Delft University of Technology

Universitat Politècnica de València

20/08/2022

Adaptation of the thermoalkaliphilic bacteria Caldalkalibacillus Thermarum TA2.A1 to different oxygen environments

Bachelor Thesis Life Science and Technology
Biocatalysis department

Trabajo de Fin de Grado realizado en la Escuela Tecnológica Superior de Ingeniería Agronómica y del Medio Natural

Supervisor: Dr Duncan McMillan
Cotutor: Ana Belén Heredia Gutierrez

Andrea Castelló Tejera

Title: Adaptation of the thermoalkaliphilic bacteria *Caldalkalibacillus thermarum* TA2.A1 to different oxygen environments.

Abstract

The thermoalkaliphilic bacterium *Caldalkalibacillus thermarum* TA2.A1, who belongs to the Bacillaceae family, grows simultaneously at alkaline environments, high temperatures and low oxygen concentrations. The microorganism needs oxygen strictly to obtain energy by the oxidative phosphorylation chain but too much oxygen produces harmful effects. However, this microorganism lives in environments with changing oxygen levels, therefore the proteome composition will be studied in order to describe the mechanisms used by the bacteria to adapt to them. Three different dissolved oxygen availabilities in the media culture will be studied: 1.05%, 10.5% and 21% and the data processing will be done by a series of python scripts designed for this proteome data set. The results show an optimum growth at 10.5% DOTs. At lower DOTs fermentation occurs and an alternative oxidative phosphorylation chain is activated in order to guarantee survival when very little oxygen is available. Lastly at 21% oxygen, when oxidative stress occurs, many antioxidant enzymes are overexpressed in order to control the potential harm, and cell division is delayed by many different mechanisms such as sporulation, mechanisms like quorum sensing and cofactors with lower redox potential so resources are not wasted. This study gives a deeper insight into the adaptation of the microorganism to the diverse oxygen concentrations.

Key words: Proteomics, microaerophile, metabolic switches, proteins, aerobic respiration, oxidative phosphorylation chain, carbon metabolism, cell division.

Author: Andrea Castelló Tejera

Location and date: Delft, august 2022

Academic tutor: Dr. Duncan McMillan

Cotutor from home university: Ana Belén Heredia Gutierrez

Título: Adaptación de la bacteria termoalcalófila *Caldalkalibacillus thermarum* TA2.A1 a diferentes ambientes de oxígeno.

Resumen

La bacteria termoalcalófila *Caldalkalibacillus thermarum* TA2.A1, perteneciente a la familia Bacillaceae, crece simultáneamente en medios alcalinos, a altas temperaturas y concentraciones bajas de oxígeno. El microorganismo necesita de manera esencial oxígeno para obtener energía vía la cadena de respiración oxidativa, pero si este se encuentra en concentraciones demasiado elevadas, el oxígeno producirá efectos dañinos para el mismo. Así mismo, el microorganismo vive en condiciones con niveles de oxígeno cambiantes, por lo que la composición del proteoma será estudiada para ser capaces de describir los mecanismos usados por la bacteria para adaptarse a las diferentes condiciones. Tres diferentes disponibilidades de oxígeno disuelto en el medio de cultivo serán utilizadas para cultivar el microorganismo: 1.05%, 10.5% and 21%. Después de realizar un análisis de masas a las muestras obtenidas, los datos obtenidos serán analizados a través de una serie de scripts creados especialmente para este experimento. Los resultados muestran un crecimiento óptimo en las condiciones de oxígeno de 10.5 DOTs. Cuando los niveles de oxígeno disminuyen, los mecanismos de fermentación son activados y una alternativa cadena de fosforilación oxidativa es expresada en la membrana. Gracias a estos dos mecanismos, el microorganismo es capaz de sobrevivir cuando muy pocas moléculas de oxígeno están disponibles. Por otro lado, cuando los niveles de oxígeno en el medio de cultivo aumentan, ocurre el estrés oxidativo, entonces bastantes enzimas antioxidantes son expresadas para controlar el potencial daño que podrían causar los radicales libres. La división celular se ve relentizada debido a varios sucesos como la esporulación, señalización por *quorum sensing* o el uso de cofactores con bajos potenciales redox. Evitando la división celular descontrolada se consigue un mayor control de los recursos disponibles. Este estudio da una visión más profunda a la adaptación del microorganismo *Caldalkalibacillus thermarum* a las diferentes concentraciones de oxígeno.

Palabras clave: Proteómica, microaerófilo, cambio metabólico, proteínas, respiración aerobia, cadena de foforilación oxidativa, metabolismo del carbono, división celular.

Autor: Andrea Castelló Tejera

Localización y fecha: Delft, agosto de 2022.

Tutor académico: Dr. Duncan McMillan

Cotutor: Ana Belén Heredia Gutierrez

Acknowledgments

I would like to express my gratitude to everyone I worked with during my research project. First, to Gabriella van Leersum for being always very kind with me and creating a very good environment. I want to also thank Sam de Jong for his daily supervision and all his help and guidance through my project. And lastly, I want to thank Dr. Duncan McMillan for offering me the opportunity to work in this department and make my research project possible. He encouraged me to keep going and to enjoy as much as possible from the project.

I want to also thank to my family that has been supporting me since the beginning. Thanks to my brother for being my reference in all aspects. To believe I can do whatever I intend to and to help me without any complaint with everything he could. I want to thank my mother for believing in me unconditionally and appreciate my job. Lastly my father for being my everyday support and my personal counsellor. Also, I want to thank all the friends I made during my stance in the Netherlands for have shared with me all the time on the library and make my weekends so joyful that I had enough strength to survive the rest of the week. Also, I am thankful that all I have learned these past months hasn't only been about *C. thermarum* but also about other nationalities and cultures.

Contents

1. Introduction 1
1.1 Extremophile microorganisms 1
1.1.1 Alkaliphiles 1
1.1.2 Thermophiles 2
1.1.3 Microaerophiles 3
1.2 *Caldalkalibacillus thermarum* TA2.A1 5
1.2.1 Central Carbon Metabolism 7
1.2.2 Oxidative phosphorylation. 8
1.3 Objectives and approaches 8
2. Materials and methods 9
2.2 Samples and growth conditions 10
2.2 Mass spectrometry 10
2.3 Proteomic data analysis 11
2.3.1 Introduction 11
2.3.2 Pathway analysis 13
3. Results and Discussion 13
3.1 Pathway analysis 13
3.1.1 Stacked bar plot 14
3.1.2 Clustermap 16
3.2 Overexpressed proteins 18
3.2.1 1.05% Oxygen Inflow 19
3.2.2 10.5% Oxygen Inflow 28
3.2.3 21% Oxygen Inflow 26
3.3 Membrane proteins characterization 29
3.3.1 Complex I 30
3.3.2 Complex II 31
3.3.3 Quinones 32
3.3.4 Complex III 32
3.3.5 Cytochrome C 33
3.3.6 Complexes IV. Terminal oxidases. 34

3.3.7 Complex V. ATP Synthase. 34
4. Conclusions 35
5. Bibliography 37

1. Introduction

1.1 Extremophile microorganisms

There is limited information available about microorganisms living in extreme conditions because of the difficulty to study them in a common facility but in fact, they are really interesting due to the specific properties that provide them the ability to grow at the limits of the environmental factors such as pH, salinity, temperature and pressure. Studying them in detail, will give us insights into the mechanisms they use to reach environmental adaptation.

Understanding their adaptation opens a lot of doors for protein discovery and usage in industry and many other fields that require extreme conditions. Also, genetic engineering could be done to create bacterial strains resistant to the same environments.

1.1.1 Alkaliphiles

An environment is considered alkaline when the pH is higher than 9 and occasionally, it can reach values higher than 12. These environments are not very common in nature - they are found in soda lakes, soda deserts, hot springs, or horse manures (Tiago *et al.*, 2004). Alkaliphilic microorganisms can be classified into obligate or facultative alkaliphiles depending on the range of pH the bacteria can tolerate. Obligate alkaliphiles require strictly a pH greater than 9 while facultative can live in a broader range of pHs (Wissink, 2021) The alkaline environment can also be defined as proton dessert due to the low availability of hydrogen ions.

Bacteria living at this pH must overcome several thermodynamic challenges in terms of chemiosmotic energy generation and solute transport driven by the proton motive force. They show an inverted gradient in which their intracellular pH is lower than the extracellular pH (Peddie *et al.*, 1999). Microorganisms living in these conditions have been evolving during centuries to adapt their mechanisms to guarantee survival. One of the most recurrent mechanisms alkaliphilic bacteria appeal is the use of Na^+ ions instead of H^+ to lead the electron transport chain (ETC) of the oxidative respiration. Another interesting mechanism could be a

H^+ sequestration, where protons are trapped by the cell wall or the cell membrane in the intermembrane space to maintain the osmotic gradient. Alkaliphiles use sodium/proton antiporters to acidify the cytosol and generate an inwardly directed sodium motive force. Alkaliphilic bacteria are studied because the ability to use the scarcely available protons in their habitat and with an inverted chemical potential gradient to harvest energy via respiration of oxygen is poorly understood and quite intriguing. (Wissink, 2021)

Furthermore, enzymes extracted from these microorganisms are important in the detergent industry. Alkaline enzymes extracted from these bacteria such as alpha- amylase, cellulase (endo-1,4-ß-glucanase), mannanase and lipase have been incorporated into heavy-duty laundry and dishwashing detergents since the 1970's. Moreover, nowadays the largest world market for industrial enzymes is the detergent industry. Detergent enzymes sum up the 30-40% of the total worldwide enzyme production except for diagnostic and therapeutic enzymes, so alkaliphilic bacteria are a good target of study (Ito, 2011).

1.1.2 Thermophiles

Thermophiles are defined as microorganisms that grow optimally in temperatures between 55 and 80ºC. Hot springs are one of the main sites where thermophiles are isolated from, although they can also thrive in man-made environments, such as compost facilities. Usually, thermophiles may also live under other harsh conditions like extreme pH or high salt concentrations (Urbieta *et al*., 2015).

Extreme temperatures, as every other extreme condition, alter the normal properties of the microorganisms. High temperatures induce a decrease in stability. This destabilization can be seen as the denaturalization of the native conformation of proteins and a leaky cytoplasmic membrane to protons (Peddie *et al*., 1999; Smith *et al*., 2000). Thermophiles are microorganism that are capable of surviving and growing optimally on those high temperatures by providing mechanisms that increase stability. First, the cytoplasmic membrane is shown to be reinforced by an increase in acyl-chain length, lipid saturation, branching and cyclization of the fatty acids upon increasing the temperature. All these features make the membrane more stable and decrease the proton permeability of the cytoplasmic membrane, important for growth at high temperatures (Tolner *et al*., 1997). Macromolecules,

such as proteins and enzymes, also need to reinforce their stability in order to preserve their native state. There are many hypotheses that explain the mechanism they use to achieve this stability, but the most important ones are based on an increase in the number of disulphide bonds in the quaternary structure and the synthesis of heat shock proteins (HSP) which protect protein complexes from being dissolved (Wang *et al*., 2013).

Thermophilic microorganisms have attracted much interest in the sector of biotechnology, especially in relation to industrial processes because of the increased stability their proteins present to high temperatures. They can be used in many fields such as bioenergy, biomining, and enzyme and biosurfactant production but the two main studied fields are bioremediation strategies and clean production technologies (Urbieta *et al*., 2015).

1.1.3 Microaerophiles

Microaerophiles are microorganisms for which oxygen is beneficial when it is available at low concentrations but harmful at the same time at high concentrations. They grow best under 2-10% oxygen availability, and they can't grow under fully oxic conditions (21%), or they can grow poorly. In figure 1 it is represented the growth rate against the dissolved oxygen tension (DOT) for different types of respiration.

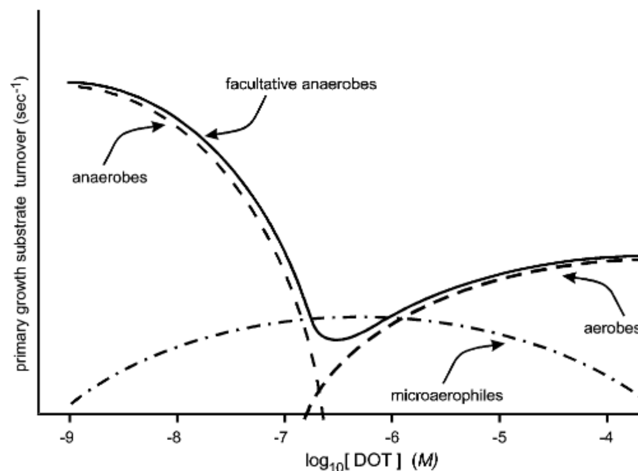

Figure 1*.* The different types of energy metabolism depending on the oxygen utilization. Anaerobes grow in absence of oxygen, aerobes need strictly oxygen, facultative anaerobes grow best under anaerobic conditions but can also respire in presence of oxygen and microaerophiles grow best under low oxygen concentrations.

