Developing an oxygen sensitive iron sulphur cluster expression system based on proteins from anaerobic protozoa

A Thesis Submitted to the University of Kent for the degree of Master of Research in Microbiology

October 2024

James Skerritt

The School of Biosciences

Declaration

No part of this thesis has been submitted in support of any other application for a degree or qualification of the University of Kent or any other University or institution of learning. I can confirm that all work submitted is my own.

James Skerritt

October 2024

Acknowledgements

Firstly, I would like to thank my supervisors, Dr Tobias von der Haar and Dr Anastasios Tsaousis, for their support throughout my project. All of the members of the Kent Fungal Group were extremely helpful, and I could not have completed this project without them. I would like to specifically mention Charlotte Bilsby, as she has always been there to aid me whenever I have needed guidance. Finally, I would like to thank all of my family and friends for supporting me this past year.

Table of Contents:

Abstract:

Cell factories are widely used to produce microbial metabolites on an industrial scale, offering a practical alternative to traditional production methods. However, some enzymes, particularly iron–sulphur (Fe-S) enzymes, remain difficult to work with due to their sensitivity to oxygen. These enzymes are crucial in many metabolic processes, but the instability of their Fe-S clusters in aerobic conditions limits their use in engineered systems.

This project explored whether the Fe-S cluster assembly protein SufCB, from the protozoan *Blastocystis* Nand II, could be used in *Saccharomyces cerevisiae* to improve Fe-S enzyme function. Specifically, we investigated whether SufCB could interact with the yeast protein Nbp35 in a similar way to the native protein Cfd1, potentially forming a complex involved in the cytosolic Fe-S assembly (CIA) pathway.

To do this, we used bacterial culturing, molecular cloning (via Gibson assembly), and protein extraction techniques. Western blotting was used to detect expression and investigate possible interactions. While the interaction between SufCB and Nbp35 could not be fully confirmed, the results suggest that SufCB may play a role in the CIA pathway, and that further work is needed to explore this possibility in more detail.

Introduction:

1.1. Aims and Objectives:

The aim of this project is to improve the ability of *Saccharomyces cerevisiae* to express and maintain the function of iron–sulphur (Fe-S) proteins by introducing components of the Fe-S cluster assembly machinery from the anaerobic protozoa, Blastocystis. This project will also address a key challenge in synthetic biology, where the heterologous expression of eukaryotic Fe-S proteins is often limited by environmental sensitivity, particularly to oxygen.

To achieve this, the project will begin by identifying Fe-S cluster assembly proteins from anaerobic protozoa that could potentially complement or enhance the existing yeast Fe-S maturation system. These selected genes will then be introduced into bacterial and yeast systems using molecular cloning techniques, with the intention of enabling functional expression.

Also, this project will aim to investigate whether the protozoan protein SufCB can interact with the native yeast protein Nbp35, in a role similar to that of Cfd1 in the cytosolic Fe-S assembly (CIA) pathway. Protein expression and interaction will be examined using methods such as western blotting, to assess whether the introduced components can form part of a functional heterocomplex.

The final objective of this project would involve the modified system being tested using selected model Fe-S proteins to evaluate any improvements in expression or activity. These findings are expected to inform future strategies for developing yeast strains with improved tolerance to oxygen-sensitive conditions and enhanced capacity for the production of Fe-S-dependent enzymes in biotechnological applications such as bioproduction and bioremediation.

1.2. Cell Factories:

Throughout history, we have been using microorganisms to produce fermented foods and beverages. More recently, microbes have been used to produce chemicals for various applications (Nielsen and Keasling, 2016). One way we use these microbes is in cell factories. Cell factories are engineered biological systems that harness the metabolic capabilities of living cells to produce desired products (Stephanopoulos and Aristidou, 1998). Cell factories provide a scalable alternative to the traditional method of synthesising chemicals, with applications ranging from simple materials to complex structures, such as monoclonal antibodies (Lee and Kim, 2015). The microbes that make up the cell factories are genetically engineered to enhance their efficiency when synthesising specific compounds. This often will involve the introduction of new metabolic pathways or the optimisation of existing ones (Wang B et al., 2017). These changes to the microorganisms often involve the recombination of DNA. This is why we refer to these microorganisms as recombinant (Pray, 2008). In some cell factories, the recombinant expression of one open reading frame (ORF) is enough, whereas some cell factories require whole biosynthetic pathways (Cho et al., 2022).

1.3. *S. cerevisiae* as a cell factory:

When producing cell factories, manufacturers select biological platforms with low cost, high productivity, and proper post-translational modifications. One such, preferred cell factory is the yeast *Saccharomyces cerevisiae*, as it meets these requirements (Huang et al., 2014). S. cerevisiae has other benefits: cheap and quick growth and maintenance, a fully sequenced genome, tolerance of high sugar conditions, and a "make-accumulate-consume" life strategy that eliminates competition (Goffeau et al., 1996; Parapouli et al., 2020; Hagman et al., 2013). Despite the advantages of using *S. cerevisiae* cell factories, recent studies show that *S. cerevisiae* metabolites make up 18.5% of all protein-based recombinant biopharmaceuticals, with mammalian cell culture and *Escherichia coli* metabolites making up a much larger 39% and 29.8%, respectively (Ferrer-Miralles et al., 2009; Nielsen, 2013).

1.4. Iron sulphur enzymes:

For cell factories to produce products effectively, they have many requirements that must be met. This includes the growth of the cell factories under precise environmental conditions, which some have been trying to work around (Wang et al., 2024). This requirement of precise environmental conditions also applies to cell factories that contain iron-sulphur enzymes. These enzymes contain an iron-sulphur cluster that is destroyed upon reacting with oxygen. Because of this, cells containing iron-sulphur clusters can only produce their desired product when under anaerobic conditions (Raj et al., 2018). As well as this, in yeast, it was shown that the maturation of iron-sulphur proteins in the cytosol requires the presence of glutathione even if the environment is anaerobic (Sipos et al., 2002). It was concluded that the glutathione protects the iron-sulphur proteins from oxidative stress (Ozer et al., 2015). Creating these environmental conditions and sourcing any extra biologicals could be costly and may require extra work.

Iron-sulphur enzymes are a group of enzymes that bind to specific co-factors, known as iron-sulphur clusters, they come under the grouping of metalloproteins (Wang et al., 2019). Metalloproteins make up almost half of all known proteins in biology. They are involved in the catalysis of many extremely important interactions, including photosynthesis, respiration, and nitrogen fixation (Lu et al., 2009). Iron-sulphur enzymes are extremely important metalloproteins that are utilised in many scenarios, including electron transport, gene expression and regulation, and DNA replication and repair (Vallières et al., 2024). In the case of both respiration and photosynthesis, the iron-sulphur enzymes are known to use chains of iron-sulphur clusters to generate ATP (Bai et al., 2018). Iron sulphur enzymes are essential for all these processes; therefore, the production of iron sulphur enzymes is extremely important. A deficiency of iron sulphur clusters can lead to several diseases. This includes diseases such as Friedreich ataxia, inherited human sideroblastic anaemia, and even myopathy, exercise-induced acidosis, and myoglobinuria in certain conditions (Rouault and Tong, 2008).

Iron-sulphur clusters (Fe-S clusters) were discovered around 60 years ago in the electron carrier protein class, ferredoxins (Mortenson et al., 1962). 20 years later, it was found that iron-sulphur clusters were not just present in ferredoxins; they were also present in the aconitase protein (Johnson M. K., 1998). Aconitase catalyses reversible isomerisation in carbon metabolism, and its iron-sulphur cluster aids in this reaction (Frazzon et al., 2003). Most recently, over 100 different iron-sulphur cluster containing proteins have been described and are extremely diverse in their roles (Wächtershäuser G., 1992). Iron-sulphur clusters continue to be important in three major processes that are required to sustain life: respiration, photosynthesis, and nitrogen fixation (Frazzon et al., 2003).

The most common iron-sulphur clusters that are found in nature are cubane [4Fe-4S] clusters or rhombic [2Fe-2S] clusters (Figure 1, Figure 2). These are most often found connected to their proteins by cysteine ligands (Frazzon et al., 2003). Though iron-sulphur clusters like this are most common, there is a wide variety of iron-sulphur cluster forms that attach to their proteins in different ways. There are lots of variations in the spacing, environment, and type of iron-sulphur clusters found in different iron-sulphur proteins. This heavily influences the wide range of chemical and electrical properties that are associated with iron-sulphur clusters (Frazzon et al., 2003).

