



**Michigan  
Technological  
University**

Michigan Technological University  
**Digital Commons @ Michigan Tech**

---

Dissertations, Master's Theses and Master's Reports

---

2022

## INVESTIGATION OF A CARBON MONOXIDE DEHYDROGENASE FROM AN UNCULTURED ARCHAEON

Luke Moore

*Michigan Technological University, lukem@mtu.edu*

Copyright 2022 Luke Moore


---

### Recommended Citation

Moore, Luke, "INVESTIGATION OF A CARBON MONOXIDE DEHYDROGENASE FROM AN UNCULTURED ARCHAEON", Open Access Master's Thesis, Michigan Technological University, 2022.

<https://doi.org/10.37099/mtu.dc.etr/1420>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etr>

 Part of the [Biochemistry Commons](#), [Biotechnology Commons](#), and the [Molecular Biology Commons](#)

INVESTIGATION OF A CARBON MONOXIDE DEHYDROGENASE FROM AN  
UNCULTURED ARCHAEON

By

Luke Moore

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Biological Sciences

MICHIGAN TECHNOLOGICAL UNIVERSITY

2022

© 2022 Luke Moore

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Biological Sciences.

Department of Biological Sciences

Thesis Advisor: *Stephen Techtmann*

Committee Member: *Rupali Datta*

Committee Member: *Amy Marcarelli*

Department Chair: *Chandrashekar Joshi*

# Table of Contents

Acknowledgements.....	v
Abstract.....	vi
1 Introduction.....	1
1.1 Carbon Monoxide’s Importance in the Environment and as a Human Resource.....	1
1.1.1 Human Emissions .....	1
1.1.2 Natural Sources of CO .....	2
1.1.3 CO and CODH as a Resource.....	2
1.2 Gene Structure and Metabolic Function.....	4
1.3 Phylogeny of CODHs.....	7
1.4 Biochemical Characterization of CODH in the Lab .....	8
1.5 Aims .....	10
2 Methods.....	11
2.1 Gene Constructs and Vector Creation .....	11
2.2 Transformations and Protein Purification .....	11
2.2.1 Transformations .....	11
2.2.2 Expression and Column Purification .....	12
2.3 Gel .....	13
2.4 Protein Characterization .....	14
2.4.1 Dialysis and BCA Quantification .....	14
2.4.2 Melt Curves.....	14
2.4.3 Enzyme Activity and Whole Cell Assays.....	15
2.4.4 Gas Chromatography .....	16
2.4.5 ICP-OES .....	17
3 Results.....	18
3.1 Genomic Context of the CODH gene clusters in Hydrothermarchaeota JdFR-17 18	18
3.2 Expression and Purification of the Hydrothermarchaeota CooS Protein .....	18
3.3 Melt Curves .....	20
3.4 CO and CO <sub>2</sub> Activity Assays .....	23
3.4.1 Carbon Monoxide Detection.....	23
3.4.1.1 CO Oxidation in the Reverse Reaction.....	23
3.4.1.2 CO <sub>2</sub> Reduction in the Forward Reaction.....	25
3.4.2 CO <sub>2</sub> Detection .....	26
3.4.2.1 CO <sub>2</sub> in the Reverse Reaction.....	26
3.4.2.2 CO <sub>2</sub> in the Forward Reaction .....	27
3.5 ICP-OES.....	28
4 Discussion .....	30

4.1	Expression .....	30
4.2	Activity.....	31
4.3	Future Directions.....	31
	4.3.1 Genetic Considerations .....	31
	4.3.2 Chemical Inhibition .....	32
4.4	Conclusion.....	33
5	Reference List .....	35
A	Appendix.....	39
	A.1 Carbon Monoxide Constructs –.....	39
	A.2 Autoinduction Media –.....	39
	A.3 BCA Assay Raw Data – .....	39
	A.4 BCA Assay Protocol – .....	39
	A.5 Melt Curve Assay PCR Settings – .....	39
	A.6 Melt Curve Assay Setup –.....	39
	A.7 GC Assay Standard Curve –.....	39
	A.8 Forward Reaction Example –.....	39
	A.9 Reverse Reaction Example –.....	39
	A.10 Raw Data for the GC Assays –.....	39
B	Copyright documentation.....	40
	B.1 Figure 1.....	40
	B.2 Figure 2.....	40
	B.3 Figure 3.....	40
	B.4 Figure 4.....	40

## **Acknowledgements**

I would like to thank Stephen Techtmann as my advisor and mentor for this project. The support provided has been monumental in my growth as a student.

Yogita Warkhade has been an extremely important and reliable lab partner, and without her this would not be possible.

Laura Schaerer, Lindsay Putman, Isaac Bigcraft, and all the other members of the Techtmann Lab have been supportive and knowledgeable colleagues in my time here. Thank you all.

I would also like to thank my committee members for their feedback and support.

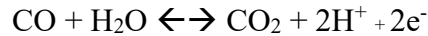
Finally, the MTU Biological Sciences department has been a great resource for this entire project.

## Abstract

The Nickel based Carbon Monoxide Dehydrogenase (CODH) is an anaerobic metalloenzyme responsible for the reversible conversion of CO and water into CO<sub>2</sub> and 2 protons and 2 electrons. This enzyme has importance in the environment as one of Earth's first carbon fixation pathways, and for human uses as a potential source of biofuels and other commodity chemicals. CODH enzymes are present in a wide array of taxa, many of which are uncultured. In this study we express and purify the catalytic subunit (CooS) of the anaerobic CODH from an uncultured Hydrothermarchaeota JdFR-17 co-expressed with the nickel insertion accessory protein (CooC) from *Archaeoglobus fulgidis* to generate a CODH complex in *E. coli*. The protein was then characterized via activity assays and thermal stability assays to test the hypothesis that this CooS was more active for CO oxidation compared to CO<sub>2</sub> reduction and functioned at elevated temperatures. A soluble CooS protein was purified that contained both nickel and iron in the active site. This provides evidence that CooC and CooS proteins from two different species can cooperate, and that it is possible to express Archaeal CODH complexes in *E. coli*. However, enzyme assays yielded inconclusive results. This led us to infer that the CODH complex was expressed, yet inactive.

# 1 Introduction

Anaerobic carbon monoxide dehydrogenases (CODHs) are protein complexes involved in carbon monoxide metabolism. The core of the CODH is the catalytic subunit called the CooS. The CooS is the CODH subunit that carries out the reversible conversion of CO to carbon dioxide. The reaction is as follows:



CODHs are very important in carbon cycling and have a role in ecosystem stability. Feng and Lindhal state that on a global scale, microbial activity in soils consumes ~10% of annual CO emissions [1]. On top of this, CODHs have been pivotal in the evolution of life on Earth. The Wood-Ljungdahl pathway for Acetyl-CoA synthesis has been coupled with the CODH in microbial metabolisms for 3.5 billion years. Previous studies have suggested that the CODH and the Wood-Ljungdahl pathway were present in the last universal common ancestor (LUCA) as one of the earliest forms of carbon fixation [2]. Understanding how life has formed on Earth could yield interesting discoveries in the field of evolution. It may even be possible to phylogenetically reconstruct an early carbon monoxide metabolizing organism, which would yield scientific novelty. Furthermore, knowing the metabolism of early life can also give clues on what to look for when searching for life on other planets[2, 3].

With such a substantial role in the environment and life's history, there is a need for a deeper understanding of the diversity and activity of CODHs. The main challenges to understanding the anaerobic CODH is that they function in a unique environment. While CODHs have been shown to present in many organisms. [4] Some of these CODHs may be vestigial structures which can lead to questions about their function and importance. Also, many organisms that encode CODHs use other pathways as their primary metabolism [5]. The function of many of these CODHs in the genome remain a mystery until lab work is done to determine enzyme activity and biochemically determine their function.

## 1.1 Carbon Monoxide's Importance in the Environment and as a Human Resource

### 1.1.1 Human Emissions

Carbon Monoxide (CO) is a gas best known for its toxicity for humans. It is a common pollutant as well as a trace gas in the atmosphere. However, many microorganisms have the ability to metabolize CO [6]. Anthropogenic emissions of carbon monoxide are produced by many industrial and commercial processes.

The main source of CO pollution is the combustion of fossil fuels. In cities, CO levels can rise to a higher level than OSHA deems safe. OSHA considers anything more than 50 ppm for over 8h exposure dangerous [7]. The CO levels in major cities can get up to 350 ppm. Atmospheric CO has harmful effects on the human body. Atmospheric CO reduces the amount of oxygen that can be carried in the circulatory system, which can cause complications in the heart and lungs. This is especially concerning for people with pre-existing heart conditions. Furthermore, long term CO exposure is damaging to the



human body even if there are no acute symptoms. CO cleanup becomes more important as the global population continues to rise, and more people live in urban areas where CO pollution is at its worst.

### 1.1.2 Natural Sources of CO

In an addition to the industrial sources of carbon monoxide in the environment, there are a few environmental sources of carbon monoxide. One natural source of CO is microbial metabolism [8–11]. Sulfate-reducers and methanogens are some of the microbial CO sinks, while glucose and formate metabolizers act as a source for CO [10]. CO can also be produced from hydrothermal vents and volcanic eruptions through abiotic processes [6]. CO is also produced by Eukaryotes such as plants and animals in their metabolisms [6].

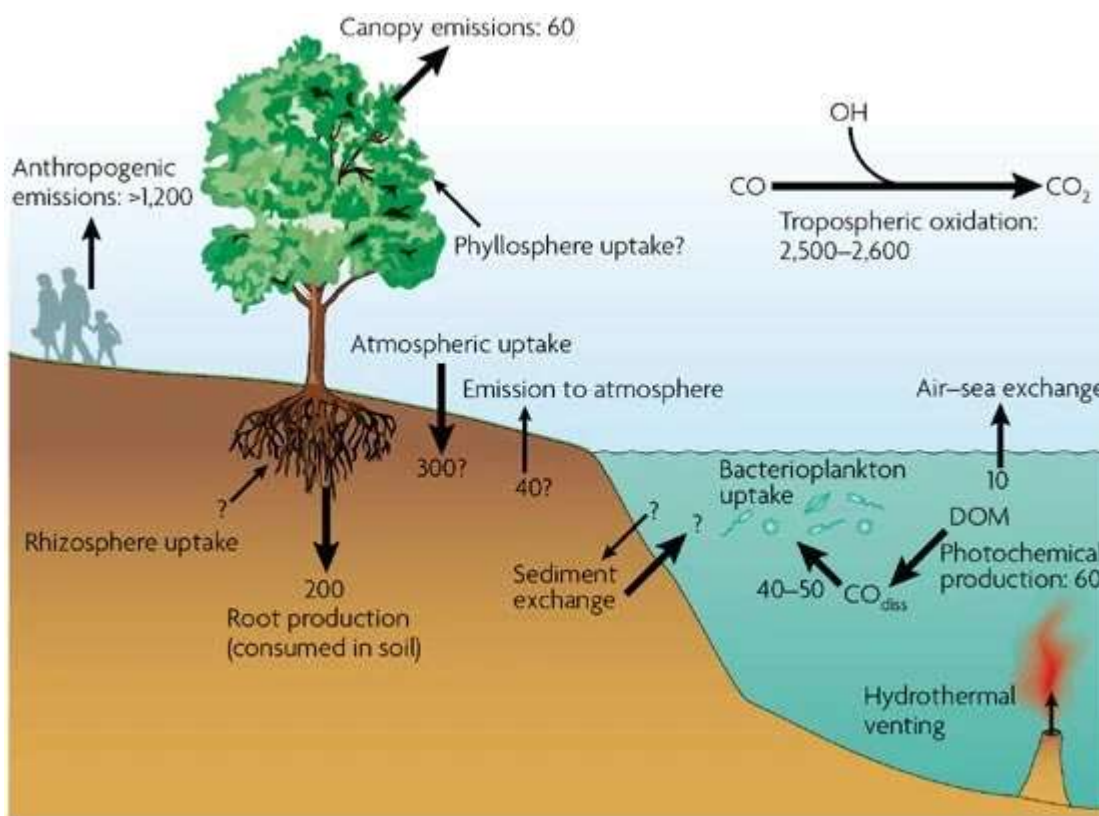


Figure 1: Shows the sources and sinks of carbon monoxide in the atmosphere [6].  
Copyright information in section B.1

### 1.1.3 CO and CODH as a Resource

Carbon monoxide is also considered a valuable product for production of value-added chemicals such as fuels [12]. Some promising strategies have been developed to convert  $CO_2$  into CO and then use electrochemical strategies to convert CO into chemicals such as fuels and polymers [13]. Several waste streams and industrial process result in carbon monoxide production. Synthesis gas (Syngas) is a gas composed of a

mixture of hydrogen and carbon monoxide with a minor amount of CO<sub>2</sub>. There has been growing interest in the production of synthesis gas as a means of producing valuable commodities from waste streams, as most waste streams can be converted into synthesis gas through gasification. These include materials that are dry, like wood chips, or non-degradable materials like plastics and rubber [14].

In synthesis gas, CO can make up more than 50% of the gas composition. Syngas can be used to produce several chemicals from NH<sub>4</sub> to fuels. The hydrogen from syngas can be used in the Haber-Bosch process for NH<sub>4</sub> production. The CO from the syngas can be used to produce fuels using the Fischer-Tropsch reaction to produce synthetic petroleum. These various processes show that carbon monoxide is a common industrial gas that is produced to be upgraded to various valuable products.

For the syngas to be used in specific industrial processes it needs to have the correct concentration of hydrogen. This very important in making the fuel volatile enough to be an energy source. The water gas shift reaction is often used to increase the hydrogen content of synthesis gas. The water gas shift reaction involves the reaction of CO with water to produce CO<sub>2</sub> and hydrogen. Currently, the water gas shift reaction is carried out using inorganic catalysts at very high temperatures. The use of biological systems for this process would be advantageous because the energy requirement to run the water gas shift reaction would be decreased. Organisms that could be used for the water gas shift reaction include hydrogenogenic carboxydrotrophs such as *Rhodospirillum rubrum* and *Carboxydotherrmus hydrogenoformans* [14].

