofunctional CO dehydrogenase CooS is essential for growth of Thermoanaerobacter kivui on carbon monoxide" *Extremophiles* **2021**, *26*, 4.
- [10] J. H. Jeoung, J. Fessler, L. Domnik, F. Klemke, M. Sinnreich, C. Teutloff, H. Dobbek, "A Morphing [4Fe-3S-nO]-Cluster within a Carbon Monoxide Dehydrogenase Scaffold" *Angew Chem Int Ed Engl* **2022**, *61*, e202117000.
- [11] M. Inoue, K. Omae, I. Nakamoto, R. Kamikawa, T. Yoshida, Y. Sako, "Biome-specific distribution of Ni-containing carbon monoxide dehydrogenases" *Extremophiles* **2022**, *26*, 9.
- [12] M. Böhm, H. Land, "Investigating the native functions of [NiFe]-carbon monoxide dehydrogenases through genomic context analysis" *eLife* **2026**, *15*, DOI 10.7554/eLife.108780.1.
- [13] M. Inoue, I. Nakamoto, K. Omae, T. Oguro, H. Ogata, T. Yoshida, Y. Sako, "Structural and Phylogenetic Diversity of Anaerobic Carbon-Monoxide Dehydrogenases" *Front Microbiol* **2018**, *9*, 3353.
- [14] D. A. Antonopoulos, K. E. Nelson, M. Morrison, B. A. White, "Strain-specific genomic regions of Ruminococcus flavefaciens FD-1 as revealed by combinatorial random-phase genome sequencing and suppressive subtractive hybridization" *Environmental Microbiology* **2004**, *6*, 335–346.
- [15] W. A. Ayers, "Nutrition and physiology of ruminococcus flavefaciens" *Journal of Bacteriology* **1958**, *76*, 504–509.
- [16] J. Seravalli, M. Kumar, W.-P. Lu, S. W. Ragsdale, "Mechanism of CO Oxidation by Carbon Monoxide Dehydrogenase from Clostridium thermoaceticum and Its Inhibition by Anions" *Biochemistry* **1995**, *34*, 7879–7888.
- [17] "Reconsidering the Enzyme Kinetics of [FeFe]-Hydrogenases: Improved Turnover Rates and New Insights into pH and Potential Dependence with Eu(II)-Based Solution Assays" *Analytical Chemistry* **2025**.
- [18] T. Inoue, T. Yoshida, K. Wada, T. Daifuku, K. Fukuyama, Y. Sako, "A simple, large-scale overexpression method of deriving carbon monoxide dehydrogenase II from thermophilic bacterium Carboxydothemus hydrogenoformans" *Biosci Biotechnol Biochem* **2011**, *75*, 1392–4.
- [19] S. M. Kim, J. Lee, S. H. Kang, Y. Heo, H.-J. Yoon, J.-S. Hahn, H. H. Lee, Y. H. Kim, "O₂-tolerant CO dehydrogenase via tunnel redesign for the removal of CO from industrial flue gas" *Nature Catalysis* **2022**, *5*, 807–817.
- [20] M. L. Overeijnder, W. R. Hagen, P.-L. Hagedoorn, "A thermostable hybrid cluster protein from Pyrococcus furiosus: effects of the loss of a three helix bundle subdomain" *J Biol Inorg Chem* **2009**, *14*, 703–710.

- [21] T. Inoue, K. Takao, T. Yoshida, K. Wada, T. Daifuku, Y. Yoneda, K. Fukuyama, Y. Sako, "Cysteine 295 indirectly affects Ni coordination of carbon monoxide dehydrogenase-II C-cluster" *Biochemical and Biophysical Research Communications* **2013**, *441*, 13–17.
- [22] Y. Basak, C. Lorent, J.-H. Jeoung, I. Zebger, H. Dobbek, "Metalloradical-driven enzymatic CO₂ reduction by a dynamic Ni–Fe cluster" *Nat Catal* **2025**, 1–10.
- [23] Z. Hu, N. J. Spangler, M. E. Anderson, J. Xia, P. W. Ludden, P. A. Lindahl, E. Münck, "Nature of the C-Cluster in Ni-Containing Carbon Monoxide Dehydrogenases" *J. Am. Chem. Soc.* **1996**, *118*, 830–845.
- [24] P. A. Lindahl, E. Münck, S. W. Ragsdale, "CO dehydrogenase from *Clostridium thermoaceticum*. EPR and electrochemical studies in CO₂ and argon atmospheres." *Journal of Biological Chemistry* **1990**, *265*, 3873–3879.
- [25] M. S. M. Jetten, A. J. Pierik, W. R. Hagen, "EPR characterization of a high-spin system in carbon monoxide dehydrogenase from *Methanotrix soehngeni*" *European Journal of Biochemistry* **1991**, *202*, 1291–1297.