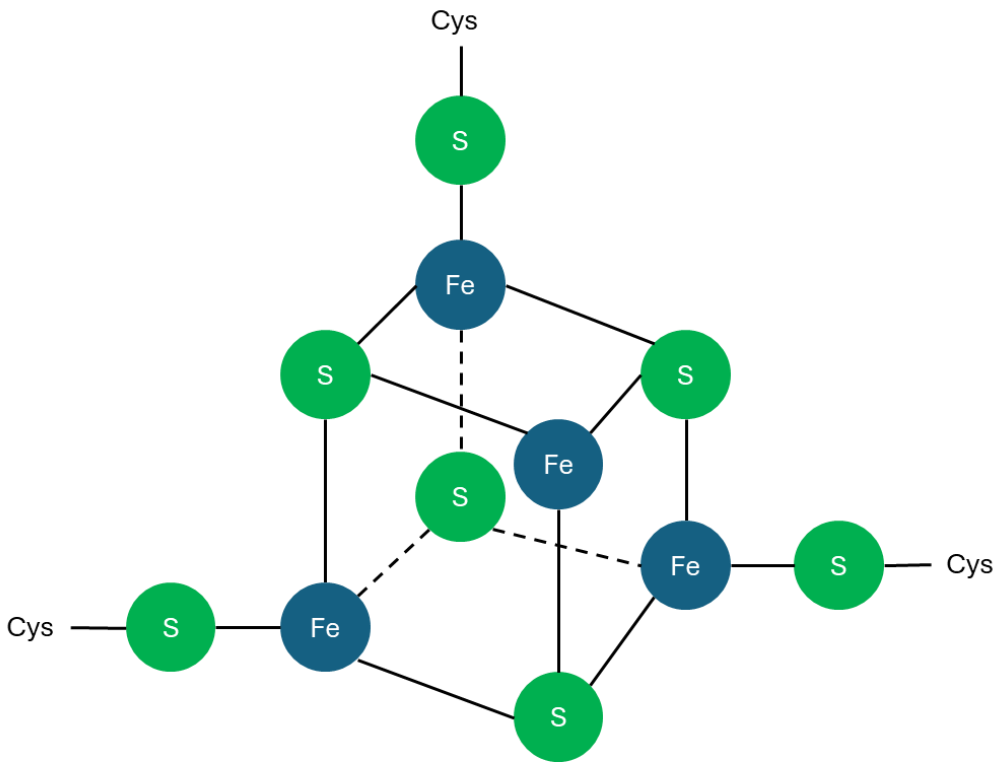

Figure 1: Iron-sulphur cluster cubane [4Fe-4S] form

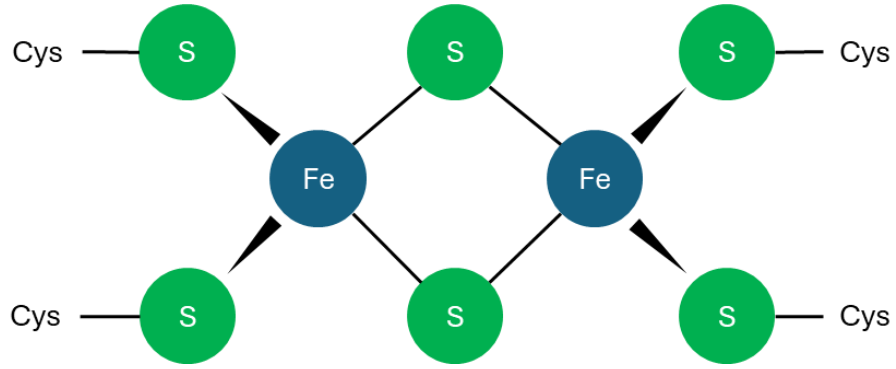

Figure 2: Iron-sulphur cluster rhombic [2Fe-2S] form

1.5. Pathways of iron-sulphur cluster formation:

The production of these iron sulphur enzymes and the iron sulphur clusters inside them is extremely controlled by extremely important systems. The formation and maturation of iron-sulphur clusters involves extremely specific pathways. In both prokaryotes and eukaryotes, there are three major systems that are responsible for managing these pathways (Lill R., 2020). The Iron-Sulphur Cluster (ISC) system, the Cytosolic Iron-Sulphur Protein Assembly (CIA) system, and the Sulphur Utilisation Factor (SUF) system.

This iron-sulphur cluster (ISC) system is a highly coordinated process that requires specialised cellular machinery (Saha et al., 2018). The ISC system is the most conserved pathway for the assembly of iron-sulphur clusters in bacteria and the mitochondria of eukaryotes. The ISC system is responsible for the initial synthesis of iron-sulphur clusters, which occurs in the mitochondrial matrix (Lill R., 2009). There are many different key enzymes that are involved in this system. Some examples include IscS, which provides sulphur and transfers it onto scaffold proteins; IscU, an example of one such scaffold protein that assembles the iron and sulphur atoms into iron-sulphur clusters with the aid of the other enzymes; and Ferredoxin, an electron donor that transfers electrons between iron and sulphur (Johnson et al., 2005). Together, these and other enzymes work to produce iron-sulphur clusters. This system works in tandem with the CIA system to produce iron-sulphur proteins.

The Cytosolic Iron-Sulphur Protein Assembly (CIA) system relies on the mitochondrially produced iron-sulphur clusters from the ISC system to produce cytosolic proteins that are involved with translation, DNA repair, and amino acid synthesis (Alhebshi et al., 2012; Bedekovics et al., 2011). Some extremely important enzymes in the CIA system include Nbp35 and Cfd1. These proteins are scaffold proteins that bind to iron-sulphur clusters and aid in their transfer to target proteins (Stehling et al., 2018). It is known that in some organisms, such as *A. thaliana*, the Nbp35 works alone, whereas in *S. cerevisiae*, it must form a complex with Cfd1 (Bych et al., 2008). The CIA

system in *S. cerevisiae* utilises a main scaffold complex (composed of Nbp35 and Cfd1) to assemble iron-sulphur clusters and then a second scaffold protein (composed of Cia1, Cia2, and Met18) to mature and transfer the iron-sulphur clusters onto target proteins (Netz et al., 2007). It has been proposed that this main oligomeric complex, composed of Nbp35 and Cfd1, has its roles separated between the two different proteins. It is suggested that Nbp35 is the main site of iron-sulphur cluster synthesis and that Cfd1 is used to assemble and transfer the clusters (Pallesen et al., 2013). Nbp35 and Cfd1 are theorised to have originated from a genome duplication even as both share a 47.3% sequence identity (Egan R., 2019). The Nbp35-Cfd1 complex is formed with two of each protein forming a tetramer that is bridged by [4Fe-4S] clusters connected by a shared C-terminal CPxC motif and two phenylalanine amino acids. These were presumed to be π-stacking to help strengthen the bonds between the proteins. (Stehling et al., 2018, Pallesen et al., 2013, Egan R., 2019).

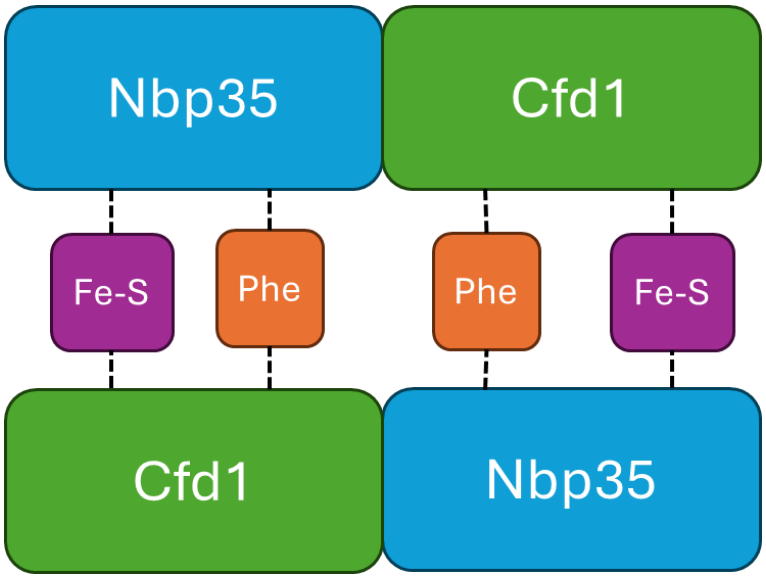

Figure 3: Nbp35-Cfd1 complex with both the iron-sulphur cluster (Fe-S) and phenylalanine (Phe) bridges.

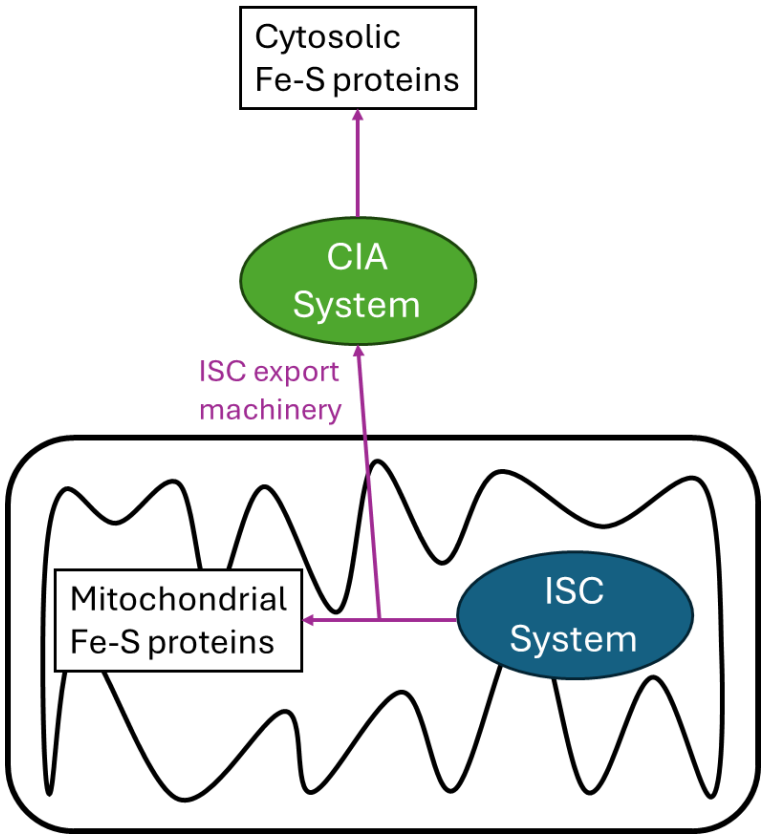

Figure 4: Diagram showing the locations of the ISC and CIA systems in iron-sulphur cluster production in the mitochondria. Purple arrows show the transport of iron sulphur clusters via different enzymes.

The Sulphur Utilisation Factor (SUF) system is an iron-sulphur cluster production pathway that is mainly found in bacteria and archaea. This system is often utilised under stressful conditions. This includes iron limitation, oxidative stress, and any times the ISC system is not functioning correctly (Outten et al., 2003). It is believed that SUF system has been an ancient method of assembling iron-sulphur clusters that was created in response to certain environmental challenges, such as iron limitation and oxidative stress (Boyd et al., 2014).

There are many enzymes that are all involved in the SUF system, some of which are similar to those present in the ISC system. One such enzyme is SufS, a cysteine desulphurase, similar to IscS, which supplies sulphur to the iron-sulphur cluster assembly. Our research involves the enzymes SufC and SufB. SufC is an ATPase that provides energy to the complex it forms; in *E.coli* it forms a complex with SufB and SufD. However, in *Blastocystis,* it forms a complex with only SufB. The SufB section of these complexes is a scaffold protein that acts as a platform where iron and sulphur ions are utilised to produce an iron-sulphur cluster before being transferred to target proteins (Outten et al., 2003). It is believed that Blastocystis has gained this sufCB gene via lateral gene transfer from a relative of the methanomicrobials (Tsaousis et al., 2012).

In normal SUF systems, for example, those found in Escherichia coli and other prokaryotes, SufC and SufB do function together. However, unlike the tightly integrated complex of sufCB, these two proteins usually function in a more flexible and modular assembly system. This SufCB complex in Blastocystis has evolved to efficiently operate in anaerobic environments. The more rigid structure of the SufCB complex ensures that the iron-sulphur clusters are assembled without being damaged by oxygen presence (Outten et al., 2003). This adaptation has allowed *Blastocystis* to become adapted to a more parasitic lifestyle to synthesise iron-sulphur clusters under oxygen stress (Tsaousis et al., 2012).

1.6. Iron-sulphur cluster assembly in different organisms:

The machinery that produces iron-sulphur clusters differs from organism to organism. As well as this, the method in which the iron-sulphur clusters are protected from oxidation also differs. This is required due to the reaction that occurs between oxidants and the cluster. The loss of an electron destabilises the cluster, thus causing it to release the catalytic iron atom and converting the cluster from, for example, [4Fe-4S] to a $[3Fe\text{-}4S]^{1+}$ form (Djaman et al., 2004).