While *Rhodospirillum rubrum* and *Carboxydotherrmus hydrogenoformans* have been tested on small scales to attempt the WGS reaction, they have not been able to be optimized for a whole industrial plant scale process. Overall, if the CODH containing organisms are to be used for industrial synthesis gas applications, they need to be optimized to grow under the conditions of the chemical processing plant. *R. rubrum* can metabolize CO well; however, it requires another carbon source to grow. This is a major drawback since it is very hard to try to keep conditions optimal for growth in an industrial setting. *C. hydrogenoformans* has some problems as well. It's optimal growing temperature is around 70°C. This leads to issues because CO has low solubility in growth medium at that temperature. *C. hydrogenoformans* also has been reported to have trouble growing in the industrial setting due to its need for minerals to grow. The syngas mixture itself can also inhibit bacteria attempting WGS. While *C. hydrogenoformans* can metabolize pure CO well, its metabolism is decreased when exposed to the other gasses in the mixture [14]. Changing the conditions to optimize growth is far too expensive and troublesome [14]. Some biotech companies have had success in engineering organisms for conversion of CO-containing waste streams into valuable products such as jet fuel [15]. There is a need for new organisms' CooS to be characterized so that an optimal candidate can be identified.

## 1.2 Gene Structure and Metabolic Function

Carbon monoxide dehydrogenases have been studied for decades [16]. These enzymes have been found in relatively common organisms like *R. rubrum*. [1] However, the majority of the CODH containing organisms are thermophiles found in hydrothermal vents and seafloor sediments. Organisms like *Moorella thermoautotrophica* and *Carboxydotherrmus hydrogenoformans* are typical CODH-containing organisms [17, 18]. Archaea and Bacteria that are extremophiles are especially hard to culture in the lab. This is due to their need for high temperature and often requirement for anaerobic conditions. This has led to a gap in our understanding of CODH complexes, as the CODH gene clusters identified in genomes far exceeds those that have characterized in the lab. To understand the unknown CODHs it is essential to build off what has already been researched.

There are two large classes of carbon monoxide dehydrogenases: aerobic or anaerobic. Aerobic carbon monoxide dehydrogenases are molybdenum containing proteins. The Mo-CODHs consist of operons encoding three subunits called the *coxSML* [19]. The anaerobic CODHs are Ni and Fe containing proteins that are found in diverse environments [6, 20]. There are two types of anaerobic carbon monoxide dehydrogenase: monofunctional and bifunctional. Bifunctional CODHs are used to assimilate carbon in the cell, whereas monofunctional CODHs are used for energy generation in the form of the electron released from the oxidation of CO.

Bifunctional CODHs are more widely studied as the products can be used downstream in the Wood-Ljungdahl pathway (Figure 2). Bifunctional CODHs run reversibly from CO to CO<sub>2</sub>, which is fed into the Wood-Ljungdahl pathway or used in methanogenesis. The Wood-Ljungdahl pathway is one of the most ancient forms of Acetyl CoA synthesis and is believed to be one of the first forms of carbon fixation on Earth. [5] With its part in such an important metabolic pathway, CODH could yield interesting applications if fully understood.

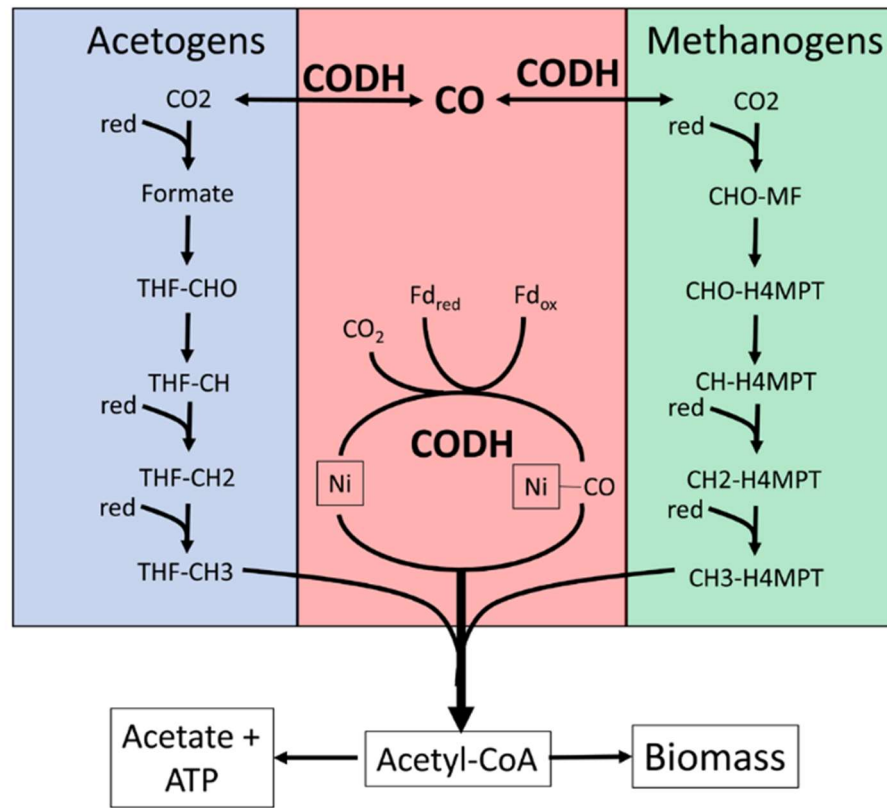


Figure 2: Shows the Wood-Ljungdahl pathway and the CODH's significance [5]. Copyright information is in section B.2.

Monofunctional CODHs on the other hand, are not as well understood. They are not connected to the Wood-Ljungdahl pathway and have not been explored as widely. However, this does not deem them unimportant. When monofunctional CODHs oxidize CO into CO<sub>2</sub>, they release electrons which can be used in other metabolic processes such as sulfate reduction, hydrogen production, iron reduction, as well as many others. It is often not obvious to discern the path of these electrons and their downstream function, and many of the monofunctional CODHs discovered have no known function [4].

The metabolic function of a CODH is hard to determine by just analyzing the genetic context of CODH itself. Many CODHs are bifunctional and take on a role in acetogenesis, methanogenesis while some of the monofunctional CODHs are in operon with hydrogenases (Figure 3). Other CODHs have no known function or are thought to be vestigial structures within the genome. Examining the genes a CODH can sometimes help to know what a CODH does in conjunction with the organism's other metabolic pathways. [20]

The gene that codes the anaerobic Ni-CODH catalytic subunit is called the *cooS* gene. This gene is surrounded by other accessory proteins that can reveal the function of the CODH. The most common genes in the vicinity of the *cooS* are the *cooC* genes.

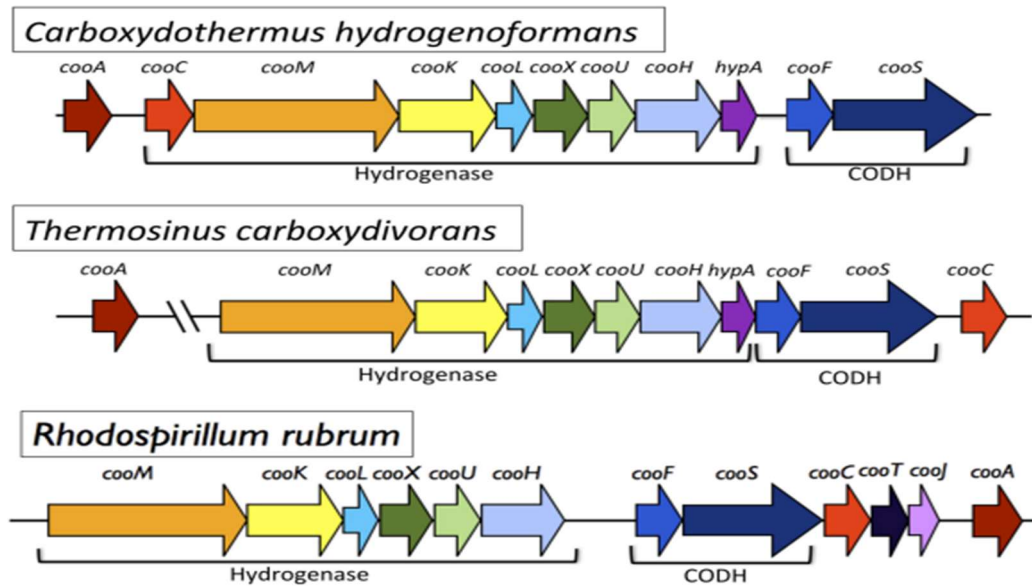


Figure 3: Shows the structure of CODHs from different organisms and the proteins surrounding them [4]. Copyright information is in section B.3.

The CooC protein is also very important in making the protein functional. The CooC protein, which is a homodimer, is a nickel chaperone protein that is involved in the active site maturation of the C cluster[21]. The CooC is involved in insertion of the Ni into the active site of the CooS, which is essential for CooS activity. *cooCs* have been found in *Carboxydothermus hydrogenoformans* and are essential to performing the water gas shift reaction. [14]

The other accessory genes such as *cooF*, *cooT*, and *cooJ* are also located near the CODH. [4] Most of the accessory genes code for metalloproteins. They are also theorized to be used for the activation of the enzyme and require nickel in solution to be activated. These have been studied and have been deemed to be less essential to the gas shift reaction than the CooC accessory protein, which is pivotal in active site maturation. [21]

One of the most well-known accessory proteins is the CooF protein. It acts as an electron shuttle, which feeds electrons from the CODH into other pathways. Ni-CODHs can use many electron acceptors. These include CO<sub>2</sub>, protons, sulfate, and ferric iron. [7]An example of one of these acceptors is the energy conserving hydrogenase metabolic pathway that turns protons into Hydrogen gas. This is one of the main interests of chemical engineers looking to fortify biofuels with Hydrogen. The CooF protein is important for monofunctional CODHs as they can determine the downstream processes catalyzed by the CooS and the fate of electrons released by CO oxidation.

While the structure of the protein has been determined: the reaction mechanism of this protein is still debated and can possibly fluctuate among different homologs of this protein. The reaction from *Rhodospirillum rubrum* is the most widely studied organism for CODH metabolism.

### 1.3 Phylogeny of CODHs

CODHs have been divided into one of seven known clades based on phylogenetic analysis[22]. These clades are labeled as clade A, B, C, D, E, F, and G.

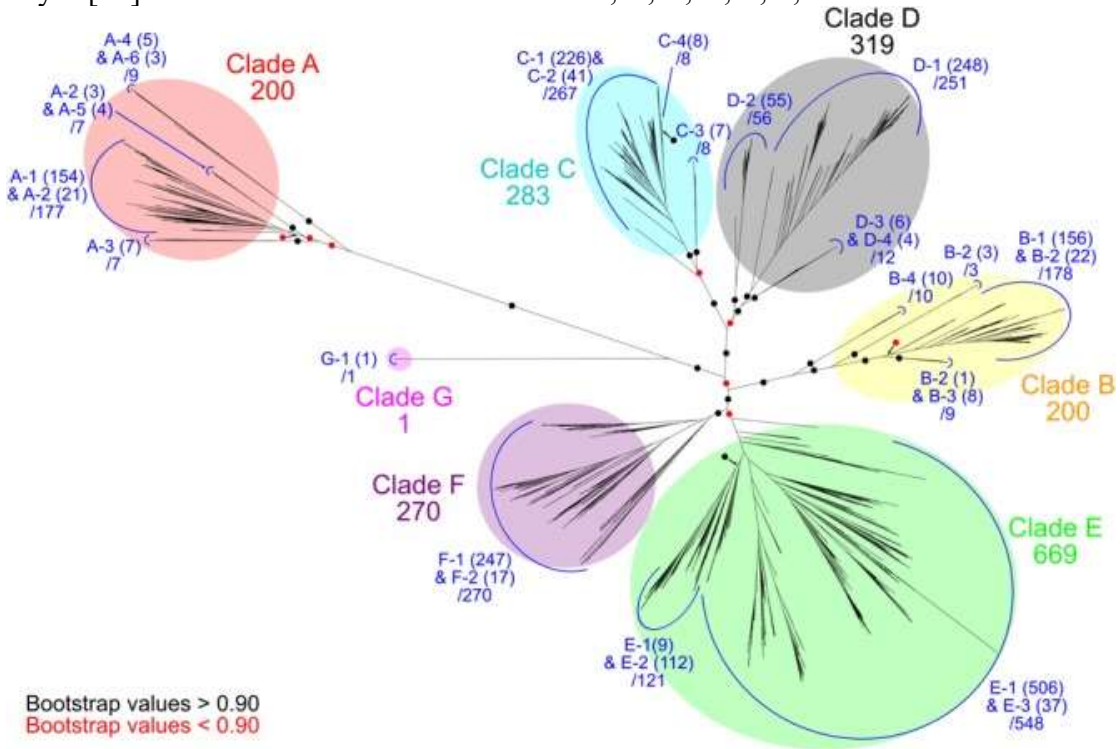


Figure 4: Shows the phylogeny of the different clades of carbon monoxide dehydrogenases. The loops that are within each clade are the subclasses. Numbers next to clades indicate proteins discovered within. [22]

Many of the CODH genes that are in clade A are connected to the Wood-Ljungdahl pathway for acetyl CoA synthesis. This makes it reasonable to believe that they are bifunctional CODHs. Clade F and clade E also have some genes that are connected to the Wood-Ljungdahl pathway. [20, 22] However, clade F also has subclasses that connect to the energy conserving hydrogenase gene. [20, 22] Clade G only contains one hypothetical protein from one organism: Deltaproteobacteria bacterium, a sulfur reducer.

Clades B, C, and D mostly do not have connections to the Wood-Ljungdahl pathway. Many of these could be monofunctional CODHs. They are differentiated by their amino acid residues. They vary in the number and location of cysteine and histidine amino acids on the protein. [20]

The organisms that contain these different clades of CODH are widespread and phylogenetically diverse as well. In a phylogenetic analysis of 1942 CODHs [22], 748 of the CODHs found belonged to the phylum Firmicutes. This phylum includes many aerobic and anaerobic bacteria. “37 hydrogenogenic CO-oxidizing isolates have been

reported in five phyla, 20 genera, and 32 species". [20] Of these, the focus has been on the bifunctional CODHs. This leaves a very large amount of discovery to be done on the front of microbiology of carbon monoxide dehydrogenases.