Microaerophiles can be classified in three different groups: the obligate microaerophiles that only grow in a determined low oxygen concentration; the microaerophiles that grow best under low oxygen concentrations but can grow aerobically if certain conditions are met; and lastly, the facultative microaerophiles that have anaerobic phenotype and metabolism but can grow at little oxygen concentrations (Kaakoush *et al*., 2007).

Some examples of microaerophile microorganisms are aquatic microorganisms, such as *Gallionella ferruginea*, found in the depths of the ocean where the levels of dissolved oxygen in the water (DOT) are not very elevated; Nitrogen fixers, like *Rhizobium*, which are microorganisms that stablish a symbiotic relationship with plants. These bacteria incite the production of root nodules on legumes, making possible the nitrogen fixation. Another example are the pathogens, these group is the most studied because the importance they have on actual diseases rather in humans than in other animals.

Campylobacter jejuni is considered an important pathogen that affects humans, for this reason, many microbiologists have been concerned and it has been deeply studied during the past years. Infection of *C. jejuni* can cause symptoms such as fever, headache, dizziness and myalgia but the most important is gastroenteritis or diarrhoea. *C. jejuni* is an obligate aerobe, doesn't survive to oxygen concentration greater than 10% and its metabolism is strictly respiratory, only oxygen is used as a terminal acceptor. There is no evidence this microorganism is capable of do fermentation. (Smith *et al*, 2000)

Helicobacter pylori another important pathogen in humans that has been deeply studied. It is a causative agent of gastric and duodenal ulcers in humans and the persistence of this bacteria in the stomach increases the risk of developing gastric cancer. The oxygen concentrations in which grows best are 4-5% (Donelli *et al.,* 1998).

The main characteristic all microaerophiles have in common is their respiratory chain: They have specific terminal oxidases with a very high affinity for oxygen. Hence, this mechanism is believed to be the main responsible to enable the utilization of the little available oxygen. However, there must be more reasons why microaerophiles grow better under low oxygen conditions (Smith *et al*, 2000).

One extended theory is the increased susceptibility of these microorganisms to the toxic reactive oxygen species or ROS. It is demonstrated microaerophiles like *C. jejuni* can tolerate less hydrogen peroxide concentrations than *E. coli*, this could be due to an increased

generation of ROS in their particular metabolism or because they lack the necessary enzymes such as catalase or superoxide dismutase to neutralize these compounds, decreasing its aerotolerance (Krieg and Hoffman, 1986). Another renowned theory is that there is a strong inhibition of certain respiratory enzymes by oxygen. All these mechanisms make it impossible for these bacteria to grow under fully oxic conditions. (Kaakoush *et al*., 2007)

Most of the microaerophiles possess pathways characteristic of anaerobic metabolism and evidence shows they are capable of using them and carry out fermentation metabolism. Therefore, why can't they grow under anaerobic conditions with lack of oxygen if they possess the mechanisms to do so? It is hypothesized that oxygen is then required for other survival processes of the microorganism such as biosynthesis of compounds.

There is probably more than one explanation for microaerophily that applies to all organisms. Even within the same genus, there is a wide variation of factors.

1.2 *Caldalkalibacillus thermarum* TA2.A1

The aerobic thermoalkaliphilic bacteria *Caldalkalibacillus thermarum* TA2.A1 is selected to be deeply studied because of its many interesting features**.** This microorganism is classified as gram positive and belongs to the Bacillaceae family. It was found in a thermal spring in Mount Te Aroha, New Zealand. Therefore, it is capable of living under high temperatures. The age of the microorganism can't be determined, but none other close relative has been found suggesting that the evolutionary line diverged long time ago. The optimal growth conditions for *C. thermarum* TA2.A1 are temperatures of 65ºC and a pH of 9.5. (De Jong, 2020). Hence, the bacteria is capable of grow under extreme pH and temperature but the most important fact doesn't rely in this. But the adaptability It possesses under different environments. It is capable of change its internal mechanisms to adapt to different environmental conditions such as oxygen availability, pH and temperature.

One of the most interesting features of *C. thermarum* TA2.A1 is the alkaliphilic profile, that has been previously described. *C. thermarum* grows under high pH conditions or lack of protons. The mechanism that most alkaliphiles use to overcome the thermodynamic challenges that it involves are the use of $Na+$ ions to create an ionic gradient (Smith *et al.,* 2000). However, there

is no evidence that any Na^+ pumping transport chain enzyme has been found on *C. thermarum,* so the microorganism still uses hydrogen to get energy from oxidative respiration and proton pumping so it needs other mechanisms to beat the thermodynamic challenges (Wissink, 2021).

The growth of the bacteria at different oxygen conditions was trying to be described by studying the optical density and the cell dry weight of the cell culture in Martijn Wissink thesis (2021). Then, the overall bacterial population density was calculated by using these measures. Being able to hypothesize the optimal growth conditions.

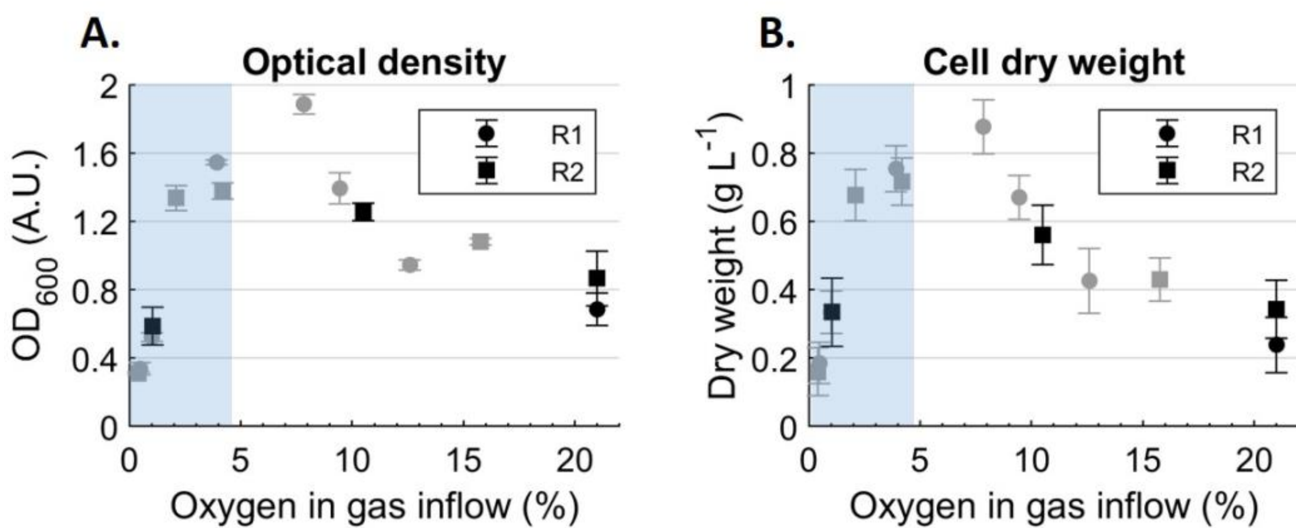

Figure 2. Figure from Martijn Wissink Thesis (2021). The effect of oxygen on the growth of the bacteria *C. thermarum* is represented. The dot shapes, represent the two different reactor, where the grey data-points represent the average of four technical replicates and the black data-points represent biological duplicates, with again four technical replicates. **A.** Optical denisity (OD at 600 nm) **B.** Cell dry weight.

It can be seen in the results that the highest cell density is shown at around 7% oxygen, but this exact number can't be stated because there is still some limiting compound that hasn't been defined yet. We can still see that when oxygen become limiting (1%), the population density decreases significantly. Same behaviour shown at atmosphere normal conditions (21%).

Hence, Oxygen is strictly required by the microorganism to grow, so *C. thermarum* is classified as an obligate aerobe although it is able to do fermentation with low oxygen concentrations, suggesting that oxygen is not necessary for respiration but for other survival processes others than energy generation. It is classified as the second type of microaerophile mentioned in the

last section, the microaerophiles that grow best under low oxygen concentrations but can grow aerobically if certain conditions are met.

Another interesting aspect about this alkaliphilic bacteria is that *C. thermarum,* even though grows better at pH 9.5, the actual range for growth using fermentable carbon sources is from pH 7.5 to 10. These fermentable carbon sources include sucrose, glutamate, trehalose, pyruvate, and common C_4 dicarboxylates. Whereas for nonfermentable carbon sources like malate or succinate a pH greater than 9 is required (McMillan *et al*., 2009). Because of this feature, *C. thermarum* is classified as a facultative alkaliphilic microorganism, depending on the carbon source used. The reason why this bacterium is unable to grow in pH lower than 9 are not known (McMillan *et al*., 2009).

1.2.1 Central Carbon Metabolism

The carbon sources the microorganism can assimilate are classified in fermentable and nonfermentable. The two preferred carbon sources are glutamate and sucrose.

Sucrose is imported via a phosphotransferase system, and it is split up into fructose and glucose once reaches the cytosol. Glucose will follow the classic glycolytic and TCA cycle path while Glutamate assimilation will be more intriguing due to it is believed to follow a path that no other microorganism has followed before. It is hypothesized that glutamate enters directly to the TCA cycle by the enzyme glutamate dehydrogenase. Moreover, *C. thermarum* shows to grow a 30% faster when growing on glutamate comparing to sucrose. This could be related to a physiologically beneficial function. Such as the 'resource allocation theory' where fewer enzymes are implicated in the glutamate digestion rather than on sucrose. Another theory relies on the ammonia released in the glutamate deamination into α-ketoglutarate by glutamate dehydrogenase, which has a pK_a of 9.25. In a natural environment, It does not matter as much as in a sealed volume, because metabolic by-products are not washed away (De Jong *et al*., 2020).

The two main products found in the cell result of the substrate intake (sucrose or glutamate) are CO_2 and acetate. CO_2 is released from the oxidative phosphorylation system and acetate comes directly from the alcoholic fermentation. CO_2 is mainly produced with higher oxygen

concentrations (21%) but on the contrary, acetate takes over when oxygen concentrations decrease. Surprisingly, all the carbon stoichiometry still doesn't sum up the numbers, suggesting that there is another compound that is formed in the cell. However, this compound is still undefined. Moreover, the uncertainty on the culture media components makes forming a balance even harder.

1.2.2 Oxidative phosphorylation.

There are previous studies about the respiratory enzymes present in the cell depending on the O2 conditions, showing that there are multiple respiration pathways the microorganism can use. The evolutionary reasons for conserving multiple respiratory chain branches is thought to enable growth in environments with changing oxygen concentrations. Three different oxygen concentrations are defined in this study: low: 1.05%, medium: 10.5% and high: 21,0%. Wissink (2021) found in his growth experiment, that oxygen concentrations of about 7.0% where when the biomass concentration is highest meaning the conditions where the most optimum and cells divided as much as possible.

A brief review of how the respiratory chain is interesting to understand the bacteria's adaptation to oxygen. The respiratory machinery consists of five membrane protein complexes. Electrons enter the electron transport chain by pairs via complex I or complex II and are eventually shuttled to complex III by menaquinones (MQ) that once are reduced, are called menaquinol. Electrons are then split up by the complex III and will be transported to the terminal oxidase or complex IV by cytochromes. Four types of cytochromes have been found in *C. thermarum* expressed in different conditions (aa_3, ba_3, bb_3, *bd*). Lastly, electrons are donated to oxygen that works as an electron acceptor. Complex V or ATP Synthase will use all the Hydrogens pumped to the intermembrane space by some of the complexes to synthesize ATP and produce energy for the cell.

1.3 Objectives and approaches

The mechanisms microaerophiles use to grow at low oxygen availability are still not completely understood. From the proteomic data obtained in a previous respiratory chain study of *C.*

thermarum, the goal is to gain insight in the metabolic switches that take place in the central carbon metabolism, the fermentation pathways and the overall processes inside the bacteria in response to changes in oxygen availability.

For this purpose, the changing expression levels of peptides will be calculated, and the proteins where the peptides belong will be characterized to see in what cell processes, they are implicated.