In eukaryotic organisms, such as *Saccharomyces cerevisiae,* iron-sulphur clusters are produced through the ISC assembly pathway present in the mitochondria or through the production of the activated sulphur substrate used by the iron-sulphur cluster assembly (CIA) pathway present in the cytosol (Hinton et al., 2022). The ISC assembly pathway avoids the oxidation of iron-sulphur clusters by being present only in the mitochondria, which provides a reducing environment (Lill & Muhlenhoff., 2006).

Prokaryotes, such as *Escherichia coli*, also have two systems for the production of iron sulphur clusters. Much like eukaryotes, they also contain an ISC assembly system and a CIA system; however, they also contain an ancient bacterial sulphur formation (SUF) pathway. The SUF pathway acts as a backup pathway that is used in emergency conditions of iron limitation or oxidative stress (Blahut et al., 2020). Despite the suggestion that the natural role of the SUF proteins is to catalyse the repair of oxidised iron-sulphur clusters, it was observed that after exposure to superoxide or hydrogen peroxide, SUF mutant cells repaired the clusters at the same rate as wild-type cells (Djaman et al., 2004).

Much like prokaryotes, the protozoan parasite Blastocystis contains the ISC CIA and SUF systems. However, Blastocystis expresses a fused version of the SufC and SufB proteins found in prokaryotes (Tsaousis et al., 2012). This fusion, named SufCB, is upregulated when under oxygen stress, thus allowing Blastocystis to synthesise iron-sulphur clusters under environmental conditions of high oxygen (Tsaousis et al., 2012). This has allowed Blastocystis to adapt to its

parasitic lifestyle. Comparing yeast to Blastocystis, it was observed that Blastocystis does not produce the Cfd1 protein, leading to the hypothesis that sufCB has replaced the function of Cfd1 (Tsaousis et al., 2014). When present in other organisms, Cfd1 is known to form a complex with nbp35, which is then involved in iron-sulphur protein assembly.

Though sufCB is a complex formed in Blastocystis, it has been translated into *Saccharomyces cerevisiae* (Egan., 2019). The results of this work showed that the addition of sufCB resulted in a range of recombinant and endogenous iron-sulphur enzymes being overexpressed. As well as this, Egan demonstrated several novel findings with regards to the role of Cfd1 and sufCB in yeast and Blastocystis, respectively.

Methods:

2.1. Health and Safety:

Any work that was related to genetic modification was performed under safe conditions. This involved the constant use of personal protective equipment. As well as the use of a category II laboratory that has been licensed for the handling and manipulation of genetically modified organisms. All procedures were risk and COSSH assessed where appropriate.

2.2. Bacterial culturing:

2.2.1. Preparation of media:

Any liquid culture that was used was prepared using the following recipes. It was routinely made in Duran bottles that were only used up to 4/5th of their nominal volume to avoid boiling over when autoclaving. When preparing solid cultures, the liquid culture had agar added and was then poured into petri dishes after autoclaving. The recipes for the different media that were used are listed below, first the solid ingredients were added, then followed by the double distilled water (ddH_2O) to reach the correct volume.

Lysogeny Broth		**Super Optimal Broth**	
Component	Concentration (Grams/Litre)	Component	Concentration (Grams/Litre)
BactoTryptone	10g/L	BactoTryptone	20g/L
Yeast Extract	5g/L	Yeast Extract	5g/L
Sodium Chloride	10g/L	Sodium Chloride	0.6g/L
Agar (if plating)	20g/L	Potassium Chloride	0.2g/L
		Magnesium Sulphate	2.4g/L

2.2.2. Usage and storage of media:

Before use, all media was sterilised in a bench top autoclave at 120 °C for around 40 minutes. Media was prepared for each experiment; sometimes excess media was created for use in other experiments. Media was stored at room temperature unless containing antibiotics, in which case, it was stored at 4°C. For creating solid media, the agar was added before autoclaving, then mixed and poured into plates shortly after removal from the autoclave. Poured plates were stored at 4°C after they had set.

2.2.3. Antibiotics usage and storage:

Antibiotic stocks were prepared in ddH_2O at 1000x concentrations and stored at 4°C for a maximum of 2 weeks. Antibiotics were added to media in the concentrations stated below.

Antibiotic	Concentration (µg/mL)	Supplier
Ampicillin	100	Melford/Roche
Chloramphenicol	25	Sigma

2.2.4. Glycerol Stocks:

Glycerol Stocks were frozen at -80°C in 1.5 ml cryotubes. They were composed of 50% sterile glycerol and 50% overnight cultures.

2.2.5. Culturing *E. coli*:

With regards to growing liquid *E.coli* samples all samples were placed in an oscillating incubator at 37°C overnight. For optimal growth these were 2ml samples in a 5ml glass test tube, leaving room in the test tube for air flow.

For growing *E. coli* cultures on solid media plates, a static 37°C incubator is used.

The *E.coli* is grown at 37°C to mimic the temperature of the human body, as this is the optimal temperature for *E.coli* growth. Occasionally, the E.coli was grown overnight at a lower temperature of 30°C to prevent protein misfolding and produce a more stable sample.

2.2.6. *E. coli* strains used:

Strain	Genotype	Source
NEB TOP10	*F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ$^-$*	KFG Communal Stock
BL21	*B F– ompT gal dcm lon hsdSB(rB– mB–) [malB+]K-12(λS)*	KFG Communal Stock
BL21-CodonPlus	*B F– ompT hsdS(rB– mB–) dcm+ Tetr gal endA Hte [argU ileY leuW Camr]*	KFG Communal Stock

2.2.7. Preparing seed stocks:

Bacterial seed stocks were produced by selecting single colonies from a stock plate of the bacteria. The single colony was scraped with the tip of a pipette and placed into 5 mL of either SOB or LB medium and incubated overnight in a 30°C shaking incubated. Afterwards the pipette tip was removed and 600 µL of glycerol was added to the culture to a final concentration of 15%. Next, 1 mL samples were aliquoted into cryotubes and then placed into a -80°C freezer for long term storage.

2.2.8. Making competent *E. coli* cells:

To prepare competent cells, SOB media was made and inoculated with the seed stock (0.4% v/v). In a 30°C incubator these cells were grown, until they reached an OD_{600} of 0.3. The CCMB80 buffer was then placed on ice. The cells were harvested in a benchtop centrifuge at 4°C, 3000 rpm for

10 minutes and resuspended in 32% v/v of CCMB80 on ice. This suspension was then incubated for 20 minutes on ice. Afterwards, the suspension was harvested at 4°C, 3000 rpm for 10 minutes. The supernatant was discarded before the pellet was resuspended in 4% v/v of ice-cold CCMB80 buffer. The OD_{600} of the resulting cell suspension was altered until 1.0 - 1.5, before being aliquoted into 1.5 mL microcentrifuge tubes and frozen at 80°C.

CCMB80 Buffer	Amount	Concentration
Components	1000 mL	
Potassium Acetate (KOAc)	0.98 g	10 mM
Calcium Chloride ($CaCl_2$*$2H_2O$)	11.8 g	80 mM
Manganese(II) Chloride ($MnCl_2$*$4H_2O$)	4 g	20 mM
Magnesium Chloride ($MgCl_2$*$6H_2O$)	2 g	10 mM
Glycerol	100 mL	10 %
ddH_2O	To 100 mL	

2.2.9. *E. coli* transformation:

2 µL (2% v/v) of miniprep-grade plasmid DNA, or around 9 µL (8% v/v) of DNA if from a ligation, was placed in an Eppendorf tube, on ice. An aliquot of competent cells was thawed on ice for around 15 minutes. To keep the competent cells from being warmed up, great care was taken to keep the cells on ice and to avoid the interference of body heat when possible. Once thawed, 100µL of competent cells were added to the plasmid DNA and incubated on ice for between 20 and 30 minutes. Then, the sample was heat shocked in a water bath at 42°C for 60 seconds and immediately placed on ice. 1ml of LB medium was added to the Eppendorf tube. This was incubated in a 37°C shaking incubator for 60 minutes. When using miniprep DNA the 100 µL of

the cells are spread, with a glass spreader, onto LB medium plates that have been treated with the appropriate antibiotics. When using ligation DNA, the cells were spun in a bench top centrifuge at 3000rpm for 5 minutes. Most of the supernatant was poured off. The pellet was suspended in the remaining supernatant and then spread onto an LB medium plate with a glass spreader.

2.2.10. Growing *E. coli*:

E. coli was grown in large batches for the production of proteins. TB medium flasks were produced at various volumes, depending on the amount of protein wanting to be produced. The phosphate buffer was made and autoclaved at 120 °C for around 40 minutes prior to adding to the other components that were also autoclaved separately. 2 ml of overnight cultures were then added to the flasks and incubated in a shaking incubator at 30°C for around 3 hours or until the OD_{600} is between 0.6 and 0.8. Flasks of grown *E. coli* were then induced with 20 µL 1M IPTG in a shaking incubator at 25°C overnight or for 3 hours at 30°C.

Terrific Broth Medium		**Phosphate Buffer**	
Components	Concentration (Grams/Litre)	Components	Concentration (grams/Litre)
Yeast Extract	24g/L	KH_2PO_4	23.14 g/L
Tryptone	20g/L	K_2HPO_4	125.41 g/L
Glycerol	4mL/L	ddH_2O	To 1000 mL
ddH_2O	900mL/L		
Phosphate Buffer	100mL/L		

2.2.11. Harvesting *E. coli* Cells:

After incubation with IPTG the *E. coli* cells were harvested in two different ways. Small amounts of the cells were collected by diluting the cells 1 in 10 with water, measuring the OD_{600} of the solution and x10 to calculate the real OD_{600}. 1÷the OD_{600} will give you the required concentration to produce an equal amount of cells for each sample. This amount of each sample was then centrifuged, had the supernatant removed and then frozen for analysis. Large amounts of cells (for purifying proteins) were made by halving the remaining volume of grown *E. coli* into centrifuge bottles and centrifuging them at 8000 rpm for 20 minutes at 5°C. The supernatant was removed, and the pellet was resuspended and placed into 50 mL falcon tubes and centrifuged at 4000 rpm for 20 minutes at 5°C. The supernatant was again discarded and frozen for purification later.

2.3. Molecular Cloning:

2.3.1. *E. coli* plasmid DNA extraction:

Plasmid DNA was purified using a Qiagen miniprep kit that utilises an overnight culture that was grown under antibiotic selection to filter the cells. All instructions included with the kit were followed as the manufacturer stated. The instructions that involve the usage of a centrifuged were followed rather than the vacuum manifold options.