While the bifunctional CODHs are well characterized in genomic analyses, monofunctional CODHs are harder to understand. There can be multiple monofunctional CODHs in one organism. *Carboxydothemus hydrogenoformans* has 5 known CODHs, and 4 of those are monofunctional. *Rhodospirillum rubrum* has monofunctional CODH as well and is one of the model organisms for studying monofunctional CODHs [23]. These organisms are both gram negative and anaerobic. It is suggested that the organisms that most heavily use monofunctional CODHs have an affinity toward anaerobic metabolism [14]. There are many organisms with monofunctional CODHs, but some seem to have no known function currently.

Monofunctional CODHs are found in a wide array of organisms, but the most archetypal CO metabolizing organisms live in the oceanic crust [17, 24]. Many CODHs have been found in uncultured organisms including archaea [24–26] One such organism is the Hydrothermarchaeota. The Hydrothermarchaeota were one of the most prevalent taxa in some marine subsurface samples, accounting for over half of the 16 rRNA in a sample taken from the Juan de Fuca Ridge flank. [24]. Carr et al 2019 use metagenomics to identify the Hydrothermarchaeota through single cell genomics and metagenome assembled genomes (MAGs). The Hydrothermarchaeota MAGs recovered from International Ocean Drilling Project (IODP) bore hole observatories near the Juan de Fuca Ridge contained the genes for carboxydutrophy [24].

The previous studies have shown the genome of Hydrothermarchaeota encodes the possibility to perform carboxydutrophy. One of the goals of this study is to provide more support for the hypothesis that the Hydrothermarchaeota perform carboxydutrophy. While genomics can provide information to the potential function of a gene based on homology comparisons, biochemical characterization of a protein is the gold standard for determining function.

## 1.4 Biochemical Characterization of CODH in the Lab

Biochemical characterization of enzymes is important for determination of protein function. Enzyme assays involve the purification of the protein from an organism that utilizes CO metabolism. The protein would then be exposed to CO and the rate of CO oxidation to CO<sub>2</sub> is measured or the rate of CO<sub>2</sub> reduction to CO is measured. [27–29] These studies have shown that CooS proteins are the catalytic subunit of the CODH complex and can perform the reversible oxidation or reduction of CO or CO<sub>2</sub> respectively.

One way to study CODH specific activity is to grow the organism that already has a native CODH and purify protein from that. In a study purifying CODH from *R. rubrum*, it was found that the native CODH had a specific activity of 5600  $\mu\text{mol min}^{-1} \text{mg}^{-1}$  for CO oxidation. The CO<sub>2</sub> reduction assay yielded a rate of 30.1  $\mu\text{mol min}^{-1} \text{mg}^{-1}$ [27].

Aside from studies that have performed purification of CODH enzymes from the native organisms, other studies have used heterologous expression of proteins in a host



like *E. coli* to express and purify enzymes from diverse organisms. This has been done by first extracting DNA from an organism with a known CODH. This DNA is sequenced and then the *cooS* sequence is added to an expression vector for expression in *E. coli*. Previous work has shown that expression of the *cooC* is essential for production of a function CooS protein due to the need for the Ni in the active site of the CooS [28].

To aid in the purification of the proteins once expressed in *E. coli* affinity tags are often added to the proteins for affinity protein purification. These affinity tags are often tags such as a six-histidine tag which binds to a nickel affinity column or a FLAG Tag or GST tag. Another important thing to consider when heterologously expressing proteins is the codon usage of the organism from which the CooS is derived and how it compares to the codon usage of the expression host (*E. coli*). If the codon usage of the *cooS* is different from that of *E. coli*, it is possible that the protein will not be expressed. One way to address this is to synthesize the gene adapting the codon usage to match that of the expression host. This will choose codons that are known to be the most common in *E. coli* to provide the same sequence as the *cooS* from the original organism. [29]

Once transformed into *E. coli*, the protein is purified, and the rate of reaction determined using enzyme assays. One study characterized the CODH from *C. hydrogenoformans* grown and purified from *E. coli*. The study reported rates from of 8500-9600  $\mu\text{mol min}^{-1} \text{mg}^{-1}$  CO oxidation into  $\text{CO}_2$ . 16.9  $\mu\text{mol min}^{-1} \text{mg}^{-1}$   $\text{CO}_2$  reduction into CO in a CODH from *C. hydrogenoformans* grown and purified from *E. coli* [29].

CODHs from other organisms have been used to determine a CODH's rate as well. *D. vulgaris* has been used to determine the effect CooC has on CODH activity. In this study it was found that the CODH purified from *E. coli* had a CO oxidation rate of 1660  $\mu\text{mol min}^{-1} \text{mg}^{-1}$ . This number decreased to less than 5  $\mu\text{mol min}^{-1} \text{mg}^{-1}$  when the experiment was run without the CooC insertion. The regular assay had a  $\text{CO}_2$  reduction rate of 1.4  $\mu\text{mol min}^{-1} \text{mg}^{-1}$ . This decreased to 0 when the experiment was run without the CooC insertion [28].

Another method of detecting CODH activity is by using a whole cell assay. This is done by growing a strain of CO metabolizing bacteria or Archaea and exposing them to CO or  $\text{CO}_2$ . In one study [8], *M. smegmatis* was exposed to CO after being grown to post stationary phase and the CODH activity was monitored. This was done anaerobically in a nitrogen filled atmosphere [8]. The CO substrate concentration in this study was particularly low; about 200 ppm but it was shown that the  $V_{\text{max}}$  was about 3.13  $\text{nmol gdw}^{-1} \text{min}^{-1}$ . This is much lower than the enzyme only assays described above.

Finally, cell lysates have also been used to measure CODH activity. This is done by growing up a culture to late exponential phase and harvesting the cells anaerobically. They are then disrupted using a bead beater to break the cells open. The lysate is then assayed using a methyl viologen reduction spectrophotometric assay to measure the activity of the CODH [30]. This was done in *C. amalonaticus*. The activity of the CO reduction of this extract was around 49.3  $\mu\text{mol min}^{-1} \text{mg}^{-1}$ .



## 1.5 Aims

The vast majority of microorganisms from the environment are difficult to culture in the lab [31]. This greatly restricts our ability to understand the functions that these environmental organisms play in the environment. Metagenomics has enabled deep insights into the functional potential of environmental microbial communities. However much of the conclusions that are drawn from metagenomics involves comparison of sequence similarity of predicted genes from the metagenomes with biochemically characterized proteins from select isolates. While this approach can provide insights into the function of genes in uncultured organisms, it is essential to biochemically characterize proteins from uncultured organisms to gain insights into their functions beyond sequence similarity.

The aim of this study is to characterize a CooS from an uncultured organism. Due to the extensive amount of labor required to characterize a protein from an uncultured organism; there is a lack of lab work in the protein chemistry from many CODHs given their host organisms' affinity toward extremophilic conditions. In this study a CooS from an uncultured organism is characterized via expression in *E. coli*. This characterization will help further the understanding of how a CODH functions and the activity of these enzymes. Ideally this knowledge can be applied to create methods of utilizing CODHs to better understand their ecology and evolution as well as employ them towards renewable energy, or pollution cleanup. We hypothesized that the CooS from the Hydrothermarchaeota JdFR-17 would perform CO oxidation at faster rates than CO<sub>2</sub> reduction and that the CooS would have an optimal temperature at elevated temperatures in the thermophilic temperature ranges.

## 2 Methods

### 2.1 Gene Constructs and Vector Creation

Hydrothermarchaeota JdFR-17 contains multiple *cooS* genes as well as other proteins involved in the Wood-Ljungdahl pathway. Hydrothermarchaeota JdFR-17 contains the *cdhA*, which is a bifunctional *cooS* homolog associated with the Wood-Ljungdahl pathway. Hydrothermarchaeota JdFR-17 also contains two monofunctional *cooS* sequences [32, 33]. While Hydrothermarchaeota JdFR-17 encodes two *cooC* homologs for maturation of the CooS, these *cooC* homologs [33, 34] are associated with bifunctional CODH gene cluster and not the monofunctional *cooS* gene cluster.

The Hydrothermarchaeota JdFR-17 metagenome assembled genome is highly fragmented and only 57% complete. Here we had one of the monofunctional Hydrothermarchaeota JdFR-17 *cooS* genes synthesized and codon optimized for balanced codon usage for expression in *E. coli* [33].

We also synthesized the *cooC* from the archaeon *Archaeoglobus fulgidis* with balanced codon usage for expression in *E. coli*. These constructs were chosen as part of a larger project to character the CODH function across the seven CODH clades. It was for this reason that the *Archaeoglobus fulgidis cooC* was chosen rather than the *cooC* from Hydrothermarchaeota JdFR-17.

The Hydrothermarchaeota JdFR-17 *cooS* was cloned into the pET15b expression vector with ampicillin resistance. The *Archaeoglobus fulgidis cooC* was cloned into the pACYCDuet-1 expression vector which has chloramphenicol resistance.

The CooS was his-tagged on the N-terminus of the protein. Plasmids were provided in *E. coli* Top10 cells and placed on a master plate. The constructs used in this experiment were labeled as Batch292\_p009 (A2) construct, which was the *cooS*. The *cooC* was the Batch292\_p007 (G1) construct. The full list of constructs is shown in the appendix (A.1).

Table 1: Shows the different vectors that were created in the JGI batch and their antibiotic resistances

Vector	Antibiotic Resistance
pet15b	Ampicillin
pACYCDuet-1	Chloramphenicol

### 2.2 Transformations and Protein Purification

#### 2.2.1 Transformations

Plasmids were purified from the *E. coli* Top10 glycerol stocks. The transformed *E. coli* from the master plate were used for plasmid purification. The cells were grown up aerobically in a 37 °C shaker to >0.6 OD in 5 mL LB medium. Antibiotics were present

in the medium with 50 µg/mL ampicillin for *Hydrothermarchaeota* Jd-FR17 *cooS* (A2), and 50 µg/mL chloramphenicol for *Archaeoglobus fulgidis cooC* (G1). These antibiotic concentrations were kept constant for all the growth media in the experiment. The cells grown from A2 and G1 were both purified in the same way. The purification was done using a Zymo Research Zippy™ plasmid mini-prep kit. All the instructions were followed except that 2 mL cells were spun down before adding the first lysis buffer [35]. This increased the amount of plasmid in the eluant.

The purified plasmids were then transformed into the expression strain BL21 (DE3). The G1 plasmid was purified first, and once the plasmid was purified 5 µL of plasmid was mixed with 20 µL of BL21 (DE3) cells ordered. The mixture was kept on ice for 30 min. The mixture was heat shocked for 45 seconds at 42 °C in a water bath. The mixture was then returned to ice for 2 min. 600 µL of SOC media was then added to the mixture. This was incubated in a 37 °C shaker for 1 hour to allow outgrowth. 250 µL of the cells were then streaked on a LB plate with chloramphenicol and incubated overnight at 37 °C. 100 µL of the cells were placed in 10 mL liquid LB media with chloramphenicol and grown to 0.6 OD. This was to confirm the chloramphenicol resistance and to provide cells for the next transformation.

Once growth on chloramphenicol was confirmed, the G1-containing cells were made chemically competent and stocks to be stored in the -80°C freezer were made from the extra cells grown. The G1-containing cells were then transformed with the A2 plasmid following a similar protocol to the one described for transformation of the *E. coli* BL21 (DE3) with G1. Once transformation was complete, these cells were grown on ampicillin/chloramphenicol LB media. One liquid culture and one plate was created in the same fashion as above.

## 2.2.2 Expression and Column Purification

A2/G1 transformed cells were grown aerobically; shaking at 37 °C overnight in 10 mL of an autoinduction media to induce expression of the proteins. Autoinduction media was prepared with the specifications from the protocol found in A.2, [36] with chloramphenicol and ampicillin added. Autoinduction medium is a medium for heterologous protein expression that contains glucose and lactose in the medium and allows for high cell densities compared to IPTG induction [37]. This allows for a constant expression as the cells grow, rather than inducing expression at a certain cell density like in IPTG induction.

The 10 mL starter culture was then used to inoculate 100 mL of autoinduction media and grown overnight in the same fashion. The 100 mL culture was aliquoted into 50mL falcon tubes and then centrifuged for 10 min at 12000 x g. The supernatant was discarded, and the pellets were each resuspended in 10 ml of equilibration buffer (300mM NaCl, 50mM Tris/HCl-pH 7.5, 1mM NaDT, and 10mM imidazole). The resuspended pellets were combined and then sonicated for 3 seconds 3 times on ice to lyse the cells. The 20 mL of sonicated lysate was then spun down at 12000 x g and the supernatant was collected. This supernatant was used for the affinity purification.

The column was prepared by adding 6 mL HisPur™ Ni-NTA Resin slurry to a column. This was then drained and then the column was rinsed with 7.5 mL equilibration buffer. The supernatant was added to the column. Once the supernatant had moved through the column, 15 mL wash buffer was added (300 mM NaCl, 50 mM Tris/HCl-pH 7.5, 1 mM NaDT, and 25 mM imidazole). The flow through from the wash buffer was collected for an analysis using polyacrylamide gel electrophoresis. Once the wash solution had passed through the column, 7.5 mL of elution buffer was added (300 mM NaCl, 50 mM Tris/HCl-pH 7.5, 1 mM NaDT, and 250 mM imidazole). The elution through was collected. All fractions were analyzed by polyacrylamide gel electrophoresis.