2. Materials and methods

2.1 Samples and growth conditions

The strain *Caldalkalibacillus thermarum* TA2.A1 was cultured by the master student Martijn Wissink in TU Delft, the Netherlands in 2021. Briefly, 250 mL shake flasks were used to grow the bacteria with an alkaline basal medium of pH 9.5 and a temperature of 65ºC. The carbon source added was either glutamate or sucrose. Then, this culture was used to inoculate the bioreactors. Two bioreactors of 3L each, with a working volume of 1L, were used to grow the bacteria under changing total dissolved oxygen or DOTs. They were constantly stirred and maintained at 65ºC by pumping hot water. The temperature was maintained by an embedded thermometer that was used to give feedback.

The amount of the dissolved oxygen (DO) was measured using DO sensors. First, the dissolved oxygen values were calibrated, and then, the volumetric mass transfer coefficient (K_La) values were defined in both reactors. The reactors were supposed to be in the same conditions, to obtain two identical replicates, but the K_LA values were unexpectedly different, and because they affect the oxygen tension in the reactor, results will end up being slightly different affecting relevant statistical analysis. Different DO values were applied to the bioreactors during time: Low 1.05%, medium 10.5% and high 21%. And a sample was taken from each DO value from both reactors. Then, samples were tagged using the TMTplex™ kit by Thermo Fischer (explained in the next section) for subsequently multiplexing in the mass spectrometry analysis.

After samples being tagged, the cells were processed. Cell membranes were extracted leaving the cytosol in the sample, so each sample was divided into membrane proteins, and whole cell proteins. Therefore, for membrane proteins 3 different conditions will be studied in the two different reactors so 6 samples will be studied, same number of samples for the whole cell proteins.

2.2 Mass spectrometry

Samples needed to be identified to be analysed at the same time, by multiplexing, in the mass spectrometer. Only this way, samples could be accurately compared. The tag system used was the TMTplex^{TM} kit by Thermo Fischer. This system consists in TMT tags, each composed of an amine-reactive N-Hydroxisuccinimide (NHS)-ester group, a spacer arm and a mass reporter. All tags are isobaric, meaning that the overall molecular weight is the same, but when a fragmentation occurs in the spacer arm, every tag results on a different molecular weight permitting sample differentiation. TMTplex^{TM} was used with 8 different tags. A different tag for every oxygen condition and every reactor. Tags were added before the sample separation between membrane and whole cell proteins.

Table 1. Characterization of the different samples in the right column and the tags added to them at the left.

Label	Condition
TMT10-126C	1.05% O_2. Reactor 1
TMT10-127N	1.05% O_2. Reactor 2
TMT10-127C	10.5% O_2. Reactor 1
TMT10-128N	10.5% O_2. Reactor 2
TMT10-128C	21% O_2. Reactor 1
TMT10-129N	21% O_2. Reactor 2
TMT10-129C	Membrane bridge (100% intensity)
TMT10-130N	Membrane bridge (100% intensity)

When all tags were added, all samples were sent to the Microbial Proteomics Group from TU Delft to introduce the samples in the mass spectrometer (Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™, ThermoFisher Scientific). Membrane proteins were analysed separately from the

whole cell proteins samples. Therefore, two assays were performed, with the samples described in table 1.

The data output obtained from the mass spectrometer analysis consisted in an excel file with many columns. The first one was the gene tag of the peptide. These were related by comparing the sequences of the MS/MS analysis with the sequences found in the Uniprot database. Then, it could be found the intensities from each peak, representing the proteins in every condition, the mean between the intensity of the two replicates and the ratio between the samples in between different oxygen concentrations. It was set the default ratio value 1 to the membrane proteins sample at 21% of oxygen. Then, a membrane bridge was set in the whole proteome sample so ratios from both samples could be compared.

2.3 Proteomic data analysis

2.3.1 Introduction

From the data output of the mass spectrometer, all analysis are made. All the changes the microorganism experiences are expected to be identified and an overview analysis of what pathways have changed will give us an insight of how the bacteria is adapting to changing oxygen environments.

Before all statistical tests and data analysis. Each protein had to be paired with its description, the protein unit where it belongs, the principal modules where it participates and the pathways where it is involved. For that reason, the kegg.py script was created using the Anaconda program. The intention was to correlate the Uniprot accession numbers to the KEGG IDs and retrieve all the information about the protein from the KEGG database and import it into our data files.

For all data relation. First, Uniprot accession needed to be related with the KEGG ID accession. By using the KEGG conversion commands in the Bio library from python, was impossible to obtain the relation. Then, the KEGG mapper conversion tool was loaded with all the Uniprot accessions, and no answer was either obtained. It was then found, that Uniprot database did not use the same Reference genome as the KEGG database, therefore the annotations were different, and no relation could be done. Then, from the mass spectrometry department in TU Delft, the data output was provided again, now with NCBI identifiers. By using the same script

in python, we were able to fix this problem and then get the necessary dictionaries to relate the data assigning to each peptide its related Unit, Modules and Pathways as explained before.

Subsequently, the peptides identified in the mass spectrometer were first classified depending on the tendency they showed on their expression levels between the different oxygen conditions. This classification was made by the script ratios.py and the peptides could be classified as showing: 'Greater protein expression in limiting oxygen conditions', meaning that the greatest expression was shown when samples were cultivated in 10.5% O_2; 'Greater protein expression with decreasing oxygen conditions', meaning greatest expression with lowest oxygen concentration (1.05%); 'Greater protein expression with increasing oxygen conditions', when greatest expression was shown at 21% O_2 and 'non-significant', expressing a change in their expression levels of more than a 30% between samples. This threshold was established based on similar proteomic analysis. By applying this script, we could classify all proteins by similar trend under different oxygen conditions.

When the output data was obtained, even though proteins could be classified by trends, a more reliable statistical analysis was required in order to proceed with the data analysis. Then, by the statistical.py script, a one-way anova test was performed to every protein in order to detect disparities between conditions. The test was made using the statistical library spicy.stats from python. The principal hypothesis was that all samples were considered similar. The hypothesis was rejected when samples showed significant discrepancies. Then, the statistical p-value could be used to sort the proteins that expressed more significant changes. As low as p-value is, less probability that the differences between the samples are due to random sampling. Therefore, more certainty to say the values are significantly different.

An additional output file was created with the proteins that showed low p-values ($p < 0.05$) so they could be manually annotated. The proteins with less p-values were displayed in tables 2,3 and 4 to notice the proteins that showed a more significant different expression.

The membrane proteins were selected to be represented in the three different conditions because of the importance of the ETC to the bacteria environmental adaptation. Using matplotlib to crease a bar plot of every protein unit.

2.3.2 Pathway analysis

To study not only the single protein changes but the overall pathway swifts. An overall pathway analysis was performed by using the pathway_analysis.py script in python. This script creates two visual graphs and gives an overall picture of what is happening in the cell, when the DOT changes.

The first graph is a stacked bar plot in which the overexpressed pathways of each condition could be compared by the counting the number of overexpressed proteins that participated in the respective pathway. To do so, of all proteins, only the proteins with a p-value less than 0.20 were selected to continue with the analysis. This selection resulted in 317 proteins out of 776, having enough data to study the most changing pathways. The pathways were assigned to the greatest expression condition, and then each pathway was represented by the number of proteins that participated in that specific pathway. Then, the pathways that contained more proteins implicated were selected for doing the graph. The 'not defined' pathways were discarded. The output is represented in table A.1 (ANNEX) Then, the data was plotted in a stacked plot to have a visual representation in every condition in figure 3.

The second visual graph created is a cluster map of all the overexpressed pathways using also the same script pathway_analysis.py. Only the proteins with the p-value less than 0.20 were as well selected to perform this graph. This special graph was selected because it enables to put together the pathways that show similar trends. To do so, the ratios were first normalized by dividing each value between the sum of the replicates in each condition. Then the normalized data could be compared in for each oxygen condition. The normalized data range varied between 0.2 meaning low expression and 0.5 meaning high expression.

3. Results and Discussion

3.1 Pathway analysis

The two visual graphs explained in the previous section (2.3.2) are shown for an overall view of the pathways that are differentially expressed and participate in the bacterial adaptation to the changing oxygen concentration levels in the environment.

3.1.1 Stacked bar plot reveals overexpressed pathways for each condition

The first graph created is the stacked bar plot. The Y axis represents the number of proteins that are involved in the most expressed pathways and the X axis represents every oxygen condition. This graph was created using the pathway_analysis.py script described in methods, where the top 20th pathways overexpressed where represented in the stacked bar plot for each condition, so they could be compared.

It is important to mention that the absence of one determinate pathway in a condition doesn't mean the non-appearance. Only the overexpressed pathways are represented on this graph, so no presence means proteins involved in that pathway can be present on the cell they are just not overexpressed.

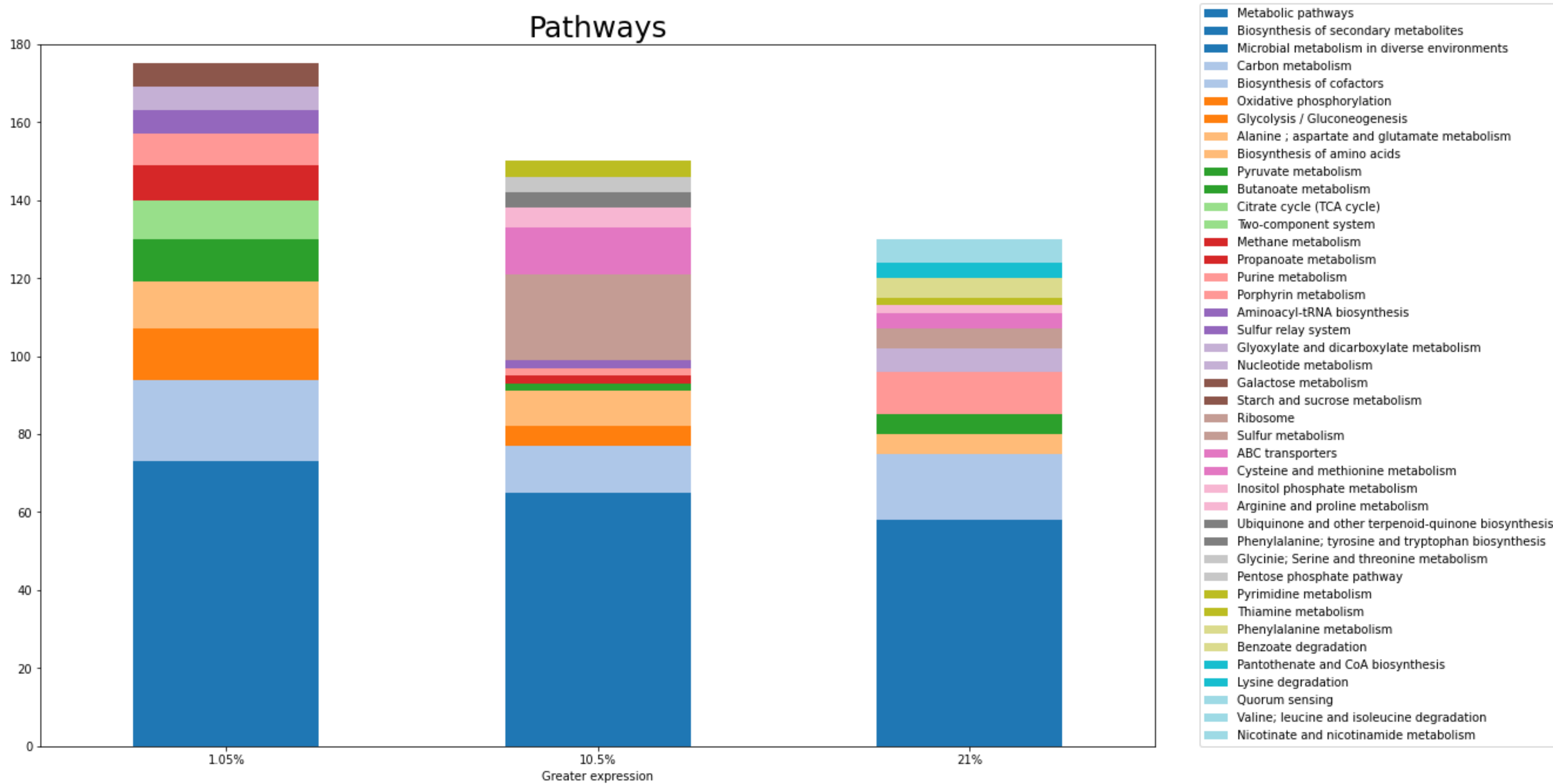

Figure 3. Comparison of the most differentially expressed pathways in *C. thermarum* in the different oxygen conditions. The Y-axis represents the number of proteins, and the three vertical bars represent the three oxygen conditions. Every color comprehends one or two pathways, and the size of the fragment is directly correlated with the overexpression in each condition.