- [26] J. A. Krzycki, L. E. Mortenson, R. C. Prince, "Paramagnetic centers of carbon monoxide dehydrogenase from aceticlastic *Methanosarcina barker*" *Journal of Biological Chemistry* **1989**, *264*, 7217–7221.
- [27] W. P. Lu, P. E. Jablonski, M. Rasche, J. G. Ferry, S. W. Ragsdale, "Characterization of the metal centers of the Ni/Fe-S component of the carbon-monoxide dehydrogenase enzyme complex from *Methanosarcina thermophila*" *Journal of Biological Chemistry* **1994**, *269*, 9736–9742.
- [28] A. Ciaccafava, D. Tombolelli, L. Domnik, J. Fessler, J. H. Jeoung, H. Dobbek, M. A. Mroginski, I. Zebger, P. Hildebrandt, "When the inhibitor tells more than the substrate: the cyanide-bound state of a carbon monoxide dehydrogenase" *Chem Sci* **2016**, *7*, 3162–3171.
- [29] M. V. Cherrier, X. Vernède, D. Fenel, L. Martin, B. Arragain, E. Neumann, J. C. Fontecilla-Camps, G. Schoehn, Y. Nicolet, "Oxygen-Sensitive Metalloprotein Structure Determination by Cryo-Electron Microscopy" *Biomolecules* **2022**, *12*, 441.
- [30] R. A. Warmack, B. B. Wenke, T. Spatzal, D. C. Rees, "Anaerobic cryoEM protocols for air-sensitive nitrogenase proteins" *Nat Protoc* **2024**, *19*, 2026–2051.
- [31] I. Hands-Portman, S. E. Bakker, "Customising the plunge-freezing workflow for challenging conditions" *Faraday Discuss.* **2022**, *240*, 44–54.
- [32] B. D. Cook, S. M. Narehood, K. L. McGuire, Y. Li, F. Akif Tezcan, M. A. Herzik, "Preparation of oxygen-sensitive proteins for high-resolution cryoEM structure determination using blot-free vitrification" *Nat Commun* **2025**, *16*, 3528.
- [33] R. Sanchez-Garcia, J. Gomez-Blanco, A. Cuervo, J. M. Carazo, C. O. S. Sorzano, J. Vargas, "DeepEMhancer: a deep learning solution for cryo-EM volume post-processing" *Commun Biol* **2021**, *4*, 874.
- [34] E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor, J. Damborsky, "CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures" *PLOS Computational Biology* **2012**, *8*, e1002708.
- [35] F. Pazos, A. Valencia, "Similarity of phylogenetic trees as indicator of protein–protein interaction" *Protein Engineering* **2001**, *14*, 609–614.
- [36] C.-S. Goh, A. A. Bogan, M. Joachimiak, D. Walther, F. E. Cohen, "Co-evolution of proteins with their interaction partners¹" *Journal of Molecular Biology* **2000**, *299*, 283–293.
- [37] G. H. Scheepers, J. A. Lycklama a Nijeholt, B. Poolman, "An updated structural classification of substrate-binding proteins" *FEBS Letters* **2016**, *590*, 4393–4401.
- [38] D. C. Rees, E. Johnson, O. Lewinson, "ABC transporters: the power to change" *Nat Rev Mol Cell Biol* **2009**, *10*, 218–227.
- [39] R. Tam, M. H. Saier, "Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria" *Microbiol Rev* **1993**, *57*, 320–346.
- [40] G. Minasov, Z. Wawrzak, A. Halavaty, L. Shuvalova, I. Dubrovskaya, J. Winsor, O. Kiryukhina, F. Bagnoli, F. Falugi, M. Bottomley, G. Grandi, W. F. Anderson, Center for Structural Genomics of Infectious Diseases (CSGID) **2011**, DOI 10.2210/pdb3un6/pdb.
- [41] K. Katoh, D. M. Standley, "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability" *Mol Biol Evol* **2013**, *30*, 772–780.
- [42] K. Katoh, K. Kuma, H. Toh, T. Miyata, "MAFFT version 5: improvement in accuracy of multiple sequence alignment" *Nucleic Acids Res* **2005**, *33*, 511–518.
- [43] S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, "trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses" *Bioinformatics* **2009**, *25*, 1972–1973.
- [44] B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R. Lanfear, "IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era" *Molecular Biology and Evolution* **2020**, *37*, 1530–1534.
- [45] T. G. Schmidt, A. Skerra, "The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins" *Nat Protoc* **2007**, *2*, 1528–1535.
- [46] M. K. Akhtar, P. R. Jones, "Deletion of *iscR* stimulates recombinant clostridial Fe–Fe hydrogenase activity and H₂-accumulation in *Escherichia coli* BL21(DE3)" *Appl Microbiol Biotechnol* **2008**, *78*, 853–862.