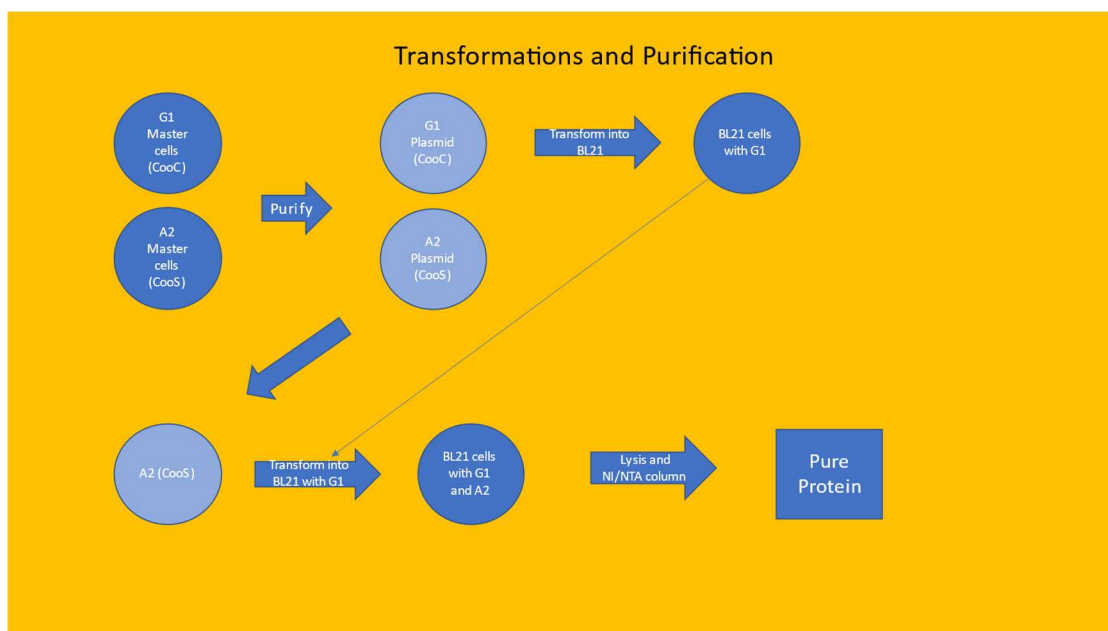


Figure 5: Shows the process of transforming the vectors into BL 21 and the purification of the protein from those cells.

## 2.3 Gel

The gel was loaded to view proteins produced under multiple different conditions to determine expression and solubility. In the first well of the gel was a Precision-Plus Protein Dual Color Standard ladder. Wells 2, 3, and 4 contained some variation of A2xG1 cells. The cells were grown in autoinduction media and then 1.5 mL cells were sonicated. 20  $\mu$ L of the lysate was used for well 2. This was added to 20  $\mu$ L of gel loading buffer and vortexed.

The remaining lysate was spun down at 12,000 xg for 5 minutes. The supernatant was collected for well 3. 20  $\mu$ L supernatant and 20  $\mu$ L of loading buffer were vortexed

together. The pellet from the lysate was resuspended in 20  $\mu$ L 3 M urea and 20  $\mu$ L of loading buffer and vortexed.

20  $\mu$ L of the Ni/NTA column eluant was reserved for the 5<sup>th</sup> well. 20  $\mu$ L loading buffer was added to the eluant and then vortexed.

All the samples were heated at 95°C for 10 minutes. During this time the gel cassette was connected to the electricity source and submerged in 1x Tris-Glycine buffer. The samples were immediately spun for 5 min at 12,000 xg and then loaded into the gel. The gel was a Mini-Protean TGx Precast gel from BioRad. The gel was run and once the run was complete the gel was placed in Coomassie blue for 30 minutes shaking. The Coomassie was destained in a solution of 50% methanol, 40% water and 10% acetic acid by volume.

## **2.4 Protein Characterization**

### **2.4.1 Dialysis and BCA Quantification**

The purified protein was dialyzed to remove the imidazole from the protein mixture prior to additional analyses. The eluant from the column was placed in dialysis tubing with a 10,000 Da molecular weight cutoff and sealed. The dialysis tubing was then placed in 250 mL dialysis buffer overnight (300 mM NaCl, 50 mM Tris/HCl-pH 7.5, and 1 mM NaDT).

The protein was then quantified using a BCA spectrophotometry assay. The kit used for the assay was the Pierce™ BCA Protein Assay Kit by Thermo Scientific. The raw data and calculations from the BCA assay are shown in the appendix in A.3. The protocol for the assay can be found in A.4. The BCA assay is a color change assay that involved creating a standard curve using Bovine Serum Albumin (BSA) and comparing the A2xG1 sample to that. The protein sample is added to a working solution and then heated at 60°C for 30 minutes and then the color change was measured at 562 nm on a spectrophotometer.

### **2.4.2 Melt Curves**

We hypothesized that the protein from Hydrothermarchaeota would function at elevated temperatures. Therefore, upon the purification and quantification of the protein it was necessary to determine the melting temperature of the protein. This was done via a Thermal Shift Assay.

A thermal shift assay uses a fluorescent dye that binds to protein and is released upon the melting of a protein. The samples are diluted to 2 mg/mL and placed in a 96 well plate with dye and a ligand. The 96 well plate is placed in a rtPCR machine and heated gradually. This causes the proteins to denature and dye to be released. The signal is measured, and a peak should appear upon protein melting.

The PCR machine used was a Thermo Scientific StepOnePlus. The assay kit was a Protein Thermal Shift™ Dye Kit from Thermo Scientific. The protocol from the dye kit was followed and applied to the A2xG1 protein sample and 2mg/mL BSA. A no protein control was also used. The manual for a complete description of the assay sample preparation can be found in S5 [38].

The run was configured using the StepOnePlus software. The run started at 25 °C and was ramped by 1% per minute until reaching 99 °C and held for 2 minutes. The dye was a ROX dye and there was no quencher. The PCR settings manual can be found in S6. All specifications were followed according to the directions for the StepOnePlus system. [39].

### 2.4.3 Enzyme Activity and Whole Cell Assays

We hypothesized that the oxidation of CO would be the dominant reaction and that the extent of the oxidation of CO would exceed that of the reduction of CO<sub>2</sub>. To test this hypothesis, we performed whole cell assays. The assays were run at 37 °C due to similarity to other experiments[27–30], and 70 °C, 80 °C, and 90 °C to account for the assumed high optimal temperature of the CooS. We chose to do whole cell assays relative to standard enzyme assays due to the oxygen sensitivity of the enzyme.

Oxygen is an inhibitor of the Nickel based CODH and short-term exposure to oxygen can cause permanent inactivation of the enzyme. These enzyme assays were designed to be oxygen free. The media was inoculated aerobically, however upon inoculation the tube was sealed with a butyl stopper and flushed with nitrogen and vacuumed via a gassing station. The resulting headspace would likely still have a very small amount of oxygen, but the *E. coli* would metabolize this early in the growth phase and cells that were created later in the growth phase or stationary phase would not be exposed to oxygen.

For each growth temperature there were 2 directions of the reaction: forward and reverse. The forward reaction would contain no CO and was meant for measuring CO production. The equation for the forward reaction is  $\text{CO}_2 + 2\text{H}^+ + 2\text{e}^- \rightarrow \text{CO} + \text{H}_2\text{O}$ . The reverse reaction would eventually be injected with CO and would measure CO consumption. The equation for the reverse reaction is  $\text{CO} + \text{H}_2\text{O} \rightarrow \text{CO}_2 + 2\text{H}^+ + 2\text{e}^-$ .

The forward and reverse reactions each had 3 replicates and 3 controls. This led to a total of 12 tubes per temperature. The control tubes were filled with 10 mL autoinduction media and inoculated with 10 µL BL21. The experimental tubes were inoculated with 10 µL A2xG1 *E. coli* and grown in 10 mL autoinduction medium with ampicillin and chloramphenicol. These cultures were sealed using a butyl stopper in a glass batch tube. To make the cultures anaerobic, the headspace was evacuated using a vacuum pump the headspace was then flushed with nitrogen. A total of ten cycles of evacuation and flushing with nitrogen was performed to remove oxygen from the headspace. The pressure in the headspace was equilibrated to approximately 1 atm after the final flush with nitrogen. The cultures were then grown in a 37 °C incubator overnight to an OD 600 of 0.589 – 0.904. At this time 100 µL of the headspace of the

tube was removed and used to measure CO and CO<sub>2</sub> using an SRI8610C Gas Chromatograph.

Each measurement took 6 minutes, and the next measurement was started directly after the first. The forward reactions were measured first. To measure CO oxidation, 50 µL of CO were added to the headspace before the first measurement. After addition of CO, 100 µl of headspace was removed and measured. After the first measurement, the tubes were placed into the appropriate temperature incubator. For the 37 °C test the tubes were placed back in the 37 °C incubator. For the 70 °C, 80 °C, and 90 °C tests the tube was placed in a water bath of that temperature to ensure rapid equilibration of the tubes to the temperature. The measurements were then repeated 2 more times. Since each measurement was about 6 minutes, by the time the last reverse measurement was taken, the first forward measurement had been incubating for around 90 minutes. A complete table of the timestamps for each measurement is included in A.7.

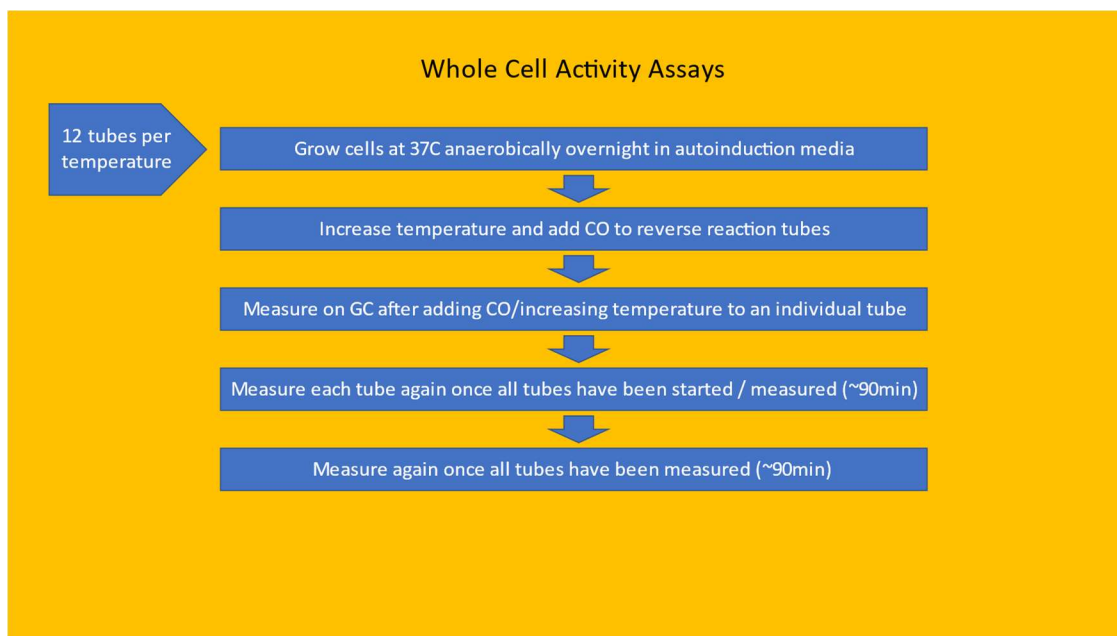


Figure 6: Shows the process used to run a whole cell rate assay.

#### 2.4.4 Gas Chromatography

To determine the concentration of CO and CO<sub>2</sub> in the headspace of the incubations, we used a GC with a HayeSep D column attached to a Thermal Conductivity Detector (TCD) and a Flame Ionization Detector (FID) attached to a methanizer. All the measurements were collected on the FID due to its increased sensitivity. The GC settings can be seen in table 2. The peaks were analyzed using PeakSimple software. The run was configured with a 6-minute hold at 60 degrees C. The signal was zeroed at the beginning of each run.

Table 2: Shows the settings used in the detection of the gasses

Setting	Status
Carrier 1 (Nitrogen)	23 psi
Hydrogen 1	20 psi
Air 1	4 psi
Valve 1	100 °C
TCD Amplifier	High
TCD Cell	200 °C
Detector 2	380°C

Prior to the assay measurements, a standard curve for CO and CO<sub>2</sub> detection were made using injections of 500-25 µL Scott™ Mini-Mix gas standard. This contained a mixture of 1% CO, CO<sub>2</sub>, methane, oxygen, hydrogen. The base for the gas standard was nitrogen.

The data was analyzed using PeakSimple software. CO was measured in the peak areas from time 1.62 to time 2.12. CO<sub>2</sub> was measured in the peak areas from time 5.03 to time 5.53. Examples of the GC chromatograms are included in the appendix. A Chromatogram of the 37°C forward reaction showing a CO<sub>2</sub> peak is in A.8 An example of the 37°C reverse reaction showing CO and CO<sub>2</sub> peaks in in A.9.

#### 2.4.5 ICP-OES

Previous work has shown that presence of nickel in the active site of the enzyme is essential for activity of the CooS. To confirm that Ni was present in the protein we used ICP – OES to determine if the enzyme had the metals needed to be functional. In this analysis the samples were run by the LEAF lab at MTU. There were two controls; a no metal control provided by the lab and the dialysis buffer used in the previous steps. The A2xG1 pure dialyzed protein was diluted 1/10 to 0.0593 mg/mL in dialysis buffer and sent to the lab along with 3 other CooS proteins that had been purified in other experiments.



### 3 Results

#### 3.1 Genomic Context of the CODH gene clusters in Hydrothermarchaeota JdFR-17

The monofunctional *cooS* from Hydrothermarchaeota used in this study is a 1929 bp region in a 6256 bp scaffold. This scaffold included 7 other genes. These were a PRC Barrel Domain containing protein, a predicted hydrocarbon binding protein, an oligoribonuclease NrnB or cAMP/CGMP phosphodiesterase, a Sulphur transport protein, a sec-independent protein translocase protein TatC, and 2 hypothetical proteins. The gene map of the *cooS* is included below (Figure 7).

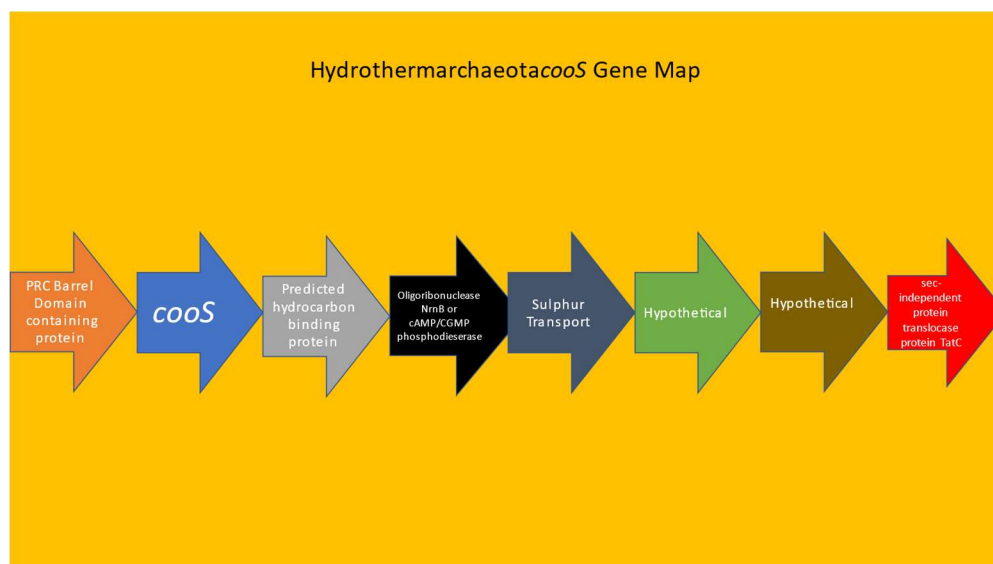


Figure 7: Shows a gene map of the *cooS* used in this study.