2.3.2. TAE gel electrophoresis:

DNA was analysed by Tris Acetate EDTA gel electrophoresis. To visualize the DNA, ethidium bromide (EtBr) was used. The EtBr was added to the agarose in its molten state to achieve a concentration of 5 µg/mL. For safety, this was completed in a fume hood with gloves on. 1% weight/volume gels were made by dissolving 0.5 g of agarose in 50 mL of TAE buffer and then boiling this in the microwave. TAE buffer was produced as a 10x solution and then diluted with dd H_2O to produce a working 1x solution. The gels were cast in a 20ml gel bed and left to cool down and set on a level surface for around 60 minutes. Once the gel was set, it was placed into the gel tank. Next, the gel tank was filled with 1x TAE buffer. The samples were mixed with 6x loading buffer before being loaded into the gel. 5 µL of sample was loaded into each well. The gel is run at 90 V, 200 mA, for around 45 minutes. After use, TAE buffer was discarded as the pH changes.

Tris-Acetate-EDTA Buffer	Concentration of working solution (mM)
Components	
Tris	40mM
Acetic Acid	20mM
EDTA	1mM

2.3.3. Amplification of DNA using PCR:

To amplify DNA the polymerase chain reaction (PCR) was used. 1 µL (2.5% v/v) of plasmid DNA and 1 µL (2.5% v/v) of each primer (forward and reverse) was added to 20 µL (50% v/v) of the Promega GoTaq Green PCR master mix along with 17 µL (42.5% v/v) of dd H_2O, for a total volume of 40 µL. This mixture was mixed on a benchtop vortex and then briefly spun in a benchtop microcentrifuge and then placed in a thermocycler with the settings mentioned below for 29 cycles.

PCR Master Mix	
Component	Volume from stock (µL)
Taq Buffer	20
dNTPs	10
$MgCl_2$	5
Taq Polymerase	1

Stage	Temperature (°C)	Duration (hr:min:sec)
Preheated Lid	105°C	00:02:00
Annealing	95°C	00:01:00
Denaturing	51°C	00:00:30
Extension	73°C	00:02:00
Final Extension	73°C	00:05:00
Hold	16°C	indefinite

Primer Sequences		
Oligonucleotide Name	Sequence 5'-3'	Scale (nMol)
Duet2_sufCB_f	AttagttaagtataagaaggagatatacatATGGCGCAGCCTATTTTG AACGTTATTG	100nMol
sufCB_Duet2_r	cgcagcagcggtttctttaccagactcgagGAACCCCTCCTTGCTCG CCTGGTC	100nMol
Duet2_Cfd1_f	attagttaagtataagaaggagatatacatATGGAGGAACAGGAGATA GGCGTTC	100nMol
Cfd1_Duet2_r	cgcagcagcggtttctttaccagactcgagTTTTGGAGATTCTATCTGG GGTTGC	100nMol
Duet1_yNbp35_f	caccatcatcaccacagccaggatccgaattcgATGACTGAGATACTA CCACATGTAAACG	100nMol
yNbp35_Duet1_r	gttcgacttaagcattatgcggccgcaagcttCTATACATCCCCCACA GCATCTCGC	100nMol
Duet1_bNbp35_f	caccatcatcaccacagccaggatccgaattcgATGAGTTCAGTTCC GAATGCAAATCCATC	100nMol
bNbp35_Duet1_r	gttcgacttaagcattatgcggccgcaagcttCTACACAATCTTCGACA CGACGACAAAC	100nMol

2.3.4. Gel extraction:

Digests and pcr products were purified with a Qiagen Qiquick gel extraction kit. All steps were followed to manufacturer's standards.

2.3.5. Restriction enzyme digest:

First, a digest master mix was made. 12 µL (60% v/v) of this master mix was then added to 8 µL (40% v/v) of DNA. This mixture was then incubated in a shaking incubator at 37°C for 60 minutes. Before analysis on an agarose gel, the digests were centrifuged for 5 seconds to resuspend the condensate.

Restriction enzyme digest master mix	
Components	Concentration(% v/v)
Restriction enzyme	9%
10x Restriction Enzyme Buffer	18%
dd H_2O	73%

2.3.6. Plasmid constructs:

Throughout this project the pDuet plasmid construct was used. This plasmid was used as it allows for the co-expression of two genes.

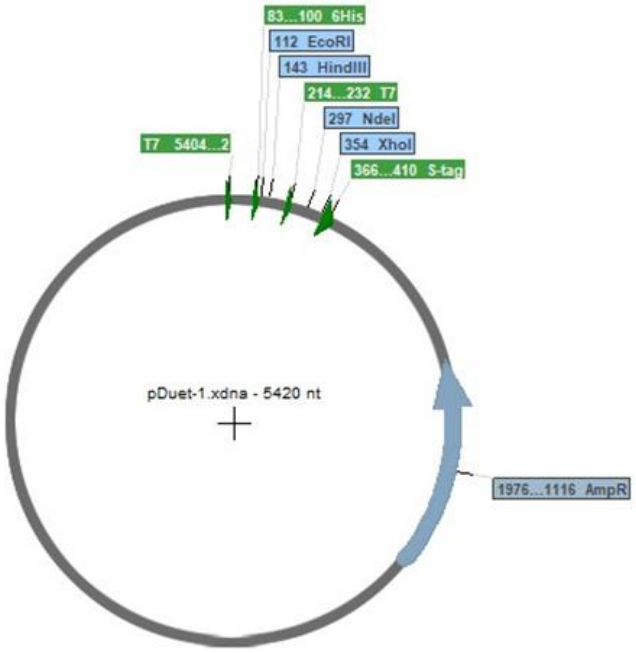

Figure 5: Empty pDuet plasmid map. Produced in Serial Cloner.

2.3.7. Gibson assembly method (ligation):

To start, a 5x isothermal buffer was made and mixed with other reagents to produce a Gibson assembly master mix that was stored at -20°C in 15 µL aliquots. The open pDuet plasmid, along with the nbp35/sufCB/cfd1 DNA template, were combined to an equimolar ratio to a volume of 5 µL. This mixture was then added to a 15 µL aliquot of Gibson master mix. This was then incubated in a 50°C water bath for 1 hour. Around 5 µL of this final mixture was used for a bacterial transformation.

5x Isothermal Buffer	
Components	Concentration (µL/mL)
50% PEG 8000	450
2M Tris-HCL pH 7.5	250
500mM $MgCl_2$	100
1M DTT (Dithiothreitol)	50
50mM NAD (Nicotinamide adenine dinucleotide)	100
100mM ATP, CTP, GTP & TTP (PCR grade)	10
ddH_2O	10

Gibson assembly master mix	
Components	Concentration (µL/mL)
5x isothermal buffer	160
1U/ µL T5 exonuclease	3.2
2U/ µL Phusion polymerase	10
40U/ µL Taq ligase	80
ddH_2O	346.8

2.3.8. DNA sequencing:

For Sanger sequencing 5 µL of miniprep grade DNA and 5 µL of sequencing primer, diluted to a concentration of 3 µM, were mixed on a benchtop vortex and sent to Eurofins genomics to be sequenced. The results were analysed using the BLAST (Camacho C. et al, 2009) and ClustalW (Thompson J.D. et al, 1994) programs.

2.4. Protein work:

2.4.1. Protein Extraction with B-PER:

To extract the protein from the bacteria without mechanical disruption, Bacterial Protein Extraction Reagent (B-PER) was used (thermos scientific). Initially regular B-PER was used. Sonication with a Soniprep 150 was used to remove the viscosity of the samples after the use of B-PER. The B-PER was later switched to B-PER complete, which included necessary lysozymes and DNase. This allowed the sonication step to be skipped as the added DNase removed the viscosity of the samples. 4 mL of B-PER per gram of sample was initially used. This was later increased to 10 mL per gram to aid in reducing the viscosity. After the addition of B-PER, the samples were incubated in an end over end mixer at room temperature for 15 minutes. Then these samples were mixed appropriately to form the samples aiming to be purified and further incubated in and end over end mixer at room temperature for another 10 minutes.

2.4.2. Native magnetic bead NiNTA purification:

Immediately after the protein extraction the proteins were purified from the samples via native magnetic bead NiNTA purification. First, 40 µL of Hispur™ NiNTA magnetic beads were added to 1.5 ml microcentrifuge tubes. Then, 160 µL of 1x equilibration buffer was added to the tubes. This was mixed with a vortex mixer for 10 seconds and the supernatant was removed via a magnetic stand. The protein samples were diluted with equilibration buffer at a ratio of 1:1. Next, 400 µL of the diluted protein extract was added to the microcentrifuge tubes. This was mixed with a vortex mixer for 10 seconds and incubated in an end over end sample mixer for 30 minutes. The supernatant was again removed with a magnetic stand. Then, 400 µL of wash buffer was added to the tubes. This was again mixed with a vortex mixer for 10 seconds and the supernatant was removed with a magnetic stand. These steps with the wash buffer were repeated once more. Finally, 25 µL of elution buffer was added to the tubes. This was then mixed with a vortex mixer for

15 seconds and incubated in an end over end sample mixer for 15 minutes. The supernatant was then removed with a magnetic stand and transferred to another microcentrifuge for freezing to analyse later.

Equilibration Buffer	
Components	Final Concentration
Sodium Phosphate (Na_3PO_4)	100 mM
Sodium Chloride (NaCl)	600 mM
Polysorbate 20 (Tween™-20 Surfact-Amps™ Detergent Solution)	0.05%
Imidazole ($C_3N_2H_4$)	30mM

Wash Buffer	
Components	Final Concentration
Sodium Phosphate (Na_3PO_4)	100 mM
Sodium Chloride (NaCl)	600 mM
Polysorbate 20 (Tween™-20 Surfact-Amps™ Detergent Solution)	0.05%
Imidazole ($C_3N_2H_4$)	50mM

Elution Buffer	
Components	Final Concentration
Sodium Phosphate (Na_3PO_4)	100 mM
Sodium Chloride (NaCl)	600 mM
Imidazole ($C_3N_2H_4$)	250mM

2.4.3. Protein separation with SDS-PAGE:

For separating and analysing proteins SDS-PAGE was used. Gels were either produced at a specific acrylamide % v/v or pre-cast Novex 4-20% gels (ThermoFisher) were used. Resolving gel was poured up to 75% of an empty cassette. Industrial methylated sprit was added on top to level the resolving gel and remove any bubbles. Once this was set, the stacking gel was then added to fill the cassette, and the gel comb was placed into the cassette. When producing the resolving and stacking gel, the APS and TEMED were added last and straight before pouring into the gel casing. Once the gel was set, the samples were loaded after being mixed with 4x sample buffer. This gel is then run in an SDS-PAGE gel tank filled with running buffer (MOPS running buffer for pre-cast Novex 4-20% gels) for around an hour (depends on the % v/v of the gel) at 160V. This gel was then stained overnight on a rotating platform with a Coomassie blue stain, and destained the next morning on a rotating platform.