- [47] M. V. Pavliuk, M. Böhm, J. Wilhelmsen, S. Hardt, H. Land, H. Tian, "Photobiocatalytic CO₂ reduction into CO by organic nanorod-carbon monoxide dehydrogenase assemblies: surfactant matters" *Chemical Science* **2024**, *15*, 16789–16795.
- [48] M. Senger, T. Kernmayr, M. Lorenzi, H. J. Redman, G. Berggren, "Hydride state accumulation in native [FeFe]-hydrogenase with the physiological reductant H₂ supports its catalytic relevance" *Chemical Communications* **2022**, *58*, 7184–7187.
- [49] S. Stoll, A. Schweiger, "EasySpin, a comprehensive software package for spectral simulation and analysis in EPR" *Journal of Magnetic Resonance* **2006**, *178*, 42–55.
- [50] E. Sayers in *Entrez Programming Utilities Help [Internet]*, National Center For Biotechnology Information (US), **2022**.

- [51] E. W. Sayers, E. E. Bolton, J. R. Brister, K. Canese, J. Chan, D. C. Comeau, R. Connor, K. Funk, C. Kelly, S. Kim, T. Madej, A. Marchler-Bauer, C. Lanczycki, S. Lathrop, Z. Lu, F. Thibaud-Nissen, T. Murphy, L. Phan, Y. Skripchenko, T. Tse, J. Wang, R. Williams, B. W. Trawick, K. D. Pruitt, S. T. Sherry, "Database resources of the national center for biotechnology information" *Nucleic Acids Research* **2022**, *50*, D20–D26.
- [52] M. Böhm **2025**, DOI <https://doi.org/10.5281/zenodo.16736722>.
- [53] M. Böhm **2025**, DOI <https://doi.org/10.5281/zenodo.16736767>.
- [54] S. Kalyanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermin, "ModelFinder: fast model selection for accurate phylogenetic estimates" *Nat Methods* **2017**, *14*, 587–589.
- [55] E. Paradis, K. Schliep, "ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R" *Bioinformatics* **2019**, *35*, 526–528.
- [56] R Core Team **2023**.
- [57] H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Golemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, "Welcome to the Tidyverse" *JOSS* **2019**, *4*, 1686.
- [58] W. Shen, S. Le, Y. Li, F. Hu, "SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation" *PLOS ONE* **2016**, *11*, e0163962.
- [59] L. J. Revell, "phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things)" *PeerJ* **2024**, *12*, e16505.
- [60] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstern, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Židek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J. M. Jumper, "Accurate structure prediction of biomolecular interactions with AlphaFold 3" *Nature* **2024**, *630*, 493–500.
- [61] A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, "Jalview Version 2—a multiple sequence alignment editor and analysis workbench" *Bioinformatics* **2009**, *25*, 1189–1191.
- [62] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, "Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*" *Acta Crystallogr D Struct Biol* **2019**, *75*, 861–877.
- [63] T. I. Croll, "ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps" *Acta Crystallogr D Struct Biol* **2018**, *74*, 519–530.
- [64] P. Emsley, K. Cowtan, "Coot: model-building tools for molecular graphics" *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2126–2132.
- [65] E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, T. E. Ferrin, "UCSF CHIMERAX: Structure visualization for researchers, educators, and developers" *Protein Science* **2021**, *30*, 70–82.
- [66] T. Bepler, A. Morin, M. Rapp, J. Brasch, L. Shapiro, A. J. Noble, B. Berger, "Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs" *Nat Methods* **2019**, *16*, 1153–1160.
- [67] A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination" *Nat Methods* **2017**, *14*, 290–296.
- [68] J. Xu, V. Srinivas, R. Kumar, L. Pacoste, Y. Guo, T. Yang, C. C. Sun, M. Högbom, X. Zou, H. Xu, "Unveiling the Structure of Anhydrous Sodium Valproate with 3D Electron Diffraction and a Facile Sample Preparation Workflow" *ACS Cent. Sci.* **2025**, *11*, 960–966.
- [69] M. Böhm **2026**, DOI <https://doi.org/10.5281/zenodo.18506471>.
- [70] H. Wickham **2025**.
- [71] D. J. Winter, "rentrez: an R package for the NCBI eUtils API" *The R Journal* **2017**, *9*, 520–526.
- [72] H. Pagès, P. Aboyoun **2017**, DOI [10.18129/B9.BIOC.BIOSTRINGS](https://doi.org/10.18129/B9.BIOC.BIOSTRINGS).
- [73] O. Wagih, "ggseqlogo: a versatile R package for drawing sequence logos" *Bioinformatics* **2017**, *33*, 3645–3647.
- [74] U. Bodenhofer, E. Bonatesta, C. Horejš-Kainrath, S. Hochreiter, "msa: an R package for multiple sequence alignment" *Bioinformatics* **2015**, *31*, 3997–3999.