#### 3.2 Expression and Purification of the Hydrothermarchaeota CooS Protein

The transformations were deemed successful after the cultures showed the ability to grow on antibiotic plates and antibiotic media.

The BCA quantification of the purified dialyzed CooS protein yielded results that indicate 0.593 mg/mL. The table of the raw data and calculations can be found in the appendix (A.3).

The Ni/NTA column eluent was run through the SDS page gel along with the cells grown in autoinduction media, and the soluble and insoluble fractions of those induced cells. The gel's results showed a distinct band under the 70 kDa marker on the ladder in lane 5. This band in the eluant is where the expected CooS would fall given it was soluble and expressed in the *E. coli*. This band is also present in all the wells of the gel that have sample in them. CooS is expected to occur at 67 kDa and has shown to be present in the

soluble fraction and the insoluble fraction in previous studies [29, 40].

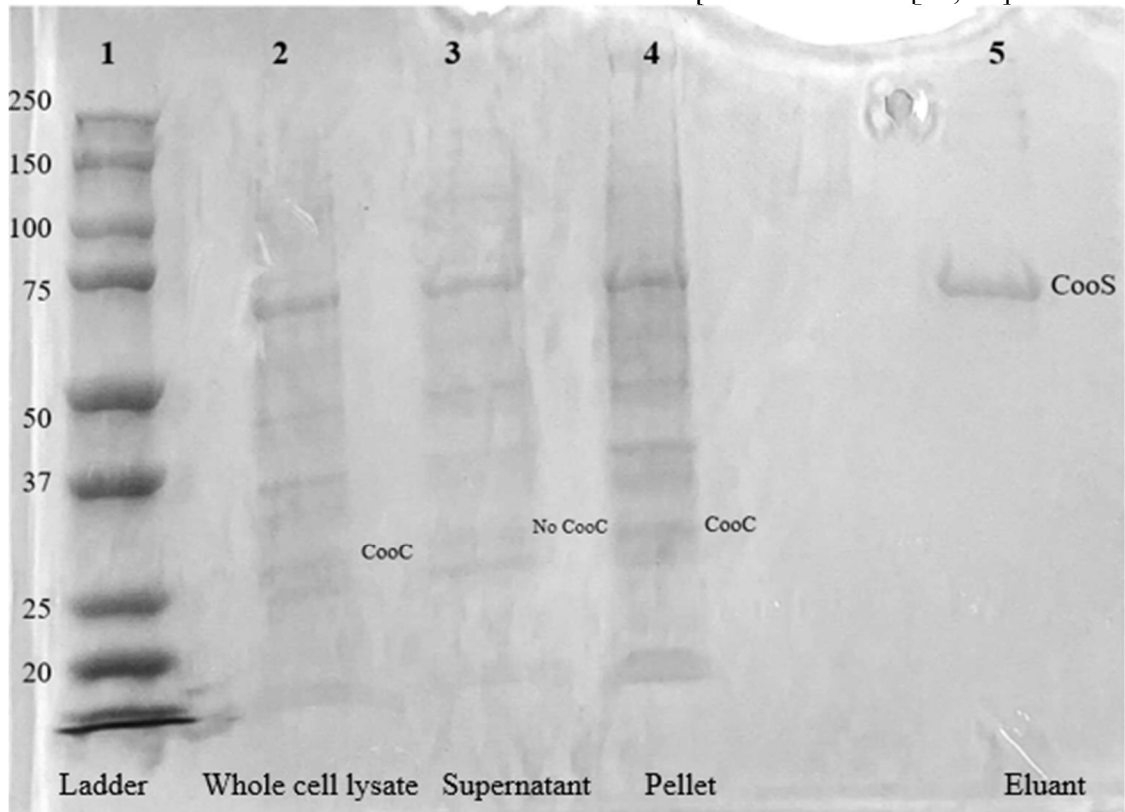


Figure 8: Shows a gel image of the A2xG1. 1 – Ladder, 2 – A2xG1 cells grown in autoinduction media and sonicated, 3 – Supernatant fraction from A2xG1 grown in autoinduction media and sonicated, 4 -resuspended pellet from A2xG1 grown in autoinduction media and sonicated, 5 – Elution flow through from the Ni-NTA column.

The CooC band would be expected to occur at 26 kDa. There are 2 bands in columns 2, 3, and 4 that are in the space where CooC is expected show up. The CooC however, would not be present in the eluant. This is because CooC has only been documented to appear in the insoluble fraction of the cell lysate [28, 40]. Since columns 3 and 5 were only from the soluble fraction, it is unlikely CooC would be present [40].

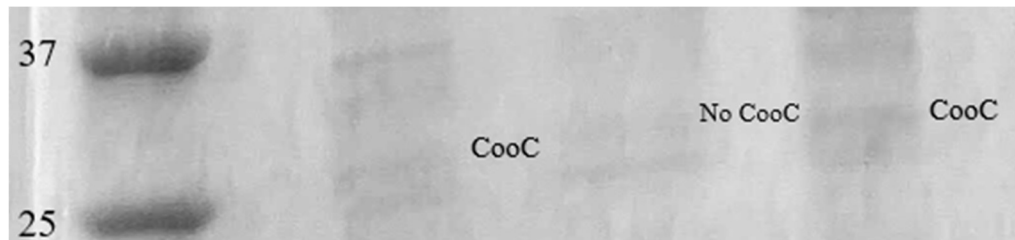


Figure 9: Shows a zoomed version of Figure 8. This focuses on the region where the expected CooC would be. The CooC is the larger of the two bands that are present near the 25kDa marker

### **3.3 Melt Curves**

Melt curves were used in attempt to determine a  $T_m$  for the CODH. This would assist in determining the optimal temperature of the protein. The melt curve was run in the same batch for all samples, but the results have been split to make viewing easier. Figure 10 shows the results for the no protein controls. 3 no protein controls are represented by different. The signal in the derivative reporter for these peaks at about 9000 and has a minimum of about -4000. This gives a reference to the signal noise, which can then be accounted for in the experimental results.

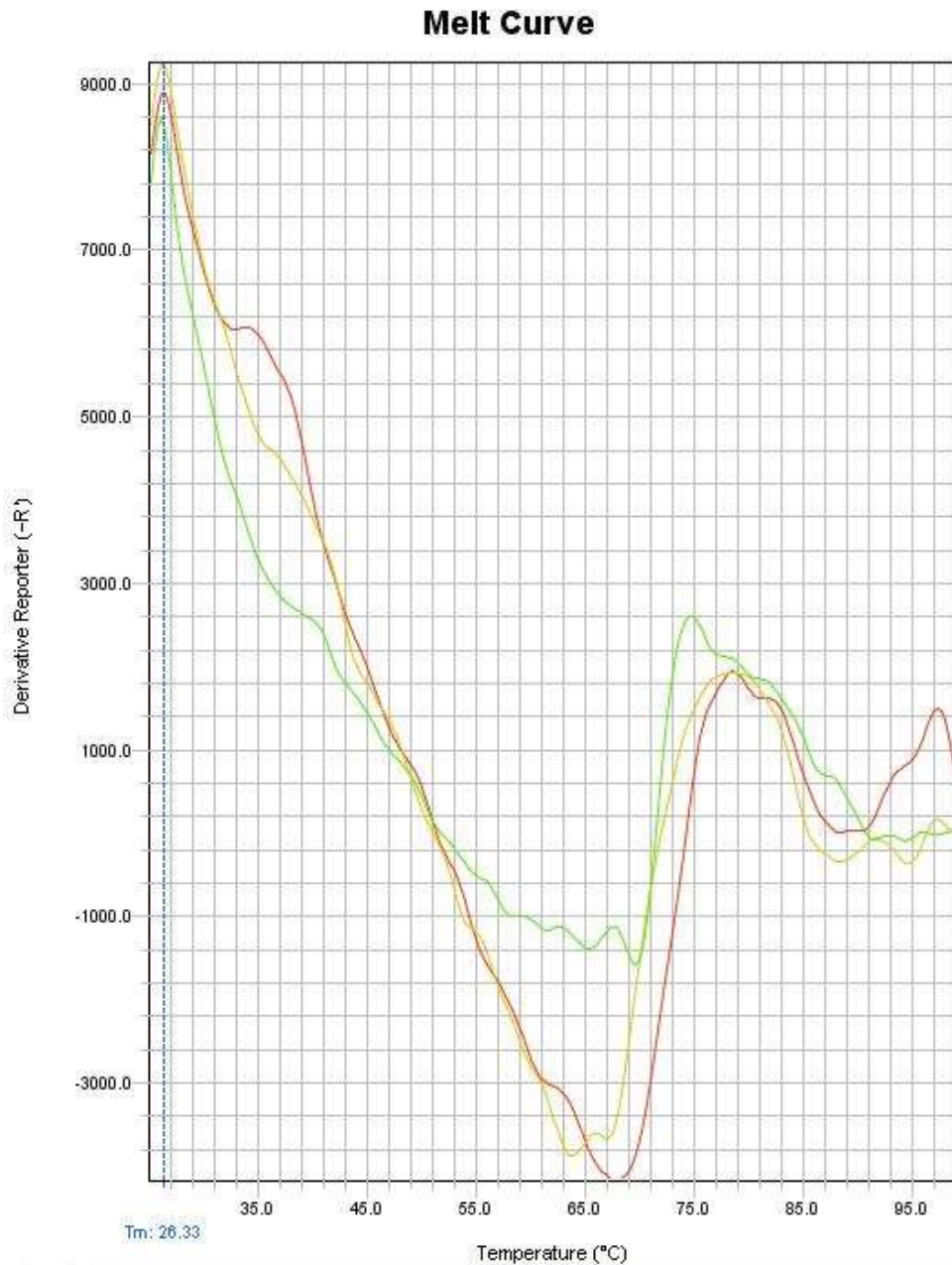


Figure 10: Shows the melt curve of the 3 blank cells used as controls.

The melt curve for the BSA and the A2xG1 purified protein is shown in figure 11. For the thermal shift assay a positive peak is expected when the protein melts, as melting of the protein exposes hydrophobic residues in the protein's core. This fluorescent dye in the thermal shift assay will bind to hydrophobic residues and begin to fluoresce. So, the positive peak of the derivative of the fluorescence indicates binding of the dye to the melted protein. Large positive peaks were present in the BSA samples. The most prominent peaks were around the 65 °C mark and almost 100,000 signal. The melting

point of BSA is 63 °C [41]. This indicates the thermal shift assay was successful and can be used to determine the melting temperature for BSA.

The A2xG1 samples have lower signal and have negative peaks. The signal for the 3 A2xG1 peaks around 35,000 at 95°C. The distinction in these peaks is much less than the BSA. The Peak in the BSA is over 3x as large as the A2xG1 peak. The A2xG1 peak could still be a relevant peak however it is not as defined as the BSA peak.

Since there were no significant positive peaks in the A2xG1 purification, the melting point of the CooS was inconclusive. However, since the BSA and the CooS were at similar concentrations and had similar molecular weights, the lack of positive peak could indicate that the melting of the CooS protein occurred at higher temperatures than the 95 °C tested in the thermal shift assay.

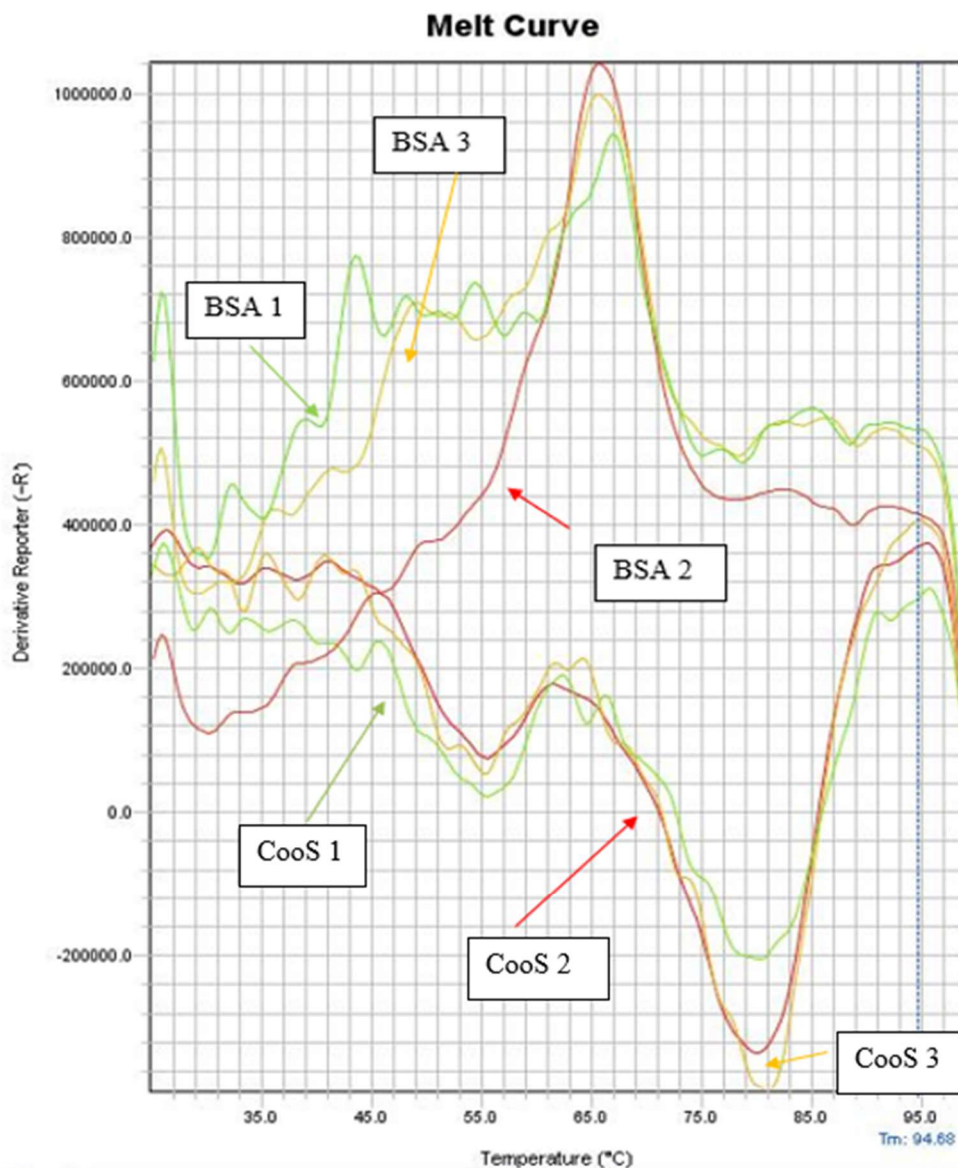


Figure 11: Shows the melt curve from BSA (BSA) and the purified protein (CooS).