Resolving Gel			
Components	Volume of components to produce 1 gels worth		
	7.5%	10%	12.5%
30% acrylamide:Bisacrylamide 29:1	2.7 mL	3.6 mL	4.5 mL
Lower Tris-Buffer	2.7 mL	2.7 mL	2.7 mL
ddH_2O	5.4 mL	4.5 mL	3.6 mL
40% ammoniumpersulphate (APS)	40 µL	40 µL	40 µL
TEMED (Tetramethylethylenediamine)	5 µL	5 µL	5 µL

Stacking Gel	
Components	Volume of components to produce 1 gels worth
30% acrylamide:Bisacrylamide 29:1	1.0 mL

4x Upper Tris-Buffer	1.75 mL
ddH_2O	4.2 mL
40% ammoniumpersulphate (APS)	40 µL
TEMED (Tetramethylethylenediamine)	5 µL

4x Lower Tris-Buffer	
Components	Concentration
Tris pH 8.8	1.5 M
SDS (Sodium Dodecyl Sulphate)	0.4%

4x Upper Tris-Buffer	
Components	Concentration
Tris pH 6.8	1.0 M
SDS (Sodium Dodecyl Sulphate)	0.4%

Running Buffer	
Components	Amount per litre
Tris-Base	3 g/L
Glycine	14.4 g/L
SDS (Sodium Dodecyl Sulphate)	1.5 g/L

MOPS running buffer	
Components	Amount per litre
MOPS (Morpholinepropanesulonic acid)	209.3 g/L

Tris-Base	121.14 g/L
SDS (Sodium Dodecyl Sulphate)	20 g/L
EDTA (Ethylenediaminetetraacetic acid)	7.6 g/L

4x sample buffer	
Components	Amounts to produce 10ml
Tris-HCl pH 6.8	2.5 mL
SDS (Sodium Dodecyl Sulphate)	1 g
Bromophenol Blue	0.8 mL
Glycerol	4 mL
β-mercaptoethanol	2 mL

Coomassie Stain	
Components	Concentration
Coomassie R250	0.1%
Acetic Acid	10%
Methanol	40%
ddH_2O	49.9%

Coomassie Destain	
Components	Concentration
Acetic Acid	10%
Methanol	20%
ddH_2O	70%

2.4.4. Protein analysis with western blotting:

For the analysis of protein content, western blotting was used with both His tagging and S-tagging. These different tags were used due to the different tags that are present in both nbp35 and sufCB/cfd1. After the SDS-PAGE separation the acrylamide gel was removed from the cassette and measured, rather than being stained. Replicating these measurements a rectangle of polyvinylidene fluoride (PVDF) membrane and two rectangles of blotting paper were cut with scissors. Nitrile gloves were worn during these processes to avoid the contamination of the proteins on the PVDF membrane. The gel was then placed onto a semi-dry power blotter in a sandwich composed of the different materials. First, one of the blotting paper rectangles was submerged in transfer buffer and placed onto the power blotter tray. Then, the PVDF membrane was submerged in >99% methanol and placed on top of the first blotting paper. Next, the unstained SDS-PAGE gel was placed on top of the membrane. Finally, the second piece of blotting paper was submerged in transfer buffer and place on top of the gel. The whole sandwich was rolled with a roller to remove any air bubbles. This was then closed into the power blotter and ran for 10 minutes at 25V and 1.3A. After the transfer had been completed, the PVDF membrane was removed and washed in 2.5% blocking buffer (2.5% skimmed milk in PBS tween) on a rotating platform at room temperature for 60 minutes. PBS tween was formed with Oxoid Phosphate buffered saline tablets, 10 tablets/L) After the blocking step, the PVDF membrane was washed 3 times with PBS tween for 5 minutes. The PVDF membrane was then incubated in the desired concentration of primary antibody overnight, on a rotating platform, in a cold room. After incubation in the primary antibody, the PVDF membrane was then washed in PBS tween, once for 5 minutes, and twice for 15 minutes. Then, the PVDF membrane was incubated in the required concentration of secondary antibody in blocking buffer on a rotating platform at room temperature for 60 minutes. The PVDF membrane was once again washed in PBS tween for 5 minutes three times. Finally, the washed PVDF membrane was developed in ECL.

2.4.5. Enhanced Chemiluminescence Detection (ECL):

ECL allows for the detection of antibodies conjugated with horse-radish peroxidase. The ECL final solution was composed of two separate solutions (S1 & S2) that were prepared beforehand. These progenitor solutions were wrapped in aluminium foil and stored at 4°C to protect the solutions from light. In a dark room, these solutions were mixed together on top of the PVDF membrane and incubated for a minute. The ECL signal was then analysed at varying exposure times.

S1	
Components	Volume
1M Tris pH 8.5	1 mL
90 mM Coumaric Acid in DMSO	44 µL
250 mM Luminol in DMSO	100 µL
ddH_2O	To 10 mL

S2	
Components	Volume
1M Tris pH 8.5	1 mL
30% hydrogen peroxide	6.8 µL
ddH_2O	To 10 mL

2.4.6. Analysis with gel Syngene G:box gel doc:

After incubation in ECL the PVDF membrane was placed into the tray of the Syngene G:box gel doc and analysed with accompanying Genesys software.

2.4.7. Analysis with Optimax 2010 film processor:

After incubation in ECL the PVDF membrane was placed into a light safe box in a sandwich of plastic sheets and taped down so that it would not move. Next, a sheet of film was placed on top of this sandwich. This was then shut and placed in a dark draw for varying exposure times to expose the membrane on the film. Afterwards, the film was removed and immediately placed into the optimax 2010 film processor.

2.4.8. Stripping PVDF membrane for reprobing:

The PVDF membrane was washed with PBS tween on a rotating platform for 5 minutes at room temperature, to remove any excess ECL. Then, the PVDF membrane was incubated in stripping buffer (pH 2.2) on a rotating platform for 10 minutes at room temperature. This buffer was discarded, and this step was repeated with fresh buffer. Next, the PVDF membrane was again washed in PBS tween on a rotating platform for 5 minutes at room temperature. This was discarded and replaced with fresh PBS tween four times. To begin the reprobing as instructed previously, the PVDF starts again from the blocking step.

Stripping Buffer	
Components	Final Concentrations
Tris pH 6.8	62.5mM
SDS	1%
2-Mercaptoethanol	0.8%
ddH_2O	To final volume.

Results:

Acquisition of genes that code for NBP35, SufCB & *CFD1*:

DNA encoding for Yeast *NBP35* and *CFD1*, and Blastocystis SufCB, had been amplified by PCR prior to the start of this project. As a part of this PCR, Yeast genomic DNA (gDNA) and Blastocystis complementary DNA (cDNA) were used as templates. The integrity of the stored PCR products was validated by gel electrophoresis. This allowed us to show that all products were of the expected sizes for the corresponding Open reading frames (ORF's). This experiment was completed prior to insertion of the genes into the pDuet plasmid to make sure that the correct genes were being inserted. From Figure 6, it is evident that all three genes are present in their respective samples. The expected sizes of each gene are as follows; CFD1 – 882 nucleotides, NBP35 – 984 nucleotides, SufCB – 2100 nucleotides. Both the CFD1 and NBP35 bands are shown at slightly higher than their expected weight at around 900 base pairs and 1100 base pairs respectively. There is a slightly heavier band present in the SufCB sample, around 7,000 base pairs. This is unlikely to be an issue due to the correct band weight of 2100 base pairs being a much higher concentration.

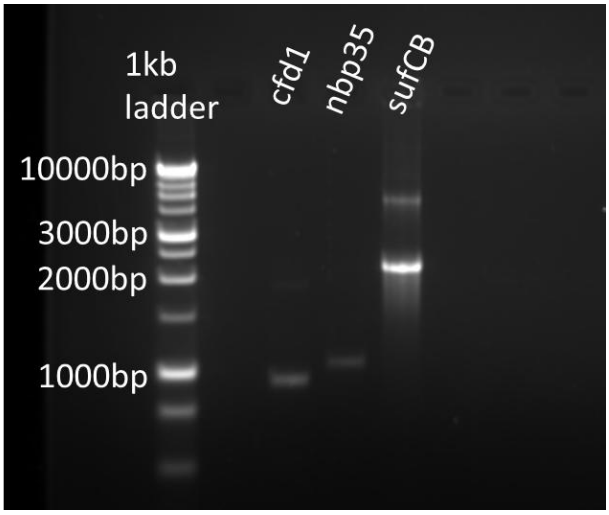

Figure 6: 1% agarose gel showing the presence of CFD1, NBP35 and SufCB. CFD1 & NBP35 bands are shown accurately at around 100bp. The SufCB band is shown around 7000bp, this is much higher than expected.

Insertion of SufCB and *CFD1* coding DNA into pDuet plasmid:

Having gathered the correct DNA samples and confirmed they are pure, both the SufCB and *CFD1* genes were inserted into the pDuet plasmid (*Figure 7*) using the Gibson assembly method. Initially the plan was to clone both genes simultaneously using PCR. The PCR primers would add complementary sequences adjacent to both, the NdeI and XhoI sites that are present in the SufCB and *CFD1* genes, and the EcoRI and HindIII sites from the *NBP35* gene, in pDuet. However, because there is an NdeI site in the *NBP35* gene, this had to be inserted last, after the cloning of both SufCB and *CFD1* genes. In addition, another two plasmids are produced that will not have *NBP35* inserted, so that it will only code for SufCB and *CFD1* on their own.

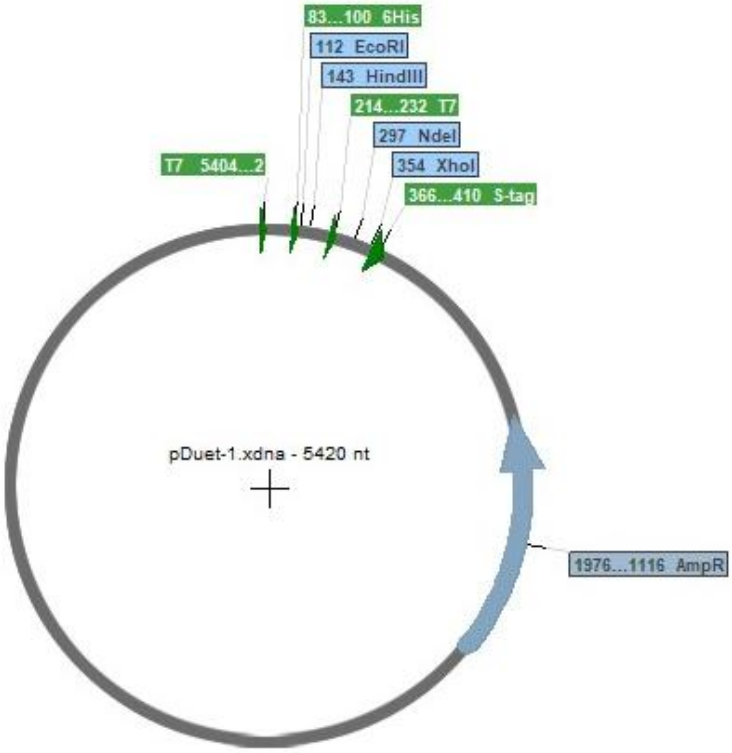

Figure 7: Empty pDuet plasmid map. Produced in Serial Cloner.