In total 36 pathways are represented in the graph. The higher number of proteins expressed are involved in the metabolic pathways. In every condition, different metabolic pathways take place and the bacterial internal functioning changes. Other pathways that show an important difference in the expression levels are the biosynthesis of secondary metabolites, the microbial metabolism in diverse environments, the biosynthesis of cofactors, the carbon metabolism, the oxidative phosphorylation, and the ribosomes.

At 1.05% oxygen availability, the pathways related with the carbon metabolism such as methane biosynthesis, the citrate cycle, glycolysis, pyruvate metabolism, oxidative phosphorylation and butanoate metabolism are overexpressed. A significant change on how the carbon is processed looks like it is taking place. The carbon metabolism will be studied in more detail when the individual proteins are described (Section 3.2.1).

Also, the signalling pathways belonging to the two-component system appeared to increase the expression levels, this is the system the microorganisms use for detecting the oxygen concentrations in the environment and will also be discussed in the next section (3.2.1).

An unexpected pathway is found at this condition: The methane metabolism. The production of methane is a feature of anaerobic bacteria, and the bacteria that possess this characteristic is known as methanogenic bacteria. On the other hand, the bacteria that use methane to obtain energy are called methanotrophs. As far as it is known, *C. thermarum* hasn't any of these features. So, a closer look to the proteins involved in this pathway was taken in order to try to understand the reason why this pathway was overexpressed. The protein 2-3-bisphosphoglycerate independent phosphoglycerate mutase is a glycolytic enzyme that catalyses the interconversion of 3-phosphoglycerate and 2-phosphoglycerate; The acetate kinase participates in the synthesis of the central metabolic intermediate acetyl-CoA from acetate or for generation of ATP from excess acetyl-CoA; The enzyme 2-oxoacid: acceptor oxidoreductase family protein in *H. pylori*, is an enzyme that forms part of the major routes for the generation of acetyl-CoA and succinyl-CoA. Lastly, the pyruvate synthase catalyses the direct conversion of pyruvate into acetyl-CoA. By looking into these proteins, we conclude that the methane metabolism is not the main function of those enzymes. They belong to the glycolytic pathway and other carbon pathways, and the methane metabolism is only a very small size contribution, therefore it is considered insignificant.

At 10.5% oxygen availability many pathways indicating cell division is taking place are shown: these include biosynthesis of amino acids, ribosomes and cysteine and methionine biosynthesis. Also, other pathways such as ABC transporters and inositol phosphates were activated. Inositol phosphates function as second messengers for a variety of extracellular signals. Also triggers numerous cellular processes by regulating calcium release from internal stores. Being probably one of the most important regulation mechanisms in the cell. ABC or ATP-binding cassette transporters constitute a ubiquitous superfamily of integral membrane proteins that are responsible for the ATP-powered translocation of many substrates across membranes. They can be used as importers and exporters. We will see in the protein analysis what specific compounds they transport.

At 21% oxygen availability unusual pathways are differentially expressed. Some of them are porphyrin and propanoate metabolism, quorum sensing and valine, leucine and isoleucine degradation. Quorum sensing in biology, is a phenomenon in which signalling pathways are activated in response to cell population density. All these pathways seem to be related with a control and regulation of cell division, as contrary as showed at 10.5% oxygen conditions, cell division is slowed down in order to enable survival by optimizing the existing resources. It is seen also a slight overexpression of methane as well. This could be related with the methane formation fund in *Bacillus subtilis* result to free ion and reactive oxygen species. In this microorganism It has been shown how ROS-induced methyl radicals, derived from organic compounds, are key intermediates that lead to methane production (Ernst *et al*., 2022).

3.1.2 Cluster map defines two bacterial behaviours according to the oxygen availabilities

The clustermap is the second graph used in this project and was created in order to compare a bigger number of pathways and grouping them by similar trends. In this clustermap, represented in Figure 4, the 30th most overexpressed pathways of each condition were compared. Every color is correlated with a level of expression. Red colors represent a higher overexpression ratio while blue colors represent a lower expression ratio. On the other hand, white colors mean neutral expression.

By this graph It can be seen how even though we have three conditions, the data is really divided in two sample groups. In oxygen limitation (1.05% DOT) or in oxygen availability (10.5 and 21% DOT).

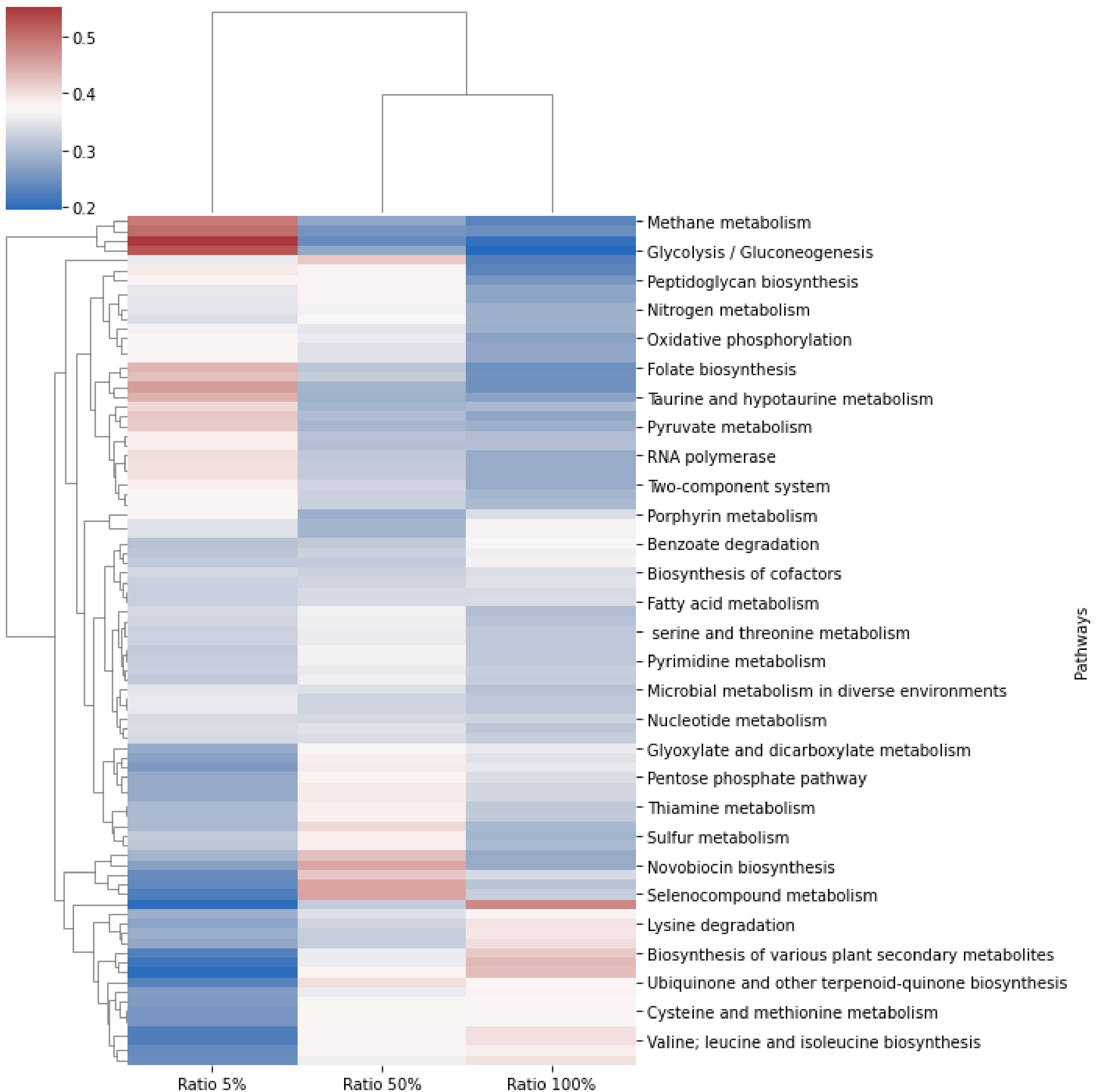

Figure 4. Clustermap of the normalized ratios of the most overexpressed pathways (p-value < 0.2) in every oxygen condition. The values are calculated based on overall changes in protein abundances that are involved in each pathway.

In conditions of limiting oxygen, other pathways different than the ones mentioned in the last section (3.1) such as the peptidoglycan biosynthesis, the nitrogen metabolism, the folate biosynthesis and the taurine and hypotaurine metabolism are overexpressed. The most important facts to point out are that peptidoglycans are one of the major components of the cell wall, taurine and hypotaurine are used to synthesize new amino acids and folate is needed for cell division.

On the other hand, in conditions of oxygen availability at 10.5% the novobiocin metabolism and selenocompund metabolism are expressed. The novobiocin is an aminocoumarin antibiotic which acts against gram-positive bacteria and inhibits bacterial DNA synthesis by targeting the bacterial DNA gyrase and the DNA topoisomerase IV. (PUBCHEM, 2022). Also a few studies have reported that certain selenocompounds have shown an interesting antibacterial activity specially against gram-positive bacteria (Mosolygó *et al.*, 2019). It is interesting this microorganism wastes resources on the synthesis of antibacterial compounds. This synthesis could mean that the bacteria is found at near optimal conditions so, it can waste resources on eliminating competitors and improve its own growth.

At oxygen availability of 21%, pathways such as the biosynthesis of secondary metabolites and the ubiquinone and other terprenoid-quinone biosynthesis are also overexpressed.

3.2 Overexpressed proteins

The 10 proteins with lowest p-values from each sample: membrane and whole cell were harvested in different tables for each oxygen condition. Low p-value equals to most significant differential expressions. Therefore, these proteins will be the ones that present the most significant changes between the three conditions.

The three columns in the right represent the ratios on the three different conditions so comparison is more visual. The proteins with more significance, more interesting or unknown functions will be explained.

3.2.1 1.05% Oxygen Inflow

Table 2. The top 20 proteins that show lowest p-value under 1.05% oxygen conditions. The first 10 proteins come from the whole cell proteome sample and the last 10 proteins come from the membrane protein sample. The membrane protein sample at 21% oxygen is the reference point so all conditions even from the whole cell proteome can be compared.

NCBI Accession	Protein Description	p-value	Ratio 1.05%	Ratio 10.5%	Ratio 21%
QZT34701	Alpha-glucosidase	0.00049	4.84	2.47	1.99
QZT33205	AdoMet dependent heme synthase	0.00051	15.1	4.4	2.77
QZT33839	Acyl-CoA dehydrogenase family protein	0.00105	1.63	1.1	1.14
QZT33206	DUF2249 domain-containing protein	0.00139	2.63	0.77	0.55
QZT34259	Adenylosuccinate synthase	0.00166	1.52	1.06	1.09
QZT33847	Uroporphyrinogen decarboxylase	0.00227	1.17	0.81	0.72
QZT33190	Molybdenum cofactor guanylyltransferase	0.00339	2.66	0.99	0.84
QZT34311	D-alanyl-D-alanine carboxypeptidase	0.00532	0.98	0.65	0.52
QZT35443	Crp/Fnr family transcriptional regulator	0.00799	4.16	1.15	0.95
QZT34586	Cyclic pyranopterin monophosphate synthase	0.01238	4.17	1.5	1.45
QZT34159	NADH-quinone oxidoreductase subunit NuoI (NDH-I)	0.00018	2.24	1.9	1
QZT33192	FAD-dependent oxidoreductase	0.00556	6.64	1.73	1
QZT33643	Metalloregulator ArsR/SmtB family transcription factor.	0.00648	3.75	1.66	1
QZT33452	Cytochrome c oxidase ba3 subunit I	0.01227	19.13	2.92	1
QZT33194	Electron transfer flavoprotein subunit beta/FixA family	0.01267	7.46	1.9	1
QZT34156	NADH-quinone oxidoreductase subunit L (NDH-I)	0.01423	1.7	1.55	1
QZT33206	DUF2249 domain-containing protein	0.01452	4.19	1.34	1
QZT33883	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	0.02046	2.12	1.43	1
QZT35056	YtxH domain-containing protein	0.02219	1.41	1.38	1
QZT34161	NADH-quinone oxidoreductase subunit D. (NDH-I)	0.02219	1.57	1.22	1

Too many proteins showed significant changes in the expression levels when oxygen becomes limiting. On Table 2, only the proteins with most different expression levels can be seen. Highest ratios are always found at 1.05% oxygen column because the proteins that showed an overexpression at 1.05% oxygen condition were selected.