### 3.4 CO and CO<sub>2</sub> Activity Assays

#### 3.4.1 Carbon Monoxide Detection

##### 3.4.1.1 CO Oxidation in the Reverse Reaction

To determine the activity of the expressed CooS, we used whole cell assays for detection of activity. The CooS can catalyze both the oxidation of CO to CO<sub>2</sub> and the reduction of CO<sub>2</sub> to CO. While either CO or CO<sub>2</sub> could be measured to determine CODH activity, we have focused our measurements on the detection of CO for the assays. The measurement of CO production or consumption had fewer confounding

variables compared to CO<sub>2</sub> measurements. The CO<sub>2</sub> measurements were affected by the growth of the *E. coli*, as *E. coli* will produce CO<sub>2</sub> in their metabolism of the growth media. The reverse reaction ( $\text{CO} + \text{H}_2\text{O} \rightarrow \text{CO}_2 + 2\text{e}^- + 2\text{H}^+$ ) was measured by determining the removal of CO in the solution.

For there to be evidence of activity for CO oxidation we would expect to see a significant decrease in CO in the experimental tubes.

The only temperature where activity was possibly indicated was the 37-degree experiment. This experiment showed a large difference of CO from the controls. A one tailed Welch's T-test with significance  $\alpha = 0.05$  and 2 degrees of freedom was run on the data from this run and indicated that this change was not statistically significant. The T-test statistic for the percent change from start to finish was 0.128. The T-test statistic for the percent change from start to measurement 1 was 0.0877. The change likely occurred due to a high variance in the sample.

Raw data including calculations can be found in the appendix under A.10

All the other tubes saw a similar decrease in CO except for the experimental 70-degree tubes. These saw a large increase in CO but had very high error values. It is unexpected to see an increase in CO in this experiment. A one tailed Welch's T-test with significance  $\alpha = 0.05$  and 2 degrees of freedom was run on the data from this run and indicated that there was no statistically significant change in CO levels. The T-test statistic for the percent change from start to finish was 0.180. The T-test statistic for the percent change from start to measurement 1 was 0.152. This is due to the high variance in the sample.

Raw data including calculations can be found in the appendix under A.10

Table 3: Shows the results of the CO changes in the reverse reaction whole cell assays.

Each percent change corresponds to the average of the percent changes from the 3 replicates for each temperature. 0-1 is the percent change between start and time 1, and 0-2 is from start to time 2. The standard deviation was reported from the raw GC data used to calculate the percentages.

Change in CO Reverse Reaction		Control	Standard Deviation of Raw Data	Experimental	Standard Deviation of Raw Data
37	0-1	-7%	42%	-56%	6%
	0-2	-18%	48%	-62%	15%
70	0-1	-80%	10%	81%	193%
	0-2	-60%	31%	67%	169%
80	0-1	-85%	9%	-65%	38%
	0-2	-83%	11%	-54%	53%
90	0-1	-84%	6%	-72%	17%

	<b>0-2</b>	-84%	7%	-71%	19%
--	------------	------	----	------	-----

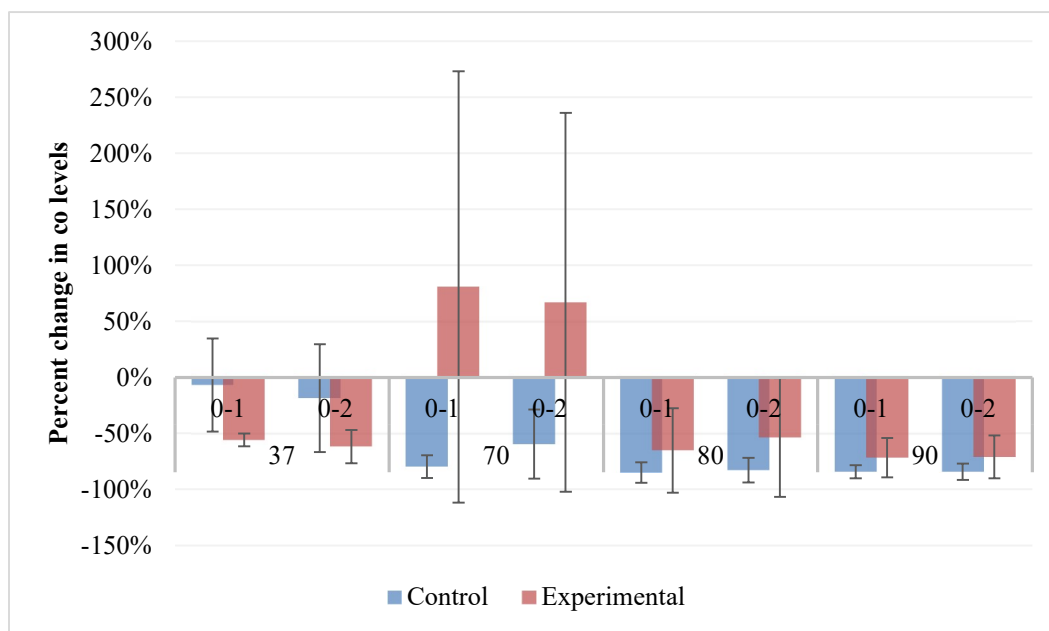


Figure 12: Graph of table 3. This shows the results for the percent change in CO levels in the reverse reactions. The x-axis separates percent changes by the different temperatures, and further by the different time stamps. 0-1 is the percent change between start and time 1, and 0-2 is from start to time 2. The error bars are calculated from the standard deviation of the raw data.

### 3.4.1.2 CO<sub>2</sub> Reduction in the Forward Reaction

The forward reaction for the CODH ( $\text{CO}_2 + 2\text{e}^- + 2\text{H}^+ \rightarrow \text{CO} + \text{H}_2\text{O}$ ) would result in the formation of CO from reduction of CO<sub>2</sub>. Activity would be detected if CO appeared in solution after the cells were placed in the experimental conditions. No CO peaks were measured in the entire experiment. This indicates no activity in the forward direction.

Table 4: Shows the CO levels for the forward reaction. No peaks were detected throughout the experiment.

Change in CO Forward		Control	Experimental
37	0-1	0%	0%
	0-2	0%	0%
70	0-1	0%	0%
	0-2	0%	0%
80	0-1	0%	0%
	0-2	0%	0%
90	0-1	0%	0%



	<b>0-2</b>	0%	0%
--	------------	----	----

### 3.4.2 CO<sub>2</sub> Detection

We also measured CO<sub>2</sub> in our experiments in attempt to confirm the CO results. CO<sub>2</sub> detection would have been coupled with the CO detection if the CO levels indicated signs of activity. Since no significant changes occurred, the CO<sub>2</sub> level changes were not deemed to be as useful. Furthermore, CO<sub>2</sub> measurements were highly variable. This could be in part due to the high levels of CO<sub>2</sub> produced as *E. coli* consumes media.

#### 3.4.2.1 CO<sub>2</sub> in the Reverse Reaction

If the reverse reaction showed signs of CO consumption; a CO<sub>2</sub> production would be the second thing to look for regarding activity. The CO levels did not decrease in the experiment, so CO<sub>2</sub> changes were likely due to *E. coli* metabolizing media.

Table 5: Shows the CO<sub>2</sub> levels for the reverse reaction. Each percent change corresponds to the average of the percent changes from the 3 replicates for each temperature. 0-1 is the percent change between start and time 1, and 0-2 is from start to time 2. The standard deviation was reported from the raw GC data used to calculate the percentages.

Change in CO <sub>2</sub> Reverse		Control	Standard Deviation of Raw Data	Experimental	Standard Deviation of Raw Data
<b>37</b>	<b>0-1</b>	14%	49%	-29%	25%
	<b>0-2</b>	-3%	40%	-19%	44%
<b>70</b>	<b>0-1</b>	-21%	29%	138%	266%
	<b>0-2</b>	16%	41%	135%	250%
<b>80</b>	<b>0-1</b>	-38%	17%	-6%	12%
	<b>0-2</b>	-21%	35%	44%	42%
<b>90</b>	<b>0-1</b>	-32%	7%	-2%	42%
	<b>0-2</b>	-31%	4%	15%	73%

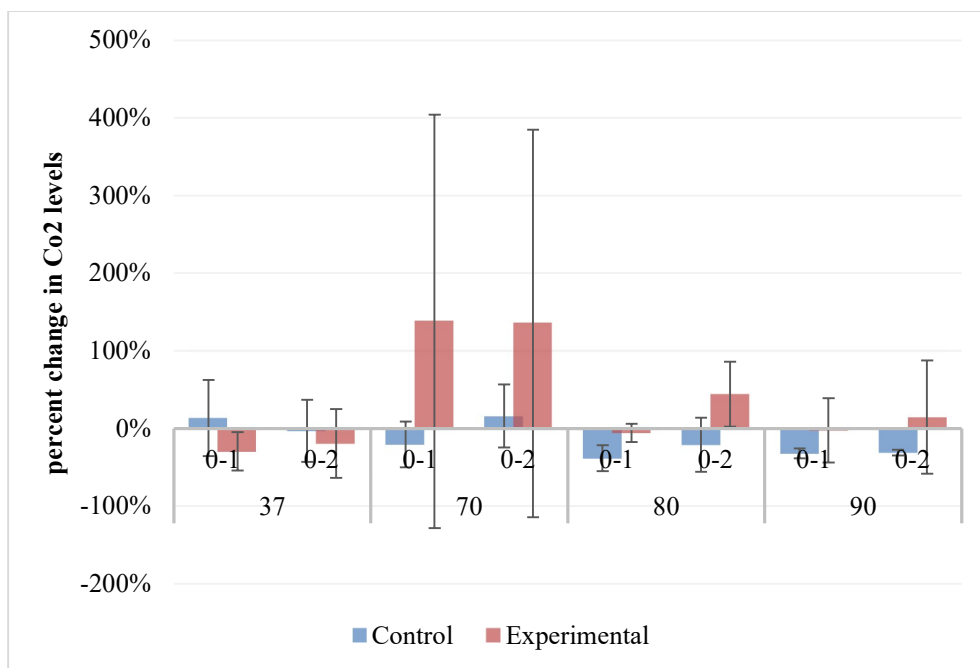


Figure 13: Graph of table 5. This shows the results for the percent change in CO<sub>2</sub> levels in the reverse reactions. The x-axis separates percent changes by the different temperatures, and further by the different time stamps. 0-1 is the percent change between start and time 1, and 0-2 is from start to time 2. The error bars are calculated from the standard deviation of the raw data.

### 3.4.2.2 CO<sub>2</sub> in the Forward Reaction

If the forward reaction showed signs of CO production; a CO<sub>2</sub> consumption would be the second thing to look for regarding activity. The CO levels did not increase in the experiment, so CO<sub>2</sub> changes were likely due to *E. coli* metabolizing media.

Table 6: Shows the results of the CO<sub>2</sub> changes for the reverse reaction whole cell assays.

Each percent change corresponds to the average of the percent changes from the 3 replicates for each temperature. 0-1 is the percent change between start and time 1, and 0-2 is from start to time 2. The standard deviation was reported from the raw GC data used to calculate the percentages.

Change in CO <sub>2</sub> Forward		Control	Standard Deviation of Raw Data	Experimental	Standard Deviation of Raw Data
37	0-1	50%	121%	84%	144%
	0-2	112%	212%	62%	113%
70	0-1	1%	5%	9%	77%
	0-2	-24%	53%	36%	43%
80	0-1	37%	16%	24%	25%
	0-2	38%	9%	29%	10%
90	0-1	16%	12%	33%	59%

	0-2	25%	16%	30%	9%
--	-----	-----	-----	-----	----

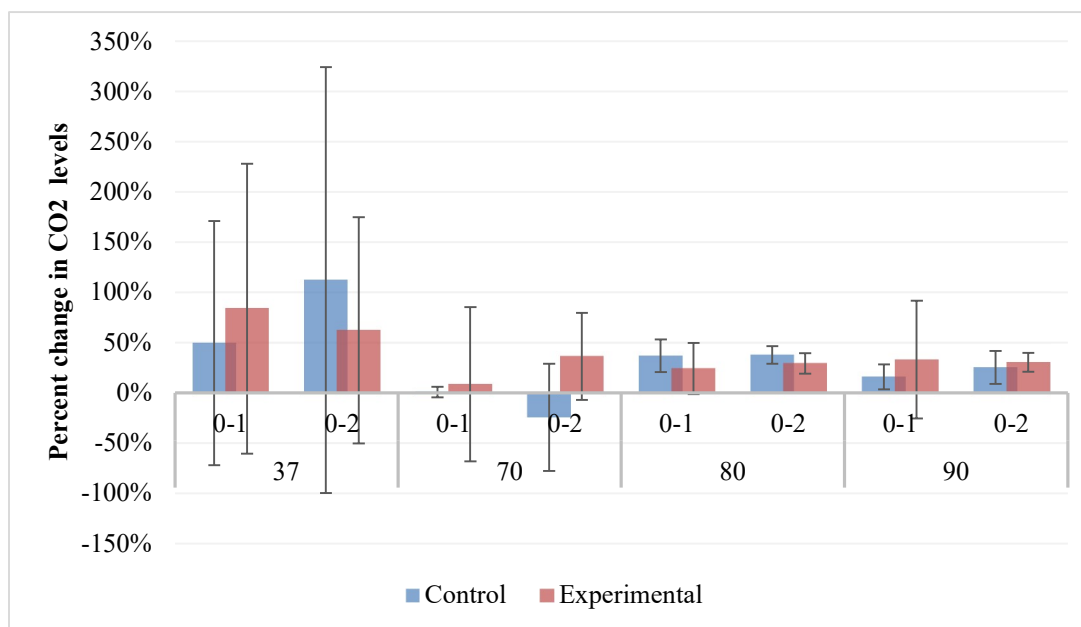


Figure 14: Graph of table 6. This shows the results for the percent change in CO<sub>2</sub> levels in the forward reactions. The x-axis separates percent changes by the different temperatures, and further by the different time stamps. 0-1 is the percent change between start and time 1, and 0-2 is from start to time 2. The error bars are calculated from the standard deviation of the raw data.