After the assembled pDuet plasmid was produced and transformed into competent *E. coli* cells, individual colonies were selected to produce overnight cultures. After these had grown, the DNA was re-isolated with a miniprep kit. Then, genes were cut out via restriction enzyme digests to check that the plasmid had taken up inserts of the correct size. For this, EcoRI and HindIII were used.

In Figure 8, there are bands representing the genes for SufCB, at around 2000 base pairs, in the first three SufCB samples. These show the presence of the SufCB coding gene as this gene is around 2100 base pairs long. As well as this, there were bands showing the presence of *CFD1*, at around 900 base pairs, in all four *CFD1* samples. Again, we know that these bands represent the *CFD1* coding gene because this gene is around 880 base pairs long. The heavier and brighter bands that are between 3000 and 10,000 base pairs show the presence of the rest of the pDuet plasmid that after digestion does no longer contain the SufCB and *CFD1* coding DNA. We know

this because this band is also present in the sample that only contained an empty pDuet plasmid. This pDuet plasmid was considered empty because it has not had the SufCB, *CFD1* or *NBP35* genes inserted into it.

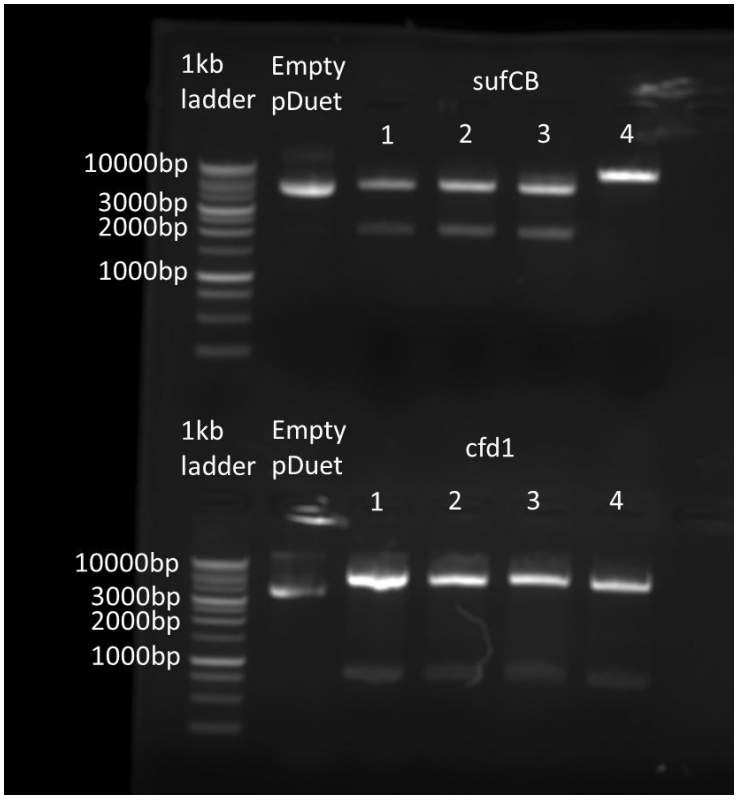

Figure 8: 1% agarose gel showing the presence of SufCB and CFD1 coding genes, after their insertion into the empty pDuet plasmid and replication of said plasmid in an overnight culture followed by the restriction digest using EcoRI and HindIII. SufCB bands are seen at around 2000bp (expected size = 2100bp). CFD1 bands are seen at around 900bp (expected size = 880bp).

Insertion of *NBP35* coding DNA into pDuet plasmids:

After the insertion of the SufCB and *CFD1* genes, the *NBP35* gene is inserted. Again, using the Gibson assembly method and the steps after it, as mentioned previously. This produces the plasmids shown in *Figure 9* & *Figure 10*.

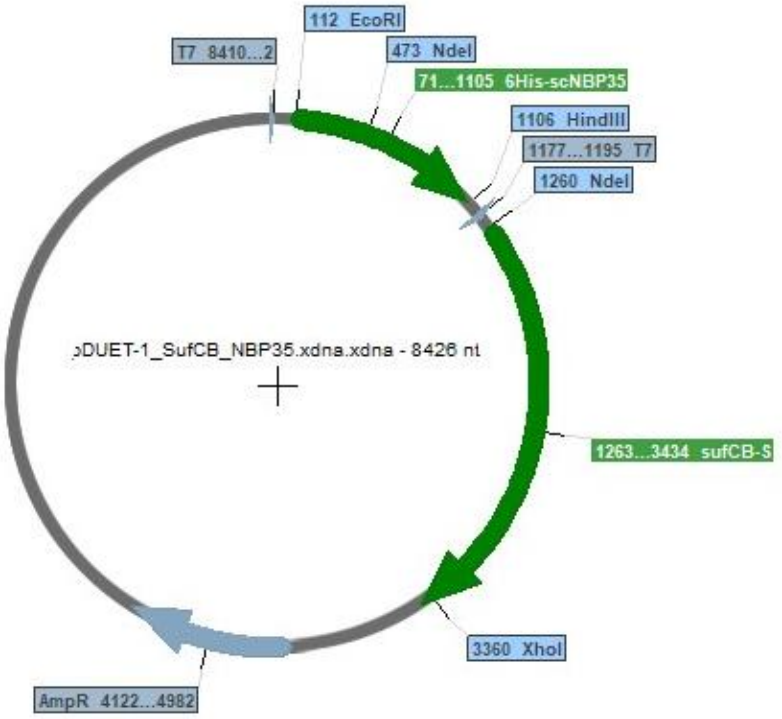

Figure 9: SufCB & NBP35 containing pDuet plasmid map. Produced in Serial Cloner.

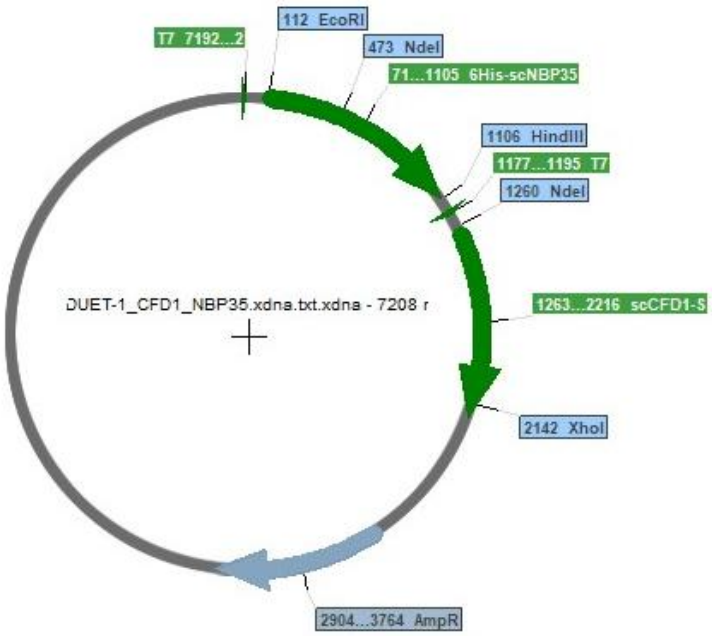

Figure 10: CFD1 & NBP35 containing pDuet plasmid map. Produced in Serial Cloner.

Figure 11 and Figure 12 show the presence of *NBP35* genes after their insertion into both the empty pDuet plasmid and the pDuet plasmids that had SufCB and *CFD1* inserted previously. To produce these samples, the plasmids were analysed by PCR with primers that were specific to *NBP35*.

In Figure 11, the first row shows three samples of pDuet containing only Nbp35. Samples in lanes 1 and 2 show a band at around 1000 base pairs which likely represents the presence of the *NBP35* coding gene which is around 984 base pairs long. However, the bottom row shows something different. For the pDuet plasmid that already contained the SufCB coding gene, no *NBP35* gene was present. Despite this, the *NBP35* coding gene can be seen in the pDuet plasmid that contained *CFD1* before adding *NBP35*. Specifically in lanes 1 and 2, represented by the bands around 1000 base pairs. Due to the lack of *NBP35* coding genes found in the SufCB & Nbp35 samples, we decided to attempt the insertion of *NBP35* into this plasmid again. In Figure 12 all three samples show a band at around 1000 base pairs, showing the presence of the *NBP35*

coding genes. With regards to both Figure 11 and Figure 12, the lighter bands below 250 base pairs are showing the presence of the dimer's that are formed from the primers used during PCR.

To confirm the integrity of the cloned sequences, samples were sent to Eurofins Genomics to be sanger sequenced.

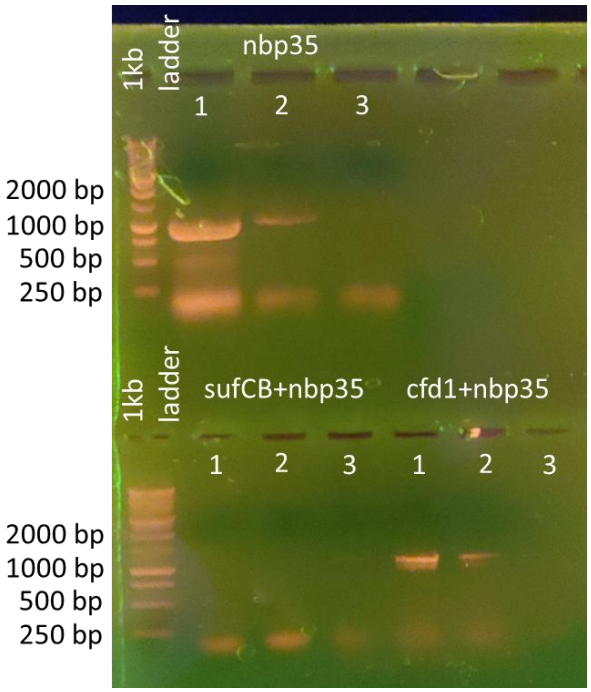

Figure 11: 1% agarose gel showing the presence of NBP35 coding genes, after Nbp35 coding DNA's insertion into pDuet plasmids that had SufCB & CFD1 inserted previously. Followed by replication of said plasmid in an overnight culture, restriction digestion, and PCR. The bands at 1000 base pairs show to presence of NBP35 coding genes. The top row contains the restriction digested pDuet with only Nbp35 coding genes inserted. The bottom row first three lanes containing restriction digested pDuet with both SufCB and NBP35 coding genes inserted. The latter 3 lanes contain restriction digested pDuet with both CFD1 and NBP35 coding genes inserted.