CARBON METABOLISM

The protein that shows a most significant change, explained by the lowest p-value, is alpha glucosidase. This enzyme is responsible for the breaking down of sucrose into glucose and fructose converting the main carbon source into digestible disaccharides participating in the carbon metabolism. One of the more interesting pathways in *C. thermarum* as it has been previously seen in the pathway analysis.

Also, a protein involved in the carbohydrate metabolism is found, the enzyme 2,3-bisphosphoglycerate-independent phosphoglycerate mutase**.** It has been explained in the previous section; but the main role, is to degrade carbon substrates.

ELECTRON TRANSPORT CHAIN

Two enzymes from the heme biosynthetic pathway appear in the table, AdoMet dependent heme synthase and Uroporphyrinogen decarboxylase. Many different proteins with heme prosthetic groups could be synthesized by this pathway, therefore by exploring all the possible products it was concluded that the most likely is the cytochrome ba_3 membrane protein due to its correlation in the overexpression trend. The other possibilities didn't make sense due that they weren't found on the proteome or had expression levels very different. Besides the synthesis pathway being activated, the cytochrome *c* oxidase ba_3 itself is as well found as overexpressed. And lastly, two subunits of the NDH-I complex I are found, subunit NuoI and subunit L.

The FAD-dependent oxidoreductase with p-value of 0.00556 catalyses ubiquinone reduction, linking oxidation of fatty acids and some amino acids to the mitochondrial respiratory chain. FAD is the electron donor to ubiquinone. So, it is found in the complexes that reduce ubiquinone: Complex I and complex II. It is unknown which one of the two forms part.

Also, the NADP-dependent isocitrate dehydrogenase is found. An ubiquitous enzyme that catalyses the oxidative decarboxylation of isocitrate, producing 2-oxoglutarate, CO_2 and NADH or NADPH. It is believed to participate in the intracellular redox balance by regulating the ratio of NAD^+/NADH.

On the other hand, the electron transfer flavoprotein subunit beta/FixA family protein participates in energy metabolism: energy production and conversion. Electron transfer flavoproteins (ETFs) serve as specific electron acceptors for primary dehydrogenases, transferring the electrons to terminal respiratory systems. They can be functionally classified into constitutive, "housekeeping" ETFs, mainly involved in the oxidation of fatty acids and ETFs produced by some prokaryotes under specific growth conditions, receiving electrons only from the oxidation of specific substrates.

SIGNALLING

An interesting family protein has also been found at 1.05% oxygen concentrations. The Crp/Fnr transcription regulator family that typically functions as an activator and responds to a variety of intracellular or extracellular signals such as anoxia, carbon monoxide, temperature, nitric oxide, oxide and nitrosative stress and control many metabolic pathways such as aromatic compound degradation and respiration. (Zhou *et al.*, 2012).

There are four different proteins that belong to this family. The most common are the cAMP receptor protein (CRP) and the fumarate and nitrate reduction regulator (FNR). CRP suffers a conformational change when Cyclic-AMP (cAMP) is bind, and then, the CRP-cAMP complex interacts with DNA and RNA polymerase to regulate transcription. Some roles of cAMP are the regulation of antibiotic production, phototrophic growth, pathogeneses, and nitrogen fixation but the most relevant implication is believed to be in energy metabolism. On the other hand, FNR senses oxygen via an iron-sulphur cluster ligated through cysteine residues and has more interesting roles, different in every microorganism: For example, *E. coli* adopts different metabolic modes in response to availability of oxygen. A hierarchy of metabolism exists in which aerobic respiration is preferred to anaerobic respiration. This order reflects the relative energetic benefit derived from each metabolic mode. The role of FNR in *E. coli* is to preserve this hierarchy by monitoring environmental oxygen and in combination with other

transcription factors, coordinate the switch between aerobic and anaerobic metabolic modes. Another example is *Bacillus subtilis* that the transcriptional activator FNR enables the microorganism to grow anaerobically in the presence of nitrate as a terminal electron acceptor.

Due to a sequence comparison using the BlastP tool from the NCBI revealed an amino acid sequence more similar to *B. subtilis* with a query cover of 94%. It would be interesting to check whether *C. thermarum* could use nitrogen, like its relative *B. subtilis,* to use as terminal acceptor in the ETC. (Nakano et al., 1996). However, it is very unlikely. Therefore, the most reasonable function of the Fnr transcription factor would be more similar to the one of E. coli, acting as a regulator of the metabolic switches.

To detect other signalling proteins that participate in the microorganism *C. thermarum*, all the proteins that participated in the two-component system were filtered in the data output from the proteomics analysis. An interesting protein, with accession number QZT33141, was found to be differentially expressed significantly with a p-value of 0.045. This protein was called the OmpR family, response regulator ResD and it was found to be overexpressed also at 1.05% oxygen availability. The protein ResD belongs to the bacterial two-component signal transduction system ResDE which plays a crucial role in aerobic and anaerobic respiration. This system is activated when changes in oxygen availability are detected. When ResD undergoes a ResE dependent phosphorylation, is thought to activate transcriptionally anaerobically induced genes. In each microorganism, different genes are transcribed; In *Bacillus subtilis* genes such as *fnr*, *hmp* and *nasD* are activated*,* but in other species such as *Bacillus cereus,* it has been found that activates other genes such as virulence factors. In E. coli genes like those involved in cytochrome c biogenesis, ctaA (responsible for conversion of haem O to haem A) and petCBD (encoding subunits of the cytochrome bf complex) are encoded. (Nakano *et al*., 2000).

The most probable signalling pathway showed in *C. thermarum* from this evidence is that the ResDE system is activated when there are anoxia conditions, and it would be of interest to check what genes are activated by this signalling system.

OTHERS

The enzyme cyclic pyranopterin monophosphate synthase catalyses an early step in the biosynthesis of molybdenum cofactor. More than 50 molybdoenzymes were identified to date and important redox reactions in global metabolic cycles. These proteins contain the molybdenum cofactor that performs a two-electron oxidation-reduction reaction that involves the net transfer of an oxygen atom between the substrate and water. It could be interesting to check in what functions is this cofactor involved.

The electron transfer flavoprotein subunit beta/FixA family is actually a confused term due to the electron transfer flavoprotein subunit beta possess a notable similarity with FixA but are two different enzymes that possess different functions. FixA is essential for symbiotic nitrogen fixation and the classical role of ETF is in β-oxidation of fatty acids and act as intermediate electron carriers between primary dehydrogenases and terminal respiratory systems (Roberts *et al*., 1999). To check which one of the two enzymes could be a BLAST sequence comparison was performed in the NCBI database. All results showed the same confusion, any protein could actually distinguish between the two options. It is most probable that this protein is more alike to the electron transfer flavoprotein subunit beta because any clue that *C. thermarum* is able to perform nitrogen fixation hasn't been found yet.

3.2.2 10.5% Oxygen Inflow

Table 3. Top 20 proteins overexpressed under 10.5% oxygen. Again, the first 10 proteins from the table are the proteins that show lowest p-value from the whole cell proteins sample and the last 10 proteins from the table are the proteins that show lowest p-value from the membrane proteins sample.

NCBI Accession	Protein Description	p-value	Ratio 1.05%	Ratio 10.5%	Ratio 21%
QZT32738	CCA tRNA nucleotidyltransferase	0.00629	1.56	2.1	1.13
QZT32948	cytochrome c oxidase subunit I	0.00637	0.23	0.43	0.39
QZT33398	molecular chaperone DnaJ	0.00715	0.35	0.5	0.49
QZT34788	DUF3870 domain-containing	0.00886	1.41	1.94	1.83
QZT34739	CRISPR-associated endoribonuclease	0.00926	1	1.66	1.48
QZT33306	Phosphonate ABC transporter ATP-binding protein	0.01056	0.67	0.89	0.7
QZT32575	Peroxiredoxin	0.01360	1.35	2.2	2.02
QZT34035	LCP family protein	0.01723	0.68	0.98	0.81

QZT32487	O-succinylbenzoate synthase	0.01967	1.18	2.44	2.26
QZT35208	Lysophospholipase	0.02015	0.74	0.95	0.7
QZT33911	Permease-like cell division protein FtsX	0.01862	1.39	1.64	1
QZT32756	DNA topoisomerase IV	0.02262	0.82	1.18	1
QZT34641	ABC transporter ATP-binding protein	0.02262	1.11	1.61	1
QZT32723	Prephenate dehydrogenase	0.02428	0.95	1.6	1
QZT34188	50S ribosomal protein L31	0.02487	1.31	1.43	1
QZT34388	UvrB/UvrC motif-containing	0.02491	1.48	1.85	1
QZT34262	DHH family phosphoesterase	0.02541	0.82	1.16	1
QZT34355	MazG family protein	0.02955	0.75	1.08	1
QZT34249	23S rRNA (pseudouridine(1915)-N(3))-methyltransferase	0.02998	0.9	1.44	1
QZT33026	GTP-sensing pleiotropic transcriptional regulator	0.03372	1.71	1.93	1

CELL DIVISION

Many interesting proteins appear at this stage. Most of them appear to contribute to an increase in the cell density. Some of the enzymes that show this process are CCA tRNA nucleotidyltransferase, a RNA polymerase responsible for the synthesis of the nucleotide triplet CCA at the 3′-terminus of tRNAs**;** chaperone DnaJ, also known as Hsp40 , plays a role in regulating the ATPase activity of Hsp70 heat-shock proteins. They assist a large variety of protein folding processes in the cell by transient association of their substrate binding domain with short hydrophobic peptide segments within their substrate proteins; Permease-like cell division protein FtsX is also overexpressed under these conditions. This protein forms part of a membrane complex widely conserved and involved in critical processes such as recruitment of division proteins and in spatial and temporal regulation of muralytic activity during cell division or sporulation; DNA topoisomerase IV is responsible for the decatenation of interlinked chromosomes and participates in DNA replication; 50S ribosomal protein L31 and 30S ribosomal protein S4 are structural proteins that make the ribosomes. And lastly, UvrB/UvrC motif-containing protein is involved in nucleotide-excision repair participating in DNA replication.

ELECTRON TRANSPORT CHAIN

Furthermore, there are more interesting proteins that are not implicated in cell division. The cytochrome *c* oxidase subunit I is found in the analysis, and It Forms part of the cytochrome aa_3. Showing a second type of complex IV or terminal acceptor.

The enzyme O-succinylbenzoate synthase codified by the menC gene is involved in one of the steps of the menaquinone biosynthesis pathway; The DHH family phosphoesterase are enzymes that catalyze the hydrolysis of phosphodiester bonds via a mechanism involving two Mn^{2+} ions. This way, enzymes that have been phosphorylated loss their phosphate group altering its function (Breaker and Joyce, 1995).

OTHERS

The antioxidant enzyme peroxiredoxin is a peroxidase and participates in the oxygen protection. Its function is the reduction of peroxides, specifically hydrogen peroxide, alkyl hydroperoxides, and peroxynitrite.

The DNA integrity scanning di-adenylate cyclase DisA converts two ATP or ADP molecules into one cyclic-di-AMP molecule. The di-adenylate cyclase activity is strongly suppressed by binding to branched DNA, but not to duplex or single-stranded DNA, suggesting a role for DisA as a monitor of the presence of stalled replication forks or recombination intermediates via DNA structure-modulated Cyclic di-AMP synthesis. Cyclic di-AMP is a bacterial secondary messenger molecule, which is associated with various physiological functions. It is involved in several important cellular processes, such as cell wall metabolism, maintenance of DNA integrity, ion transport, transcription regulation, and allosteric regulation of enzyme function. The overexpression of this enzyme could be due to cell division or to harm in the DNA by oxygen ROS.

MazG family protein, also overexpressed at 10.5% oxygen conditions, It Is actually an unexpected finding. This protein has a nucleotide pyrophosphohydrolase activity being toxic for the cell when is found in abundancy affecting cell growth and survival. Doesn’t make sense with the rest of proteins, but another role was found, suggesting its possible ‘house-cleaning’ function by hydrolysing and elimination of abnormal NTPs that are result of oxidative damage.

It could be possible, the bacteria starts to present moderate oxygen damage at 10.5% (Gross *et al.*, 2006).

The ABC transporters found seem to transport sulfonates and phosphonates. Both inorganic compounds are essential elements for growth so, this ATP- dependent transporters seem to be the principal point of entry of inorganic sulphur and phosphate.

3.2.3 21% Oxygen Inflow

Table 4. Top 20 proteins expressed under 21% oxygen. The 10 proteins that show lowest p-value from the whole cell proteins sample and the membrane proteins sample.