### 3.5 ICP-OES

To potentially explain the non-significant changes in CO and CO<sub>2</sub> observed in the whole cell assays, we sought to determine if the *CooC* from *Archaeoglobus fulgidis* could successfully insert nickel into the *CooS* from Hydrothermarchaeota JdFR-17. ICP-OES was run on the purified protein to detect the presence of iron and nickel in the protein. These metals are vital for the proper function of the protein. *CooC* inserts Nickel into the active site of the protein, so a positive ICP-OES result could indicate that *CooC* is present in the cell and functional for insertion of Ni into the *CooS* protein. The variability of the full cell assay results led to this assay being run to determine if *CooC* was present. The lack of Ni in the *CooS* would hinder activity if the *CooC* was not inserting nickel in the active site of the protein. The ICP-OES analysis results are highlighted in table 7. These results indicated that the protein contains two times the Ni concentration of the Ni in the dialysis buffer. The protein concentration used for ICP-OES was 0.0593 mg/ml. This increased concentration of nickel in the protein suggests successful insertion of the Ni in the *CooS* by *CooC*. Additionally, the iron was below detection in the dialysis buffer but was present at 0.01 mg/L in the diluted protein solution. Also indicating that iron was successfully inserted into the *CooS* protein.

Table 7: Results from the ICP-OES analysis. The batch was a mix of purified co-transformed proteins and blanks.

<b>Sample ID</b>	<b>Fe 238.204</b>	<b>Ni 231.604</b>	
lab BLANK	<0.01	<0.01	mg/L
BLANK	<0.01	0.02	mg/L
A3XF1	0.01	0.03	mg/L
A4XA1	<0.01	0.02	mg/L
A4XF1	<0.01	0.02	mg/L
A2XG1	0.01	0.04	mg/L

## 4 Discussion

### 4.1 Expression

Our results indicate that the CooS from the uncultured archaeon Hydrothermarchaeota JdFR-17 can be expressed, solubilized, and purified in the heterologous bacterial host *E. coli*. The gel results from the eluant of the column shows a very clear band under the 70kDa marker. This matches the expected molecular weight of 67 kDa for the CooS. This is a good indicator that the CooS has been expressed successfully and is soluble. This is significant because solubility of CooS is required for function of the enzyme. [40] This is likely one of the first examples of the expression of a CooS from an uncultured microorganism in *E. coli*. Previous studies have shown that CooS proteins from diverse bacterial and archaeal hosts can be expressed in *E. coli* [29, 40]

The CooC expression is likely due to the presence of the bands above the 25 kDa marker in the wells with the induced cell lysate and the pellet. There may be a band present in the soluble fraction of the induced cells, however it seems to be less defined. This would be expected due to CooC being a insoluble protein [28, 40]. It has been stated that CooC presence has had a role in increasing the expression of CooS, but CooC has not been shown to be soluble[28, 40]. The gel alone was not enough to confirm that CooC was present and able to insert Ni into the CooS, so the ICP-OES assay was used to determine if nickel was inserted into the purified protein.

The ICP-OES yielded results that indicate the nickel was present in the CooS purified in the A2xG1 protein sample. This sample has the highest metal concentrations of the batch. It is expected that in a fully mature CODH there will be ‘0.8–1.8 nickel and 16–21 iron atoms per dimer’[28] whereas a CODH without CooC present will see a decrease in the amount of nickel; with only 0-0.5 nickel atoms per dimer [28]. In the ICP-OES analysis performed in this experiment, there was more nickel than iron in the sample. This most likely indicates that the CooC is inserting nickel into the CODH. However, it is to be noted that the iron levels should be higher than the nickel levels given their similar molecular weight.

Our ICP-OES results further indicate that the CooC from a non-related archaeon can insert nickel into the Hydrothermarchaeota CooS protein. Previous heterologous expression work has used the CooC from the same organism that the CooS was derived. The ability of the *Archaeoglobus fulgidis* CooC to insert nickel into the Hydrothermarchaeota JdFR-17 CooS indicates that there is some flexibility for interspecies insertion of Ni into CooS from non-related organism. This finding is important as there is a need for biochemical characterization of proteins from the growing repertoire of uncultured organisms. The ability for the flexibility in accessory proteins is an important step in the development of a system for the screening of CooS proteins from diverse uncultured organisms. Further work should investigate how flexible the accessory proteins are in terms of their ability to insert nickel into CooS proteins from phylogenetically diverse microorganisms.

## 4.2 Activity

We hypothesized that the Hydrothermarchaeota JdFR-17 CooS would have activity at thermophilic temperatures, and that the enzyme would prefer the oxidation of CO over reduction of CO<sub>2</sub>. Our work to determine the stability of the proteins and thermal melting temperature of the CooS protein were inconclusive and no definitive melting temperature was determined. We sought to determine if the protein was active and retained activity at higher temperatures using whole cell assays.

Our activity analysis also provided inconsistent results as a function of temperature. If the enzyme was active for the oxidation of CO, we would expect a significant decrease in CO for the reverse reactions of the enzyme assays. The enzyme assays did not indicate any significant decrease in CO. The CO levels of the reverse reaction solution did decrease most notably in the 37 °C trial of the reverse reaction. This could have indicated that there is CODH activity running in the CO → CO<sub>2</sub> direction at 37 °C. However, the Welch's t-test deems this change to be non-significant. The activity of the enzyme is questionable given the experimental results.

If the enzyme was active for the reduction of CO<sub>2</sub>, we would have expected to observe an increase in the CO levels for the forward reactions. No CO was produced in the forward reactions. This indicates that the CO<sub>2</sub> → CO forward reaction showed no evidence of any activity.

This also serves as evidence against the claim that the reverse reaction had activity. The assays had large margins of error in general, however the one result that was indisputable was that there was no activity in the CO<sub>2</sub> → CO reaction. This information is more certain than the evidence for the activity in the reverse reaction. It would be conservative to deduct that since there is no activity in the forward reaction, and the reverse reaction did not show a significant decrease in CO, that there was no activity in the reverse reaction.

The counterargument could be made that there was still activity in the reverse reaction due to it being the favored mechanism in most previous studies and the large difference between the experimental and control. However, given the variability of the GC measurements, and the lack of statistical significance it would be more conservative to conclude that no activity was present in the whole cell assays.

## 4.3 Future Directions

### 4.3.1 Genetic Considerations

Our results suggest that the expressed CooS resulted in soluble protein with Ni and Fe in the protein, but the activity was rather limited. This limited activity could be due to multiple reasons and there are many avenues this research could take in the future. The first possible reason for the limited activity may be that the CODH needed more accessory proteins to function properly. The other accessory proteins such as CooJ, CooT, and especially CooF may be important to producing a functional CooS. While

previous studies have shown the CooS can oxidize CO and reduce CO<sub>2</sub> on its own, CooF is considered an essential part of the CODH along with the CooS to form a monofunctional CODH [4]. The CooF is essential for transfer of electrons to or from the CooS. This electron shuttle protein could be needed to ensure the redox reactions can occur and the protein can have function. The other accessory genes such as CooT and CooJ could also be similarly required for activity.

A second possible issue in the genes themselves is the fact that the CooS and CooC are from different organisms. Although the CooS is derived from an archaeon, the exact species is different from the species from which the CooC was derived. The CooC is from *Archaeoglobus fulgidis*. *A. fulgidis* is a sulfate reducer with an optimal growing temperature of 76 °C. It is possible that the CooS is from an organism with a higher or lower growing temperature. There is previous research on Hydrothermarchaeota, stating that most of them are found in marine environments near 65 °C. [24] However, given the fact that organisms from this clade have not been fully defined yet, it is possible that the CooS sequence from the samples used in this study were not optimized at those higher temperatures. Furthermore, the genetic context for the Hydrothermarchaeota JdFR-17 CooS and CooC genes is poorly characterized due to the fragmented nature of the Hydrothermarchaeota JdFR-17 MAG.

It is also worth noting that even though the CODH complex was expressed, it may require organism-specific proteins to be able to mature and achieve activity. In previous studies, the proteins from the CODH complexes all came from the same organism. However, the CooS and CooC are derived from organisms from different phyla. Hydrothermarchaeota JdFR-17 is a representative of the Hydrothermarchaeota phylum and *A. fulgidis* is a representative from the Euryarchaeota phylum. Given the complexity of protein folding and the maturation process,[21] it is possible that activity is not possible when using proteins from two different species. Our results indicate presence of Ni in the Hydrothermarchaeota CooS, which suggests successful nickel insertion into the CooS. However, we did not determine the location of Ni in the CooS. Further work may involve co-expressing the Hydrothermarchaeota JdFR-17 CooS and CooC.

Solutions for these problems lie in decreasing the complexity of the experiment. Using the same organism to source all the genes would remove the uncertainty of using genes from phylogenetically distinct organisms and may provide more reliable activity. Also, it would make selecting the optimal temperature much easier. Including the CooF would also be a possible solution to improve the activity of the enzyme. CooF drastically increases the expression of CooS when detecting expression on a SDS gel. [40] Increased expression of the protein should allow for it to be able to consume more substrate and have higher activity.

### **4.3.2 Chemical Inhibition**

A major inhibitor of CODHs is oxygen. It is possible that dissolved oxygen in the solution of the sealed autoinduction media would cause the enzyme to be nonfunctional. CODHs are strictly anaerobic. They are very oxygen sensitive and can lose activity after

1 min of oxygen exposure. Complete irreversible inactivation of the enzyme can occur if prolonged oxygen exposure occurs [42]. This extreme sensitivity makes oxygen removal very important in ensuing enzyme function. The intent of the sealed containers was to clear as much oxygen as possible through flushing with nitrogen gas. Furthermore, oxygen will be consumed during initial growth the *E. coli*. Therefore, we expected that the cells that grew later in the exponential phase would never have been exposed to oxygen. Even though the *E. coli* should have consumed most of the oxygen in the sealed growth containers for the whole cell assays, if any was left behind it could cause severe inhibition of enzyme activity.

A possible route of circumnavigating this could be to create a system like what was outlined in “A simple, large-scale overexpression method of deriving carbon monoxide dehydrogenase II from thermophilic bacterium *Carboxydotherrmus hydrogenoformans*”[29]. This system used a fermenter and TB medium to grow the cells. This would allow for the enzyme to mature properly without the presence of oxygen. This could lead to proper activity of the enzyme. Future work could also perform expression and purification in an anerobic chamber to ensure removal of all oxygen from the system.

Another possible solution would be to use a methyl viologen reduction enzyme assay and a carboxyhemoglobin enzyme assay. These assays would only use the purified protein rather than the entire cell. The cells would have to be grown, expressed, and purified anaerobically. This could be achieved using an anaerobic glovebox and/or a fermenter. The pure protein would then be incubated in a  $\text{NiCl}_2$  solution. For the methyl viologen assay the enzyme would be exposed to CO. The oxidation of CO would produce a color change in the methyl viologen, and that change could be measured on a spectrophotometer. The carboxyhemoglobin assay would be run in a similar fashion, but the protein would be exposed to  $\text{CO}_2$  and reduction activity would be measured by the color change produced by the hemoglobin as CO binds to it. [27–29]. One limitation of whole cell assays with thermophilic proteins are that thermal denaturation of proteins from *E. coli* at elevated temperatures may affect the activity of the thermophilic enzymes. While the thermophilic enzymes may remain stable at higher temperatures, it's possible that denatured *E. coli* proteins may affect activity of the CooS. Using enzyme assays with purified proteins removes complications of other proteins in the solution.

These solutions may help with some of the extreme variability in the results that was observed. Overall, the complexity of the experimental system needs to be reduced to mitigate the variability.

## 4.4 Conclusion

Overall CODHs are enzymes that have a very specific set of conditions to be able to function properly. Their use in industry will depend on scientists' ability to find and modify the enzyme. The optimization of these enzymes for industrial use will require specific conditions which need to be researched. The ability to prospect for novel CODH genes with unique functionalities from metagenomic data is a promising avenue for



further biotechnological development. While many hurdles still need to be overcome to achieve this goal, there are many possible applications CODHs could be useful in providing a biotechnological solution to issues that currently are unsolved.