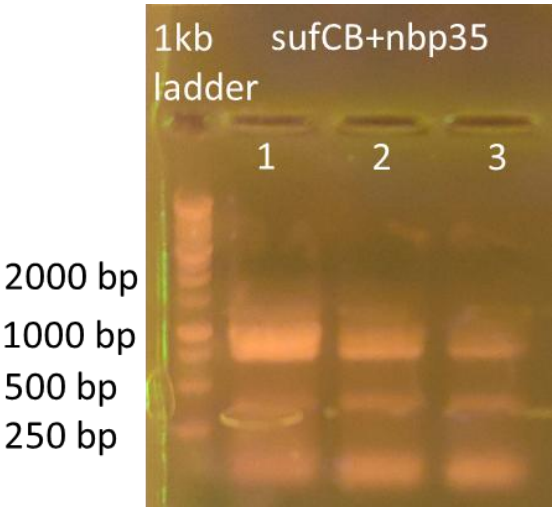

Figure 12: 1% agarose gel showing the presence of Nbp35 coding genes, after NBP35 coding DNA's insertion into pDuet plasmids that had SufCB inserted previously. Followed by replication of said plasmid in an overnight culture, restriction digestion, and PCR. Again, the three lanes containing restriction digested pDuet with both SufCB and NBP35 coding genes inserted.

pDuet plasmid transformation into BL21 *E. coli* expression strain:

Following the insertion of NBP35, SufCB, and CFD1, the pDuet plasmid was transformed into the BL21 strain of *E. coli*. After the transformation the cells were grown to an optical density 600 (oD600) of between 600 and 800. The cells were then induced with IPTG for 2 hours, and then harvested. The protein was then extracted from these harvested cells, using B-PER extraction reagent, and then seperated on an SDS-PAGE gel to be analysed.

In Figure 13, there are some bands around 80kDa (kilo Daltons). These are present in the lone SufCB sample, as well as the SufCB & Nbp35 sample. These bands likely indicate the presence of SufCB, as the weight of the SufCB proteins is 77kDa. As well as this, there is a thinner band at just over 43kDa, that is present in all the samples that should contain nbp35. This band likely represents the presence of Nbp35 as the weight of Nbp35 is 35.228kDa. This increased weight could be due to certain protein interactions or post translational modifications. On the other hand, there were no visable bands that were exclusively in the Cfd1 containing samples, that were the correct weight for Cfd1 (31.916kDa). All other bands on this gel match with those found in the empty pDuet sample, and are therefore not the Cfd1 protein that we would expect to see.

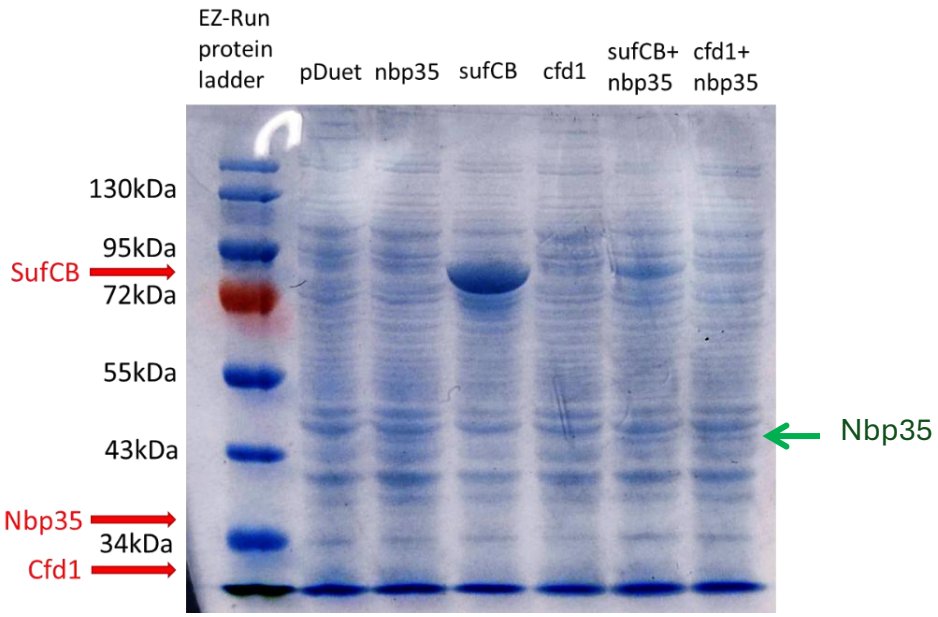

Figure 13: stained SDS-PAGE gel showing the proteins present in unpurified samples from the BL21 transformation of the pDuet plasmids produced. From left to right lanes show different samples, EZ-Run protein ladder, empty pDuet, lone nbp35, lone sufCB, lone cfd1, sufCB&nbp35, cfd1&nbp35. These samples were loaded as mentioned in the methods section. Red arrows show the expected weight of target proteins. Green arrow shows Nbp35 appear much higher than expected.

As the SDS-PAGE gel includes all the other proteins extracted from the cells, we decided to utilise western blotting to produce a clearer image of the presence of specific proteins. To target Nbp35, his-tag antibodies were used. To target both SufCB and Cfd1, s-tag antibodies were used. This required us to produce two separate western blots to show all three proteins. It is important to note that the Cfd1 sample in both of these blots had a small amount of Nbp35 in it. This is because instead of using a lone Cfd1 sample, the Cfd1 & Nbp35 sample was used.

Shown in Figure 14 is the His-tagged western blot with a 1-minute exposure time. There are bands present in the Nbp35, the Cfd1, the SufCB & Nbp35, and the Cfd1 & Nbp35 samples. This band shows the presence of Nbp35 at around 40kDa. This is close to the expected weight of Nbp35 at

35.228kDa. As expected, the band is slightly weaker in the Cfd1 sample compared to the others, due to the small amount of Nbp35 from using the Cfd1&Nbp35 sample.

Shown in Figure 15 is the S-tagged western blot also with a 1-minute exposure time. In this blot there are some bands at around 100kDa in the SufCB and SufCB & Nbp35 samples. These bands show the presence of SufCB, which has an expected weight of 77kDa. As well as this, there are two more bands lower down, at around 45kDa. These bands show the presence of Cfd1, which has an expected weight of 31.916kDa.

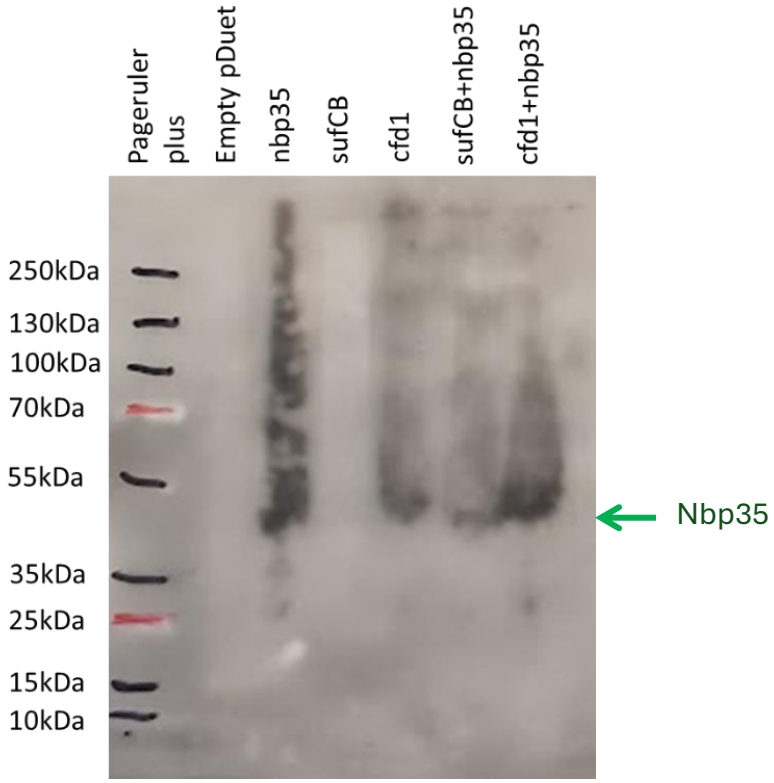

Figure 14: His-tagged Western blot showing the presence of his-tagged Nbp35 present in the BL21 transformation. From left to right lanes show different samples, pageruler plus protein ladder, empty pDuet, lone nbp35, lone sufCB, lone cfd1, sufCB&nbp35, cfd1&nbp35. These samples were loaded as mentioned in the methods section. Green arrow shows weight of nbp35 bands at around 40kDa.

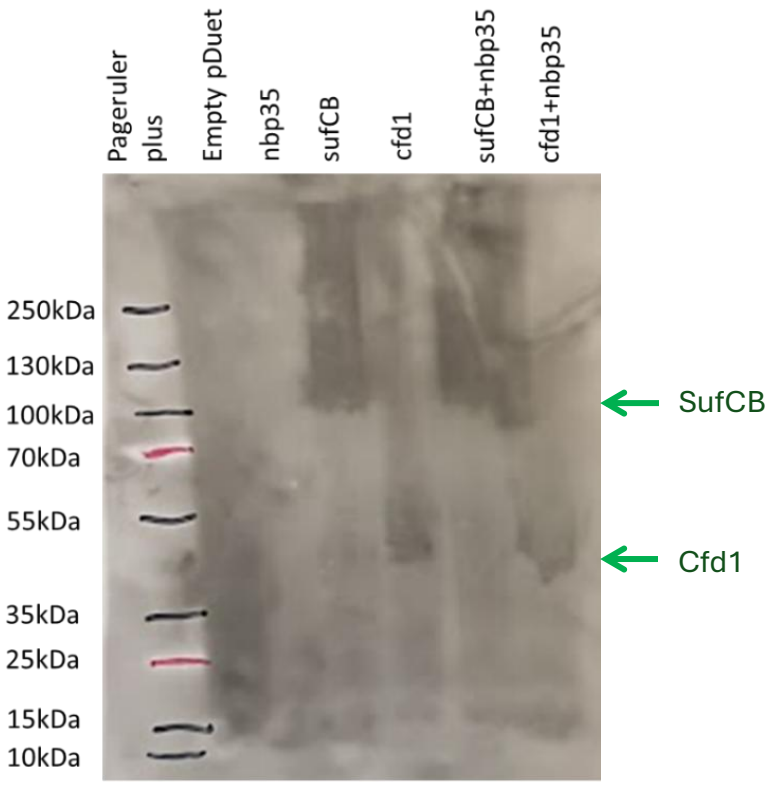

Figure 15: S-tagged Western blot showing the presence of s-tagged SufCB and Cfd1 in the BL21 transformation. From left to right lanes show different samples, pageruler plus protein ladder, empty pDuet, lone Nbp35, lone SufCB, lone Cfd1, SufCB&Nbp35, Cfd1&Nbp35. These samples were loaded as mentioned in the methods section. Green arrows show weights of SufCB and Cfd1 bands at 100kDa and 45kDa respectively.