NCBI Accession	Protein Description	p-value	Ratio 1.05%	Ratio 10.5%	Ratio 21%
QZT33602	RimK family alpha-L-glutamate ligase.	0.00212	2.3	1.96	2.51
QZT32889	Thiol peroxidase	0.00352	1.05	1.36	1.99
QZT32486	M20 peptidase aminoacylase family protein	0.00721	0.6	1.1	1.34
QZT35267	RidA family protein	0.00862	1.62	1.96	2.35
QZT33471	ABC transporter ATP-binding protein	0.01289	0.63	0.7	1.08
QZT33832	Precorrin-2 C (20)-methyltransferase	0.01486	0.92	1.05	1.41
QZT35245	RraA family protein	0.01658	1.31	2.87	4.15
QZT33817	Alpha glucosidase	0.01763	0.78	1.05	1.1
QZT35265	AbrB/MazE/SpoVT family DNA-binding domain-containing protein	0.01875	4.53	5.37	5.93
QZT35523	cobalamin biosynthesis	0.01920	0.73	0.76	1.02
QZT32757	DNA topoisomerase IV	0.00062	0.62	0.91	1
QZT34268	CoA transferase	0.01336	0.64	0.98	1
QZT34810	PAS domain S-box protein	0.01379	0.79	0.9	1
QZT32955	Stage V sporulation protein	0.02325	0.69	0.87	1
QZT34317	Dipeptide ABC transporter ATP-binding protein	0.02420	0.43	0.63	1
QZT34282	Sugar ABC transporter substrate-binding protein	0.02701	0.43	0.65	1
QZT33030	Type I DNA topoisomerase	0.03210	0.57	0.71	1
QZT34272	BMC domain-containing	0.03285	0.43	0.63	1
QZT34344	pur operon repressor	0.03415	0.67	0.8	1
QZT34518	DEAD/DEAH box helicase	0.03743	0.83	0.93	1

Very unexpected and different proteins are overexpressed in this condition, but most of them are indicators that the bacteria is livening in a hostile environment.

The RimK family alpha-L-glutamate ligase catalyzes the final reactions in the synthesis of Coenzyme F_{420} or 8-hydroxy-5-deazaflavin, an alternative coenzyme used in the ETC as an electron carrier. This coenzyme is structurally more similar to FAD and FMN but is catalytically to NAD and NADP. It has a relatively low redox potential of -340mV under standard conditions and -380mV under certain physiological conditions (Ney *et al.*, 2017). It is then, possible that F420 functions as an alternative cofactor in the ETC by reducing the electron flow due to the low redox potential and therefore slowing down the ETC.

The M20 peptidase aminoacylase family protein function involves the release of an N-terminal amino-acid, usually neutral or hydrophobic, from a polypeptide altering its normal functioning. The appearance of peptidases equals to misfolding and protein problems, most likely due to ROS.

RidA or Reactive Intermediate deaminase catalyzes the hydrolysis of enamines and imines to their ketone product. The consequences when RidA is absent are due to accumulation of the reactive metabolite 2-aminoacrylate (2AA). When this metabolic product accumulates in the cell long enough, it damages a growing list of essential metabolic enzymes. (Irons et al., 2020)

RraA family proteins catalyze the last step of the bacterial protocatechuate 4,5-cleavage pathway, one of the key catabolic routes for the degradation of aromatic compounds. These enzymes act cleaving carbon-carbon bonds. Moreover, RraA inhibits the activity of RNase E, an essential enzyme having key roles in mRNA turnover, which is determined by the rates of mRNA synthesis and degradation, which jointly adjust the level of gene expression. The capacity to adjust the stability of individual transcripts rapidly enables efficacious responses to changing environments.

There are two important signalling proteins that show an overexpression. The AbrB/MazE/SpoVT family DNA-binding domain-containing protein. The AbrB-like family proteins enable bacteria launch responses to diverse environmental stimuli. And the PAS domain S-box protein are newly recognized signalling domains that sense a wide variety of stimuli such as oxygen, redox potential or light and participates in the two-component regulatory system; a basic stimulus-response coupling mechanism to allow organisms to sense

and respond to changes in many different environmental condition. Specificity in sensing arises, in part, from different cofactors that may be associated with the PAS fold. They are fond in the internal part of the cytoplasmic membrane but can be coupled to secondary receptors of outside the membrane.

This PAS domain S-box specifically in C. thermarum is part of the KinA enzyme, forming part at the same time of the KinABCDE-Spo0FA complex and participates in the sporulation control process. *B. subtilis* cells monitor their external and internal environments with the aim of when the environment becomes hostile, they respond by forming dormant, heat-resistant endospores. The decision to sporulate is made by integrating diverse environmental and physiological signals, resulting in the activation of a key transcriptional factor, Spo0A by phosphorylation. The same process is speculated to happen in *C. thermarum*. (Eswaramoorthy *et al.,* 2009). The Stage V sporulation protein is an essential protein for sporulation in some bacteria species such as *Bacillus megaterium* and *Bacillus subtilis*. Indicates hostile environment for the bacteria so sporulation is performed.

The proteins dipeptide ABC transporter ATP-binding and peptide/nickel transport system ATP-binding protein, both participates in the quorum sensing mechanism explained previously. In biology, quorum sensing is a phenomenon in which signalling pathways are activated in response to cell population density.

OXYGEN DAMAGE AND PROTECTION AGAINS IT

Even though oxygen is necessary for obligate aerobes to survive; too much presence in the environment is actually toxic and one potential reason to produce a slowdown in the cell division. The relationship between oxygen and cell division is inverse, as oxygen availability increases grow slows down deducing oxygen is harmful for the cell (Donelli *et al*., 1998).

Those conditions of excess supply of oxygen are defined as hyperoxia. The biochemical effects hyperoxia produces in the cells are the formation of oxygen free radicals or ROS. These compounds have unpaired electrons and that makes them very unstable and very likely to react and alter the molecular compounds of the cells. The most biologically significant reactive oxidant species are the hydroxyl ion and peroxynitrite. Peroxynitrite, the product of the

reaction between superoxide and nitric oxide interacts with lipids, DNA, and proteins via direct oxidative reactions or via indirect, radical-mediated mechanism. These reactions trigger cellular responses ranging from subtle modulations of cell signalling to overwhelming oxidative injury, committing cells to necrosis or apoptosis. Although the cells possess many antioxidant systems, these are eventually overwhelmed at very high concentrations of free oxygen when the rate of oxidative damage overwhelms the capacity of the systems that prevent or repair it. Cell damage and death result (Archibald and Fridovich, 1983).

To cope with the harmful effects of oxygen and ROS, some species have developed protection systems based on the synthesis of antioxidant enzymes, such the superoxide dismutase, catalase, flavin dependent oxidases/peroxidases, thioredoxin and glutathione reductases (Ianniello *et al*., 2015).

These antioxidant enzymes are being studied in the proteome. The most abundant enzymes that participate in the reactive oxygen species (ROS) neutralization are peroxidases and catalases.

On the proteomics output there are 9 proteins with the 'peroxide', 'catalase', 'superoxide' and 'dismutase' entrance. Two of them show similar behaviour in all conditions with a high p-value. The rest 7 proteins show more significant changes in the expression level with a p-value of < 0.2 and as expected are more expressed under 21% O_2, indicating more reactive oxygen is present in the cell and more antioxidant enzymes are needed to reduce toxicity and preserve survival. These antioxidant proteins are catalase/peroxidase HPI, thiol peroxidase, glutathione peroxidase, peroxiredoxin and heme-dependent peroxidase.

Thiol peroxidase is part of an oxidative stress defense system that uses thioredoxin and thioredoxin reductase to reduce alkyl hydroperoxides. Therefore, it is an antioxidant which reduces peroxides and inhibits hydrogen peroxide response.

3.3 Membrane proteins characterization

The characterization of the respiratory chain enzymes found in the membrane in every condition is important because they are believed to be one of the main responsible mechanisms for the microorganism adaptation. Branched respiratory chains are not

uncommon; *E. coli*, for example, uses its branched respiratory chain to tune its catabolism depending on the oxygen availability. (Wissink, 2021). *C. thermarum* is another example of microorganism with a branched respiratory chain.

For the membrane protein characterization, it will be considered a statistical p-value < 0.2 the threshold for considering the results significant.

3.3.1 Complex I

Complex I is the first point of entry of electrons to the ETC. The NADH produced in the Krebs cycle is oxidized to NAD^+ and two electrons are transferred to complex I. The proteomes of *C. thermarum* TA2.A1 grown under 1.05%, 10.5% and 21% oxygen aeration, showed that two different types of NADH dehydrogenase: NDH-I and NDH-II were expressed.

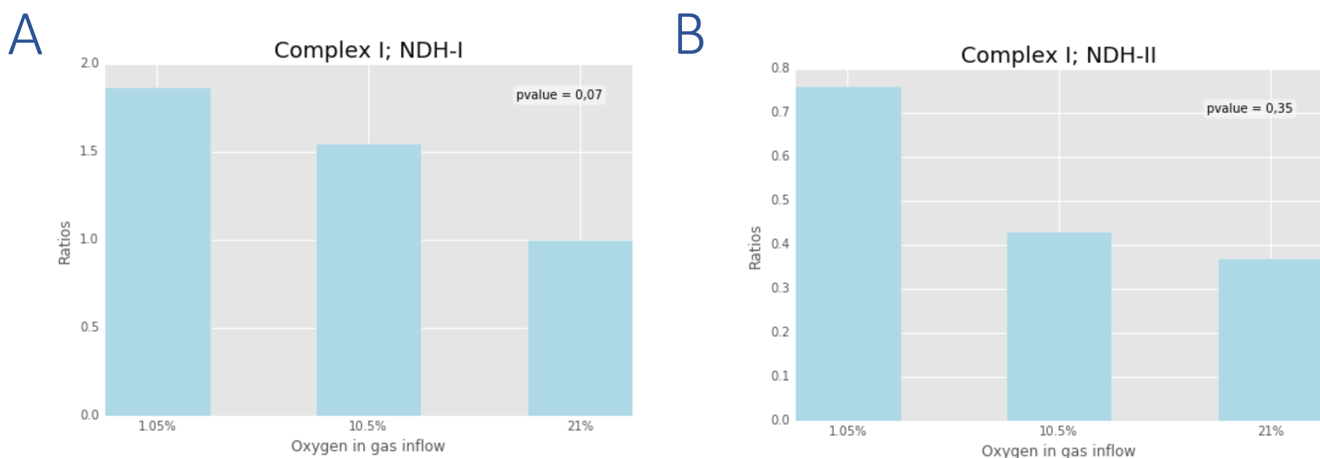

Figure 3. **A)** The graph of NDH-I is made from the mean of the 4 subunits found in the *C. thermarum* proteome: Subunit D, subunit NuoI, Subunit NuoL and Subunit NuoH. **B)** The single subunit of NDH-II was found on the proteome and the ratios are represented in the graph.

NDH-I is composed of 14 subunits in bacteria, and it is able to translocate four protons across the inner membrane every time a NADH molecule is oxidized. On the contrary, NDH-II is composed of a single subunit and does not translocate protons. Instead, it is believed to act as a regulator of the NADH, NAD^+ pool.

The two expected NDH enzymes were found on the proteome. Both look like have similar behaviours. The relative abundance of NDH-I and NDH-II increased about twofold with a decreasing oxygen availability from 21% to 1.05%. This suggests that there is an oxygen dependent regulation at play on the level of transcription or translation.

Other studies mention a third type of putative type II NDH, called YjiD. The putative enzyme would be composed of the subunit with NCBI accession number QZT33731. NAD(P)/FAD-dependent oxidoreductase. To check whether if was possible or not. The position of the gene that encodes that protein was gathered on NCBI database, because many information can be deduced. The other surrounding genes can be in the same operon or have similar regulation, so It is usual that similar genes are grouped together. In this specific gene position, many hypothetical proteins surrounded the target gene but the adjacent gene, is involved in pyridine biosynthesis. Because we don't find evidence of membrane proteins involved in ECT next to this gene, we conclude this gene is not a third type of NADH dehydrogenase.

3.3.2 Complex II

Complex II is the second component of the ETC. It is the alternative point of entry of electrons in the ETC where succinate is oxidized to fumarate and the electrons result of this reaction enter the ETC, although no ion pumping of this enzyme has been found.