While some CODHs may be useful for this from biotechnological industry; CODHs have a very important role in life's history. The connection to the Wood-Ljungdahl pathway makes this an important piece in the evolution of life on Earth. Characterizing novel organisms like *Hydrothermarcheota* JdFR-17 provides context into the history of life on Earth and can allow us to understand how life started on Earth. If enough organisms are characterized and sequenced, phylogenetic reconstruction of ancient organisms could be possible. This could lead to a greater understanding of the foundations of life which could be used to discover primitive life on other planets.[3]

## 5 Reference List

- [1] Feng J, Lindahl PA. Carbon Monoxide Dehydrogenase from *Rhodospirillum rubrum*: Effect of Redox Potential on Catalysis. *Biochemistry*. 2004;43(6):1552–1559. doi:10.1021/bi0357199
- [2] Adam PS, Borrel G, Gribaldo S. Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proceedings of the National Academy of Sciences of the United States of America*. 2018;115(6):E1166–E1173. doi:10.1073/pnas.1716667115
- [3] Cockell CS. Habitable worlds with no signs of life. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2014;372(2014):20130082. doi:10.1098/rsta.2013.0082
- [4] Techtmann SM, Lebedinsky AV, Colman AS, Sokolova TG, Woyke T, Goodwin L, Robb FT. Evidence for Horizontal Gene Transfer of Anaerobic Carbon Monoxide Dehydrogenases. *Frontiers in Microbiology*. 2012;3:132. doi:10.3389/fmicb.2012.00132
- [5] Robb FT, Techtmann SM. Life on the fringe: microbial adaptation to growth on carbon monoxide. *F1000Research*. 2018;7:F1000 Faculty Rev-1981. doi:10.12688/f1000research.16059.1
- [6] King GM, Weber CF. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nature Reviews Microbiology*. 2007;5(2):107–118. doi:10.1038/nrmicro1595
- [7] Can M, Armstrong FA, Ragsdale SW. Structure, Function, and Mechanism of the Nickel Metalloenzymes, CO Dehydrogenase, and Acetyl-CoA Synthase. *Chemical Reviews*. 2014;114(8):4149–4174. doi:10.1021/cr400461p
- [8] Cordero PRF, Bayly K, Man Leung P, Huang C, Islam ZF, Schittenhelm RB, King GM, Greening C. Atmospheric carbon monoxide oxidation is a widespread mechanism supporting microbial survival. *The ISME Journal*. 2019;13(11):2868–2881. doi:10.1038/s41396-019-0479-8
- [9] Voordouw G. Carbon Monoxide Cycling by *Desulfovibrio vulgaris* Hildenborough. *Journal of Bacteriology*. 2002;184(21):5903–5911. doi:10.1128/JB.184.21.5903-5911.2002
- [10] Rich JJ, King GM. Carbon monoxide consumption and production by wetland peats. *FEMS Microbiology Ecology*. 1999;28(3):215–224. doi:10.1111/j.1574-6941.1999.tb00577.x
- [11] Bae J, McCarty PL. Variation of Carbon Monoxide Production during Methane Fermentation of Glucose. *Water Environment Research*. 1993;65(7):890–898.

- [12] Jouny M, Hutchings GS, Jiao F. Carbon monoxide electroreduction as an emerging platform for carbon utilization. *Nature Catalysis*. 2019;2(12):1062–1070. doi:10.1038/s41929-019-0388-2
- [13] Nitopi S, Bertheussen E, Scott SB, Liu X, Engstfeld AK, Horch S, Seger B, Stephens IEL, Chan K, Hahn C, et al. Progress and Perspectives of Electrochemical CO<sub>2</sub> Reduction on Copper in Aqueous Electrolyte. *Chemical Reviews*. 2019;119(12):7610–7672. doi:10.1021/acs.chemrev.8b00705
- [14] Alfano M, Cavazza C. The biologically mediated water–gas shift reaction: structure, function and biosynthesis of monofunctional [NiFe]-carbon monoxide dehydrogenases. *Sustainable Energy & Fuels*. 2018;2(8):1653–1670. doi:10.1039/C8SE00085A
- [15] Daniell J, Köpke M, Simpson SD. Commercial Biomass Syngas Fermentation. *Energies*. 2012;5(12):5372–5417. doi:10.3390/en5125372
- [16] Ragsdale SW. Life with Carbon Monoxide. *Critical Reviews in Biochemistry and Molecular Biology*. 2004;39(3):165–195. doi:10.1080/10409230490496577
- [17] Goordial J, D'Angelo T, Labonté JM, Poulton NJ, Brown JM, Stepanauskas R, Früh-Green GL, Orcutt BN. Microbial Diversity and Function in Shallow Subsurface Sediment and Oceanic Lithosphere of the Atlantis Massif. *mBio*. 2021 Aug 3 [accessed 2022 Feb 22]. <https://journals.asm.org/doi/abs/10.1128/mBio.00490-21>. doi:10.1128/mBio.00490-21
- [18] Wu M, Ren Q, Durkin AS, Daugherty SC, Brinkac LM, Dodson RJ, Madupu R, Sullivan SA, Kolonay JF, Nelson WC, et al. Life in Hot Carbon Monoxide: The Complete Genome Sequence of *Carboxydotherrmus hydrogenoformans* Z-2901. *PLoS Genetics*. 2005;1(5):e65. doi:10.1371/journal.pgen.0010065
- [19] Kozubal MA, Romine M, Jennings R deM, Jay ZJ, Tringe SG, Rusch DB, Beam JP, McCue LA, Inskeep WP. Geoarchaeota: a new candidate phylum in the Archaea from high-temperature acidic iron mats in Yellowstone National Park. *The ISME Journal*. 2013;7(3):622–634. doi:10.1038/ismej.2012.132
- [20] Omae K, Oguro T, Inoue M, Fukuyama Y, Yoshida T, Sako Y. Diversity analysis of thermophilic hydrogenogenic carboxydrotrophs by carbon monoxide dehydrogenase amplicon sequencing using new primers. *Extremophiles*. 2021;25(1):61–76. doi:10.1007/s00792-020-01211-y
- [21] Merrouch M, Benvenuti M, Lorenzi M, Léger C, Fourmond V, Dementin S. Maturation of the [Ni–4Fe–4S] active site of carbon monoxide dehydrogenases. *Journal of Biological Inorganic Chemistry*. 2018;23(4):613–620. doi:10.1007/s00775-018-1541-0

- [22] Inoue M, Nakamoto I, Omae K, Oguro T, Ogata H, Yoshida T, Sako Y. Structural and Phylogenetic Diversity of Anaerobic Carbon-Monoxide Dehydrogenases. *Frontiers in Microbiology*. 2019;9:3353. doi:10.3389/fmicb.2018.03353
- [23] Fontecilla-Camps JC, Nicolet Y, editors. *Metalloproteins: methods and protocols*. New York: Humana Press; 2014. (Methods in molecular biology).
- [24] Carr SA, Jungbluth SP, Eloe-Fadrosh EA, Stepanauskas R, Woyke T, Rappé MS, Orcutt BN. Carboxydrotrophy potential of uncultivated Hydrothermarchaeota from the seafloor crustal biosphere. *The ISME Journal*. 2019;13(6):1457–1468. doi:10.1038/s41396-019-0352-9
- [25] Duan H, He P, Shao L, Lü F. Functional genome-centric view of the CO-driven anaerobic microbiome. *The ISME Journal*. 2021;15(10):2906–2919. doi:10.1038/s41396-021-00983-1
- [26] Baker BJ, De Anda V, Seitz KW, Dombrowski N, Santoro AE, Lloyd KG. Diversity, ecology and evolution of Archaea. *Nature Microbiology*. 2020;5(7):887–900. doi:10.1038/s41564-020-0715-z
- [27] Ensign SA. Reactivity of Carbon Monoxide Dehydrogenase from *Rhodospirillum rubrum* with Carbon Dioxide, Carbonyl Sulfide, and Carbon Disulfide. *Biochemistry*. 1995;34(16):5372–5381. doi:10.1021/bi00016a008
- [28] Hadj-Saïd J, Pandelia M-E, Léger C, Fourmond V, Dementin S. The Carbon Monoxide Dehydrogenase from *Desulfovibrio vulgaris*. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. 2015;1847(12):1574–1583. doi:10.1016/j.bbabi.2015.08.002
- [29] Inoue T, Yoshida T, Wada K, Daifuku T, Fukuyama K, Sako Y. A Simple, Large-Scale Overexpression Method of Deriving Carbon Monoxide Dehydrogenase II from Thermophilic Bacterium *Carboxydotherrmus hydrogenoformans*. *Bioscience, Biotechnology, and Biochemistry*. 2011;75(7):1392–1394. doi:10.1271/bbb.110159
- [30] Ainala SK, Seol E, Kim JR, Park S. *Citrobacter amalonaticus* Y19 for constitutive expression of carbon monoxide-dependent hydrogen-production machinery. *Biotechnology for Biofuels*. 2017;10(1):80. doi:10.1186/s13068-017-0770-8
- [31] Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L. Phylogenetically Novel Uncultured Microbial Cells Dominate Earth Microbiomes. *mSystems*. 2018;3(5):e00055-18. doi:10.1128/mSystems.00055-18
- [32] CooS1. [accessed 2022 Jul 7]. [https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=GeneDetail&page=geneDetail&gene\\_oid=2730025767](https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=GeneDetail&page=geneDetail&gene_oid=2730025767)
- [33] CooS2. [accessed 2022 Jul 7]. [https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=GeneDetail&page=geneDetail&gene\\_oid=2730027368](https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=GeneDetail&page=geneDetail&gene_oid=2730027368)

- [34] CooC1. [accessed 2022 Jul 7]. [https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=GeneDetail&page=geneDetail&gene\\_oid=2730025725](https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=GeneDetail&page=geneDetail&gene_oid=2730025725)
- [35] Corporation ZR. Zyp<sup>TM</sup> Plasmid Miniprep Kit. :12.
- [36] eLabProtocols - ZYM-5052 medium for auto-induction. [accessed 2022 Jun 24]. <https://www.elabprotocols.com/protocols/#!protocol=38033>
- [37] Lidstrom:Autoinduction Media - OpenWetWare. [accessed 2022 Jul 5]. [https://openwetware.org/wiki/Lidstrom:Autoinduction\\_Media](https://openwetware.org/wiki/Lidstrom:Autoinduction_Media)
- [38] Document Connect. [accessed 2022 Jun 30]. [https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FMSG%2Fmanuals%2FMAN0025601\\_Protein\\_Thermal\\_Shift\\_QR.pdf](https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FMSG%2Fmanuals%2FMAN0025601_Protein_Thermal_Shift_QR.pdf)
- [39] Document Connect. [accessed 2022 Jun 30]. [https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets/MSG/manuals/MAN0025600\\_Protein\\_Thermal\\_Shift\\_UG.pdf](https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets/MSG/manuals/MAN0025600_Protein_Thermal_Shift_UG.pdf)
- [40] Sundara Sekar B, Mohan Raj S, Seol E, Ainala SK, Lee J, Park S. Cloning and functional expression of *Citrobacter amalonaticus* Y19 carbon monoxide dehydrogenase in *Escherichia coli*. *International Journal of Hydrogen Energy*. 2014;39(28):15446–15454. doi:10.1016/j.ijhydene.2014.07.148
- [41] Jiang B, Jain A, Lu Y, Hoag SW. Probing Thermal Stability of Proteins with Temperature Scanning Viscometer. *Molecular Pharmaceutics*. 2019;16(8):3687–3693. doi:10.1021/acs.molpharmaceut.9b00598
- [42] Grahame DA, Stadtman TC. Carbon monoxide dehydrogenase from *Methanosarcina barkeri*. Disaggregation, purification, and physicochemical properties of the enzyme. *Journal of Biological Chemistry*. 1987;262(8):3706–3712. doi:10.1016/S0021-9258(18)61412-7

## **A Appendix**

### **A.1 Carbon Monoxide Constructs –**

[https://1drv.ms/x/s!ArhqiwxEzgvHgbtzw1\\_EUT6SQxicw?e=PcJYx0](https://1drv.ms/x/s!ArhqiwxEzgvHgbtzw1_EUT6SQxicw?e=PcJYx0)

### **A.2 Autoinduction Media –**

<https://www.elabprotocols.com/protocols/#!protocol=38033>

### **A.3 BCA Assay Raw Data –**

<https://1drv.ms/x/s!ArhqiwxEzgvHgb8LFUtAtntIJJuSEQ?e=TiiTxn>

### **A.4 BCA Assay Protocol –**

<https://1drv.ms/w/s!ArhqiwxEzgvHgb8JjShz-YIT75fOoQ?e=M6wajI>

### **A.5 Melt Curve Assay PCR Settings –**

[https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2FMAN0025601\\_Protein\\_Thermal\\_Shift\\_QR.pdf](https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2FMAN0025601_Protein_Thermal_Shift_QR.pdf)

### **A.6 Melt Curve Assay Setup –**

[https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0025600\\_Protein\\_Thermal\\_Shift\\_UG.pdf](https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0025600_Protein_Thermal_Shift_UG.pdf)

### **A.7 GC Assay Standard Curve –**

<https://1drv.ms/x/s!ArhqiwxEzgvHgb8QQ2aRZV1IT2HRAQ?e=qebl3c>

### **A.8 Forward Reaction Example –**

<https://1drv.ms/u/s!ArhqiwxEzgvHgb8gKISJbwIIIgqcAA?e=ldVCBD>

### **A.9 Reverse Reaction Example –**

<https://1drv.ms/u/s!ArhqiwxEzgvHgb8f3GpkaqGM-v7vSQ?e=cF2TY8>

### **A.10 Raw Data for the GC Assays –**

[https://1drv.ms/x/s!ArhqiwxEzgvHgb8swsV3-eEfW\\_qONag?e=NKvyjh](https://1drv.ms/x/s!ArhqiwxEzgvHgb8swsV3-eEfW_qONag?e=NKvyjh)

## **B Copyright documentation**

### **B.1 Figure 1**

Figure 1: “Main natural biospheric sources and sinks for carbon monoxide” by King and Weber in Nature Reviews Microbiology -

<https://www.nature.com/articles/nrmicro1595/figures/1> Accessed July 2022

Copyright use approval license -

<https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=762f68ff-ea8e-4352-821c-412b3b79c0c7>

### **B.2 Figure 2**

Figure 2: “Carbon monoxide (CO) metabolisms and the Wood–Ljungdahl pathway.” By Robb and Techtmann in F1000 Research.

Licensed under CC BY-SA 3.0 - <https://f1000research.com/articles/7-1981/v1>

### **B.3 Figure 3**

Figure 3: “GENE TOPOLOGY FOR THE HYDROGENASE – CODH GENE CLUSTERS FROM C. HYDROGENOFORMANS, T. CARBOXYDIVORANS, AND R. RUBRUM” by Techtmann et. al. in Frontiers in Microbiology.

Licensed under CC BY-SA 3.0 - [https://www.frontiersin.org/files/Articles/20477/fmicb-03-00132-r2/image\\_m/fmicb-03-00132-g005.jpg](https://www.frontiersin.org/files/Articles/20477/fmicb-03-00132-r2/image_m/fmicb-03-00132-g005.jpg) Accessed July 2022

### **B.4 Figure 4**

Figure 4: “Phylogenetic relationships of Ni-CODHs” by Inoue et. al. in Frontiers in Microbiology.

Licensed under CC BY-SA 3.0 - [https://www.frontiersin.org/files/Articles/429207/fmicb-09-03353-HTML/image\\_m/fmicb-09-03353-g001.jpg](https://www.frontiersin.org/files/Articles/429207/fmicb-09-03353-HTML/image_m/fmicb-09-03353-g001.jpg) Accessed July 2022