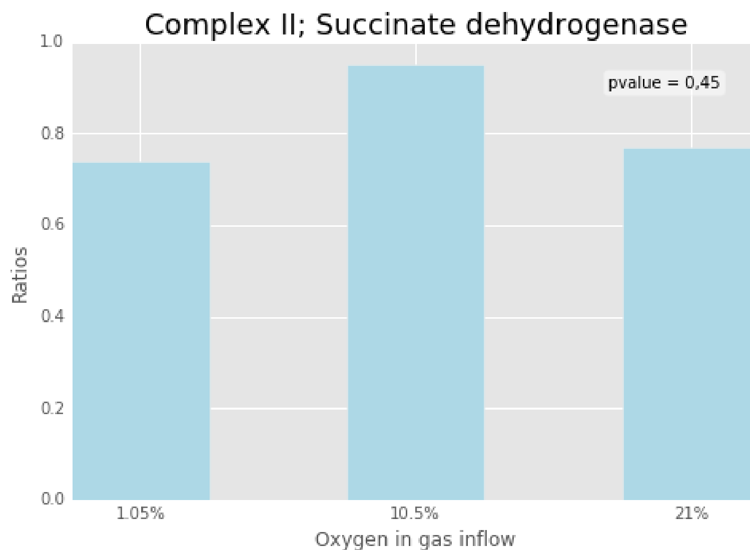

Figure 4. Only the succinate dehydrogenase iron-sulfur subunit (dshb) was found on the proteome. This graph represents the amount of this subunit in the three samples.

The p-value is too high to consider there is a significant change in the expression levels. But It is seen a slight increase of complex II at 10.5% oxygen conditions.

3.3.3 Quinones

The only type of quinones that is believed to be found in *C. thermarum*'s ECT is menaquinone due to the biosynthesis mechanisms found in the proteome. Furthermore, in the genome only menaquinone biosynthesis genes are found as well. We state menaquinone is the responsible for the electron carrier between complexes of the transport chain.

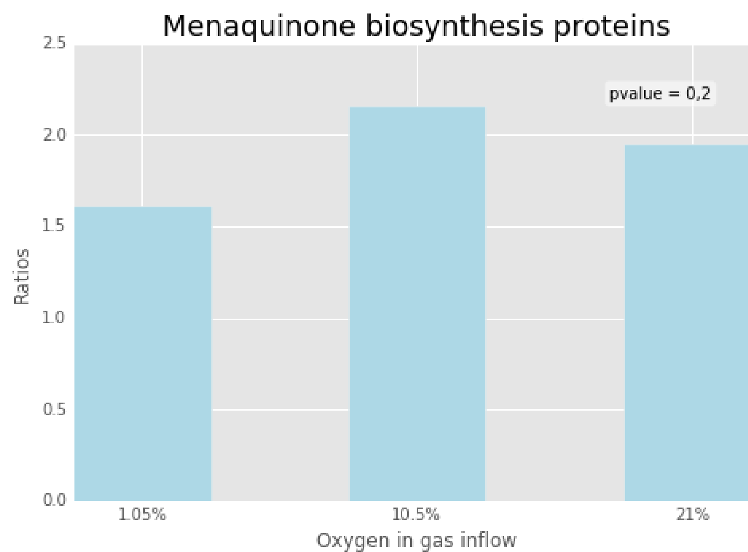

Figure 5. Graph made from the average between the menaquinone biosynthetic proteins. The proteins found on the proteome are menB, menC, menE and menF

3.3.4 Complex III

The complex III receives the pair of electrons from the menaquinones and the pair is split up into single electrons to be able to donate them to the ubiquinones and keep the electron flow translocating the electrons one by one to the next enzymes.

It is hypothesized that the complex III found in *C. thermarum* is the putative b6c1. In this project only the c1 subunit was characterized from the proteome. The b6 subunit wasn't found because we hypothesize it was lost in the MS experiment due to the high hydrophobicity of the molecule (Baniulis et al., 2008). An increase of the C1 subunit is seen with decreased oxygen availability.

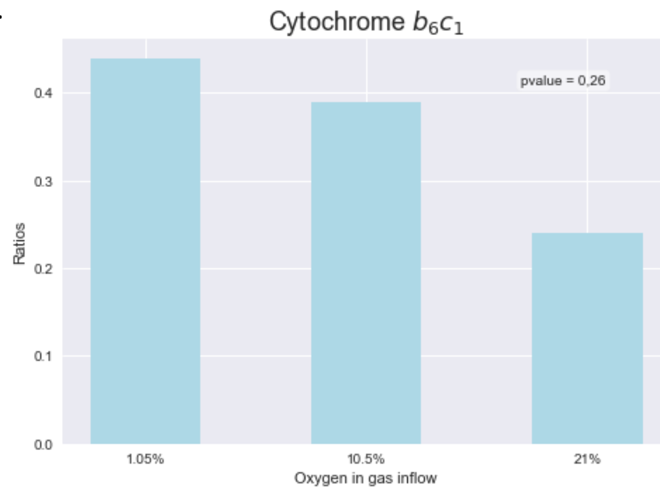

Figure 6. Abundance of the only subunit of complex III found in the proteome, the C_1 subunit.

3.3.5 Cytochrome C

The cytochrome C is the protein in the ETC that carries single electrons from complex III to the terminal oxidases or complex IV.

The cytochrome C protein itself wasn't found on the proteome. But the protein responsible for its synthesis was present. This protein is called thiol-disulphide oxidoreductase ResA, and It is differentially expressed at 1,05%.

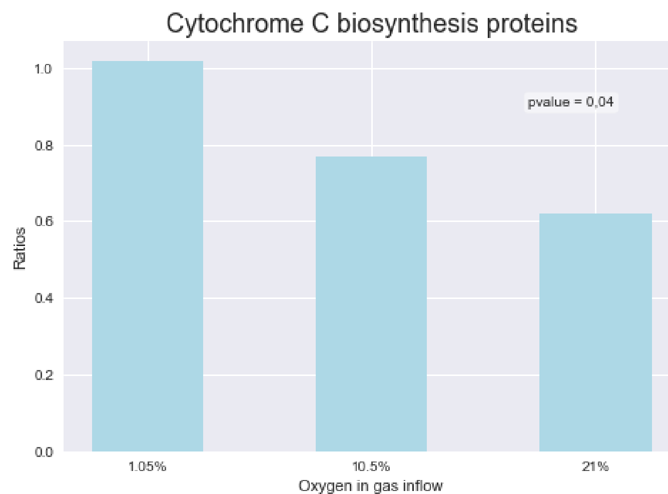

Figure 7. Ratios of the proteins involved in the cytochrome C biosynthesis. An overexpression on 1.05% oxygen condition is shown.

3.3.6 Complexes IV. Terminal oxidases.

The terminal oxidases are the enzymes responsible for the direct transference of electrons to a terminal acceptor molecule. In *C. thermarum* only terminal oxidases that transfer the electron to oxygen are found. However, there has been found proteins similar to Bacillus subtilis that activate nitrate respiration.

In *C. thermarum* there are four terminal oxidases encoded in the genome: cytochromes aa_3, ba_3, bb_3 and *bd* despite of only two of them are found in the proteome: aa3 and ba3 and interestingly they show opposite trends.

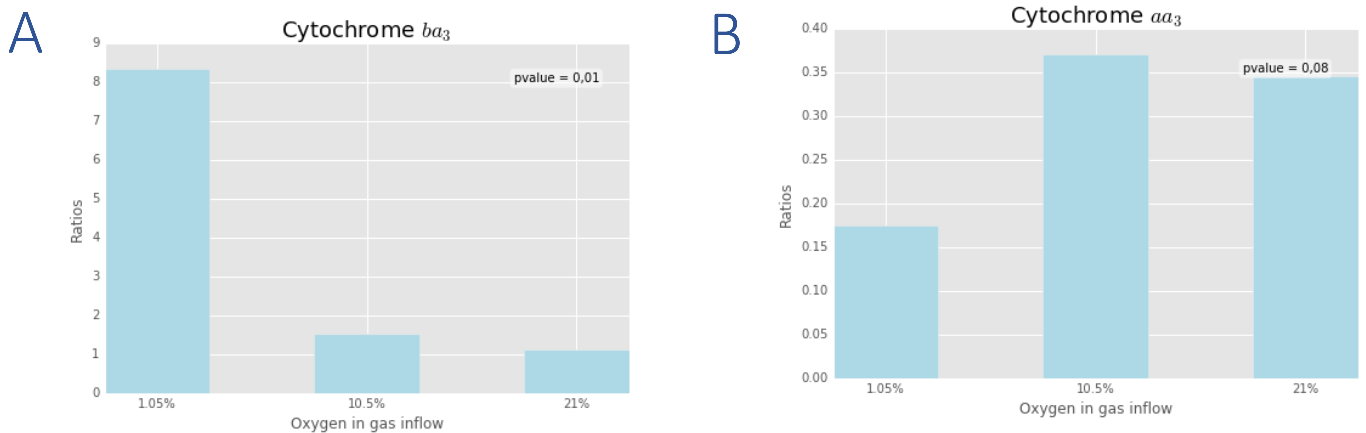

Figure 8. A) The ba3 protein is the most differentially expressed protein from the whole proteome. It is used at low oxygen concentration hypothesizing It has the highest affinity for oxygen. B) The ratio of the aa3 terminal oxidase is represented. It is overexpressed with higher oxygen concentrations, when It isn't limiting.

With low oxygen availability, cytochrome ba_3 is massively overexpressed while with higher oxygen aa_3 looks to take over. Then ba_3 shows a higher affinity for oxygen, making possible the aerobic respiration when low oxygen is available. It is unknown the reason why at higher oxygen concentration It changes of terminal oxidase and doesn't always maintain the one with high affinity.

3.3.7 Complex V. ATP Synthase.

The ATP synthase is the enzyme responsible for ATP generation. Alkaliphile ATP synthases are highly adapted due to the special conditions these are found at. The pH in alkaliphile

microorganisms are inverted. Bacterial F1F0-ATP synthases are composed of eight subunits: $\alpha_3 \beta_3 \varepsilon \delta \gamma a b_2 c_{10-15}$. The subunits found in the *C. thermarum* proteome were the a,b and the α.

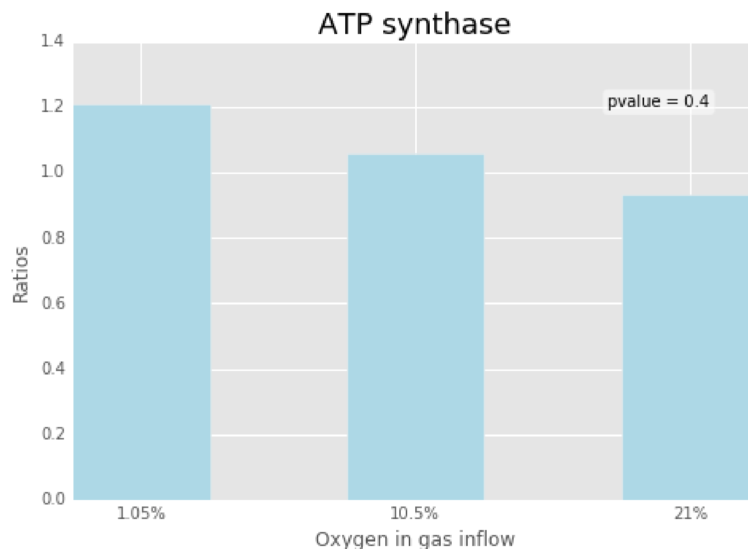

Figure 9. Average ratios of the three subunits found in the proteome, subunit a, subunit b, and subunit alpha. The decreasement of the expression levels of ATP synthase aren't very significant.

A slight decrease of the ATP synthase abundance with the amount of OD is seen. The reason why more ATP molecules are needed with less oxygen are unknown.

4. Conclusions

The microorganism *C. thermarum* TA2.A1 is capable of living under very different environmental conditions. The effect of oxygen has been studied in a proteome level to achieve a better understanding of this adaptation.

At low oxygen concentrations, the proteins that showed major differential expressions are involved in the carbon metabolism. Pathways such as glycolysis, the TCA cycle or the pyruvate metabolism are overexpressed altogether with proteins such as α-glucosidase. This over-activation of the carbon metabolism is correlated with the appearance of fermentation at this

low oxygen conditions, meaning more sucrose and glutamate need to be used in order to obtain the same energy due to lower energy obtention yields. The oxidative phosphorylation chain also undergoes structural changes; complex I, complex III and the cytochrome *c* are overexpressed. Also, the terminal oxidase ba_3 appears, indicating the standard terminal oxidase aa_3 is replaced with one with higher oxygen affinity.

At microaerophilic conditions or 10.5% DOTs, pathways and proteins indicating cell division are overexpressed indicating the microorganism is likely found at the most optimal conditions. Moreover, the phosphorylation oxidative chain undergoes again a structural change with an overexpression of menaquinones and the terminal acceptor ba_3 is replaced with aa_3. Antibiotics are synthesized to eliminate the competitors in the environment. Evidence of modest damage is found, because an antioxidant protein is overexpressed and a protein that degrades damaged NTPs is found meaning oxidative stress is starting to appear. This microorganism starts to show sensitivity of oxygen.

At standard atmosphere oxygen concentrations of around 21%, the microorganism shows the activation of sporulation, the mechanism of quorum sensing to regulate the cell density, or the overexpression of amino acid degradation pathways. All these mechanisms are indications that the environmental conditions aren't the most adequate for them. Clues that oxygen is harming the microorganism are also shown with the overexpression of antioxidant enzymes and peptidases to eradicate protein damage. Although these are not the optimum environmental conditions, the bacteria show to possess alternative mechanisms that enable survival. The coenzyme F_{420} seems to have a very important role by slowing down the ETC.

Summarizing, the most optimal conditions for *C. thermarum* growth are around 10.5%, when the DOTs levels decrease, growth also decrease until reaching anoxia conditions where the microorganism can't survive. The adaptation to low oxygen concentrations comes with sensible sensing systems that activate alternative oxidative phosphorylation chains and fermentation pathways. On the other hand, when DOT levels rise, oxidative stress is done at the cells, then many antioxidant enzymes have to be synthesized and peptidases to eliminate the harmed macromolecules. The growth is controlled by using mechanisms of quorum sensing and the use of the coenzyme F_{420} that slows down the electron transfer in the ETC.

5. Bibliography

Archibald, F. S., & Fridovich, I. (1983). Oxygen radicals, oxygen toxicity and the life of microorganisms. *Acta medica portuguesa*, 4(2), 101-112.

Baniulis, D., Yamashita, E., Zhang, H., Hasan, S. S., & Cramer, W. A. (2008). Structure–function of the cytochrome b6f complex. *Photochemistry and photobiology*, *84*(6), 1349-1358.

Bonilla, C. Y. (2020). Generally stressed out Bacteria: environmental stress response mechanisms in gram-positive Bacteria. *Integrative and Comparative Biology*, 60(1), 126-133.

Breaker, R. R., & Joyce, G. F. (1995). A DNA enzyme with Mg2+-dependent RNA phosphoesterase activity. *Chemistry & biology*, 2(10), 655-660.

de Jong, S. I., van den Broek, M. A., Merkel, A. Y., de la Torre Cortes, P., Kalamorz, F., Cook, G. M., ... & McMillan, D. G. (2020). Genomic analysis of Caldalkalibacillus thermarum TA2.A1 reveals aerobic alkaliphilic metabolism and evolutionary hallmarks linking alkaliphilic bacteria and plant life. *Extremophiles*, *24*(6), 923-935.

Dhiman, R. K., Mahapatra, S., Slayden, R. A., Boyne, M. E., Lenaerts, A., Hinshaw, J. C., ... & Crick, D. C. (2009). Menaquinone synthesis is critical for maintaining mycobacterial viability during exponential growth and recovery from non-replicating persistence. *Molecular microbiology*, 72(1), 85-97.

Donelli, G., Matarrese, P., Fiorentini, C., Dainelli, B., Taraborelli, T., Di Campli, E., ... & Cellini, L. (1998). The e¡ect of oxygen on the growth and cell morphology of Helicobacter Pilory. *FEMS Microbiology Letters 168*, 9-15.

Ernst, L., Steinfeld, B., Barayeu, U., Klintzsch, T., Kurth, M., Grimm, D., ... & Keppler, F. (2022). Methane formation driven by reactive oxygen species across all living organisms. *Nature*, 603(7901), 482-487.

Eswaramoorthy, P., Guo, T., & Fujita, M. (2009). In vivo domain-based functional analysis of the major sporulation sensor kinase, KinA, in Bacillus subtilis. *Journal of bacteriology,*, 191(17), 5358-5368.

Gross, M., Marianovsky, I., & Glaser, G. (2006). MazG–a regulator of programmed cell death in Escherichia coli. *Molecular microbiology*, 59(2), 590-601.

Ianniello, R. G., Ricciardi, A., Parente, E., Tramutola, A., Reale, A., & Zotta, T. (2015). Aeration and supplementation with heme and menaquinone affect survival to stresses and antioxidant capability of Lactobacillus casei strains. *LWT-food Science and Technology*, 60(2), 817-824.

Irons, J. L., Hodge-Hanson, K., & Downs, D. M. (2020). RidA proteins protect against metabolic damage by reactive intermediates. *Microbiology and Molecular Biology Reviews*, 84(3), e00024-20.

Ito, S. (2011). *Alkaline Enzymes in Current Detergency.* Tokyo: Horikoshi K. (eds) Extremophiles Handbook. Springer.

Kaakoush, N. O., Miller, W. G., De Reuse, H., & Mendz, G. L. (2007). Oxygen requirement and tolerance of Campylobacter jejuni. *ELSEVIER*, 644-650.

Kalamorz, F., Keis, S., McMillan, D. G., Olsson, K., Stanton, J. A., Stockwell, P., ... & Cook, G. M. (2011). Draft Genome Sequence of the Thermoalkaliphilic Caldalkalibacillus thermarum Strain TA2.A1. *American Society for Microbiology*, 2.

Krieg, N. R., & Hoffman, P. S. (1986). MICROAEROPHILY AND OXYGEN. *Annual Reviews Microbiology*, 40:107-30.

McMillan, D. G., Keis, S., Berney, M., & Cook, G. M. (2009). Nonfermentative Thermoalkaliphilic Growth Is Restricted to Alkaline. *Amercian society for microbiology*, *75*(24), 7649-7654.

Mosolygó, T., Kincses, A., Csonka, A., Tönki, Á. S., Witek, K., Sanmartín, C., ... & Spengler, G. (2019). Selenocompounds as novel antibacterial agents and bacterial efflux pump inhibitors. *Molecules*, 24(8), 1487.

Nakano, M. M., Zhu, Y., LaCelle, M., Zhang, X., & Hulett, F. M (2000). Interaction of ResD with regulatory regions of anaerobically induced genes in Bacillus subtilis. *Molecular microbiology*, 37(5), 1198-1207.

Nakano, M. M., Zuber, P., Glaser, P., Danchin, A., & Hulett, F. M. (1996). Two-component regulatory proteins ResD-ResE are required for transcriptional activation of fnr upon oxygen limitation in Bacillus subtilis. *ournal of bacteriology*, 178(13), 3796-3802.

National Center for Biotechnology Information (2022). PubChem Compound Summary for CID 54675769, Novobiocin. Retrieved August 14, 2022 from https://pubchem.ncbi.nlm.nih.gov/compound/Novobiocin.

Ney, B., Ahmed, F. H., Carere, C. R., Biswas, A., Warden, A. C. , Morales, S. E., ... & Greening, C. (2017). The methanogenic redox cofactor F420 is widely synthesized by aerobic soil bacteria. *The ISME journal, 11(1)*, 125-137.

Nitzschke, A., & Bettenbrock, K. (2018). All three quinone species play distinct roles in ensuring optimal growth under aerobic and fermentative conditions in E. coli K12. *PLoS One*, 13(4), e0194699.

Peddie, C. J., Cook, G. M., & Morgan, H. W. (1999). Sodium-Dependent Glutamate Uptake by an Alkaliphilic, Thermophilic Bacillus Strain, TA2.A. *JOURNAL OF BACTERIOLOGY*, *181*(10), 3172-3177.

Roberts, D. L., Salazar, D., Fulmer, J. P., Frerman, F. E., & Kim, J. J. P. (1999). Crystal structure of Paracoccus denitrificans electron transfer flavoprotein: structural and electrostatic analysis of a conserved flavin binding domain. *Biochemistry*, 38(7), 1977-1989.

Smith, M. A., Finel, M., Korolik, V., & Mendz, G. L. (2000). Characteristics of the aerobic respiratory chains of the microaerophiles Campylobacter jejuni and Helicobacter pylori. *Arch Microbiol*, *174*(1), 1-10.

Tank, M., & Bryant, D. A. (2015). Chloracidobacterium thermophilum gen. nov., sp. nov.: an anoxygenic microaerophilic chlorophotoheterotrophic acidobacterium. *nternational journal of systematic and evolutionary microbiology*, 65(Pt_5), 1426-1430.

Tiago, I., Chung, A. P., & Veríssimo, A. (2004). Bacterial Diversity in a Nonsaline Alkaline Environment: Heterotrophic Aerobic Populations. *American Society for Microbiology*, *70*(12), 7378-7387.

Tiwari, S., Jamal, S. B., Hassan, S. S., Carvalho, P. V., Almeida, S., Barh, D., ... & Azevedo, V. (2017). Two-component signal transduction systems of pathogenic bacteria as targets for antimicrobial therapy: an overview. *Frontiers in microbiology*, 8, 1878.

Tolner, B., Poolman, B., & Konings, W. N. (1997). Adaptation of Microorganisms and Their Transport Systems to High Temperatures. *ELSEVIER*, *118*(3), 423-428.

Urbieta, M. S., Donati, E. R., Chan, K. G., Shahar, S., Sin, L. L., & Goh, K. M. (2015). Thermophiles in the genomic era: Biodiversity, science, and applications. *ELSEVIER*, *33*(6), 633-647.

Wang, Q., Cen, Z., & Zhao, J. (2013). The Survival Mechanisms of Thermophiles at High Temperatures: An Angle of Omics. *PHYSIOLOGY*, *30*(2), 97-106.

Wissink, M. (2021). The Branched Respiratory Chain of Caldalkalibacillus thermarum TA2.A1 Enables Growth under Changing Oxygen Availabilities. 64.

Zeng, Y., Selyanin, V., Lukeš, M., Dean, J., Kaftan, D., Feng, F., & Koblížek, M. (2015). Characterization of the microaerophilic, bacteriochlorophyll a-containing bacterium Gemmatimonas phototrophica sp. nov., and emended descriptions of the genus Gemmatimonas and Gemmatimonas aurantiaca. *International journal of systematic and evolutionary microbiology*, 65(Pt_8), 2410-2419.

Zhou, A., Chen, Y. I., Zane, G. M., He, Z., Hemme, C. L., Joachimiak, M. P., ... & Zhou, J. (2012). Functional characterization of Crp/Fnr-type global transcriptional regulators in Desulfovibrio vulgaris Hildenborough. *Applied and environmental microbiology*, 78(4), 1168-1177.

6. ANNEX. Supplementary information.

A1. Pathways that show a greater number of proteins overexpressed.

Table A.1. In the table, the number of overexpressed proteins for every condition and pathways is shown. The pathways are sorted by abundance of overexpressed proteins.

Number of proteins that show greater expression	*1.05%*	*10.5%*	*21%*
Metabolic pathways	41	37	38
Biosynthesis of secondary metabolites	18	14	12
Microbial metabolism in diverse environments	14	14	8
Carbon metabolism	11	6	4
Biosynthesis of cofactors	10	6	13
Oxidative phosphorylation	7	5	0
Glycolysis / Gluconeogenesis	6	0	0
Alanine, aspartate and glutamate metabolism	6	0	0
Biosynthesis of amino acids	6	9	5
Pyruvate metabolism	6	0	3
Butanoate metabolism	5	2	2
Citrate cycle (TCA cycle)	5	0	0
Two-component system	5	0	0
Methane metabolism	5	2	0
Propanoate metabolism	4	0	0
Purine metabolism	4	2	2
Porphyrin metabolism	4	0	9
Aminoacyl-tRNA biosynthesis	3	0	0
Sulphur relay system	3	2	0
Glyoxylate and dicarboxylate metabolism	3	0	3
Nucleotide metabolism	3	0	3
Galactose metabolism	3	0	0
Starch and sucrose metabolism	3	0	0
Ribosome	0	14	5
Sulphur metabolism	0	8	0
ABC transporters	0	7	0
Cysteine and methionine metabolism	0	5	4
Inositol phosphate metabolism	0	3	0
Arginine and proline metabolism	0	2	2
Ubiquinone and other terpenoid-quinone biosynthesis	0	2	0
Phenylalanine; tyrosine and tryptophane biosynthesis	0	2	0
Glycine, serine and threonine metabolism	0	2	0
Pentose phosphate pathway	0	2	0

Pyrimidine metabolism	0	2	2
Thiamine metabolism	0	2	0
Phenylalanine metabolism	0	0	3
Benzoate degradation	0	0	2
Pantothenate and CoA biosynthesis	0	0	2
Lysine degradation	0	0	2
Quorum sensing	0	0	2
Valine, leucine and isoleucine degradation	0	0	2
Nicotinate and nicotinamide metabolism	0	0	2