Differing effectiveness of BL21 and BL21-Codonplus:

As well as BL21, we decided to transform the different plasmids into BL21-Codonplus. We decided to test whether this change would increase the expression of both Nbp35 and Cfd1. We expect this to improve expression of these proteins because codonplus cells overexpress certain tRNA's that are commonly used in non-bacterial species. Figure 16 and Figure 17 show SDS-PAGE gels comparing the protein content in the BL21 and BL21-Codonplus samples. Figure 16 shows a lower total protein content in BL21-Codonplus, whereas Figure 17 shows a lower total protein content in BL21. This could be due to human error during the normalising process.

In Figure 16, there is a clear, dark band around 80kDa in the SufCB lane of the BL21 samples. This indictes the presence of SufCB, which has an expected weight of 77kDa. This same band can just about be seen in Figure 17 in the BL21 SufCB lane. This band is not present in the BL21-Codonplus SufCB sample in either gel. Thus, showing that the BL21 is much better at producing SufCB than the BL21-Codonplus. Although, it should be noted that this band is not seen in any of the SufCB & Nbp35 samples. This lack of SufCB is seen in all the SufCB & Nbp35 samples, both BL21 and BL21-Codonplus.

In both Figure 16 and Figure 17 a strong band around 45kDa could show the presence of Nbp35 or even Cfd1. However, this is unlikely as this band is present in all samples on both gels. Either Nbp35 and Cfd1 are not shown in these samples, or, one of them is, and has contaminated the other samples.

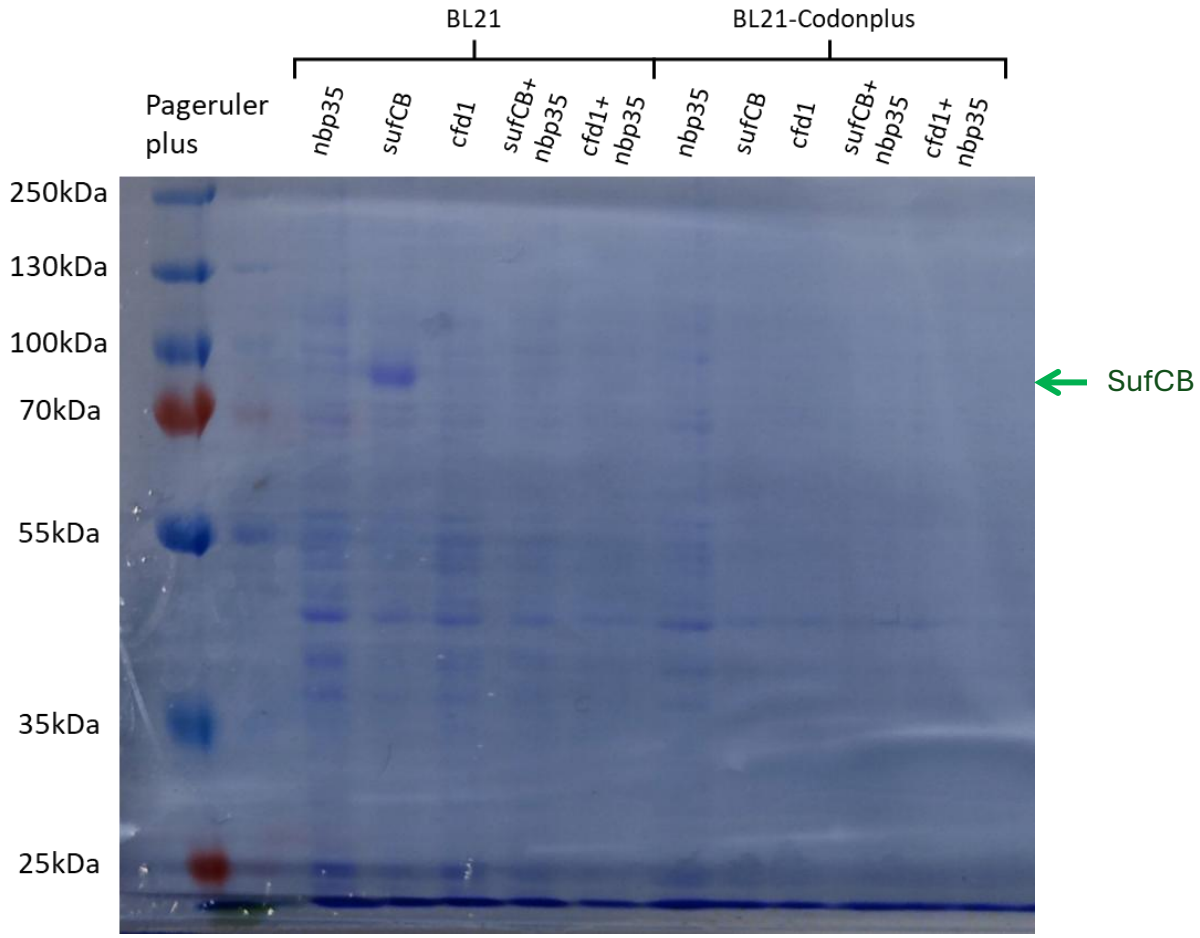

Figure 16: SDS-PAGE gel comparing the protein content of BL21 and BL21-Codonplus. From left to right lanes show different samples, pageruler plus protein ladder, the following samples for both BL21 and BL21-Codonplus respectively. Empty pDuet, lone Nbp35, lone SufCB, lone Cfd1, SufCB&Nbp35, Cfd1&Nbp35. These samples were loaded as mentioned in the methods section. Green arrow shows the weight of SufCB at around 80kDa, which matches the expected weight of SufCB, 77kDa.

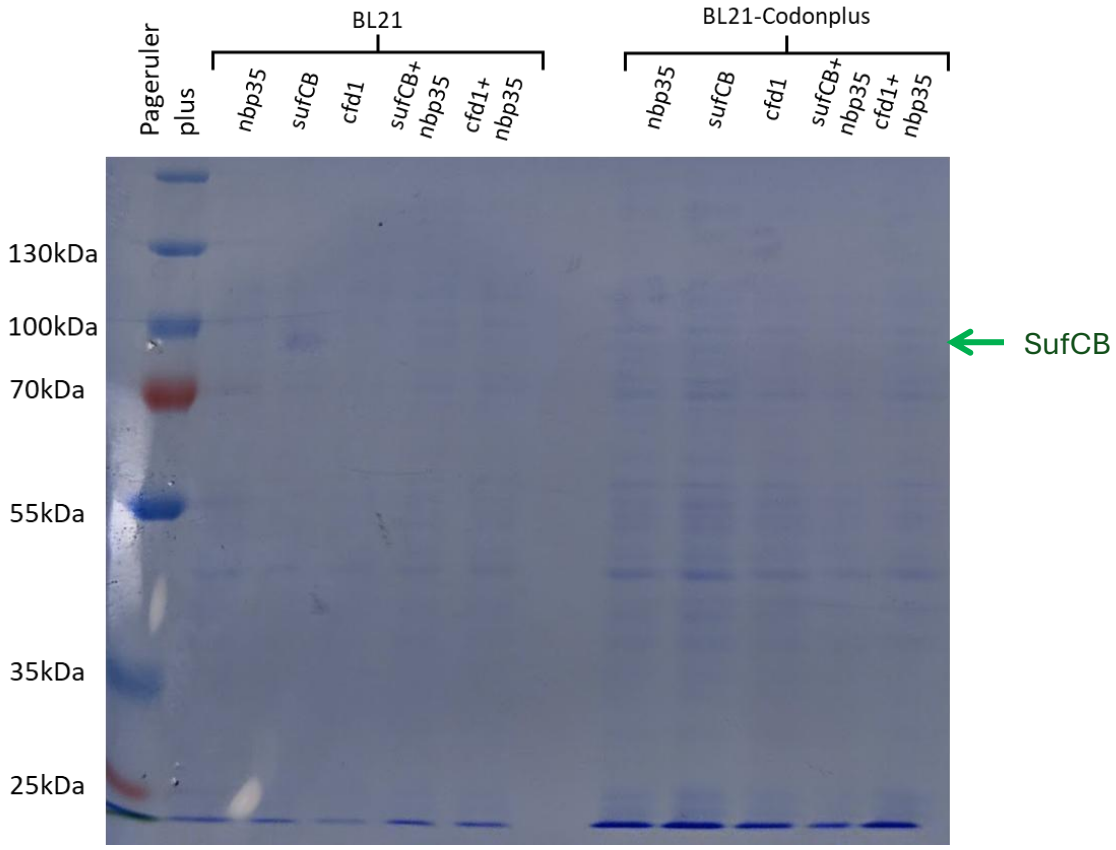

Figure 17: SDS-PAGE gel comparing the protein content of BL21 and BL21-Codonplus. From left to right lanes show different samples, pageruler plus protein ladder, the following samples for both BL21 and BL21-Codonplus respectively. Empty pDuet, lone Nbp35, lone SufCB, lone Cfd1, SufCB&Nbp35, Cfd1&Nbp35. These samples were loaded as mentioned in the methods section. Green arrow shows the weight of SufCB at around 80kDa, which matches the expected weight of SufCB, 77kDa.

After producing both Figure 16 and Figure 17, we decided to produce an S-tagged western blot showing the protein expression differences between BL21 and BL21-Codonplus. A western blot was used this time to show clearly our target proteins separate from the other proteins produced by the cells. S-tagging has reduced overall noise and has allowed us to analyse just the SufCB and Cfd1 proteins.

In Figure 18, There are clear bands present in both the BL21 and BL21-Codonplus samples. These bands that appear are present in the lone Cfd1 and Cfd1 & Nbp35 samples. Neither of them are close to the expected weight of Cfd1, 31.916kDa. The band that is present in the BL21 Cfd1 sample is around 250kDa according to the Pageruler plus ladder. The BL21-Codonplus Cfd1 sample is much lighter, at just above the 70kDa band from the Pageruler plus ladder. These changes in weight could be due to the formation of oligomers. This is still much higher than the actual weight of Cfd1. This extreme increase in weight has not occurred in the Cfd1 & Nbp35 samples of both BL21 and BL21-Codonplus. However, there is still an increase to around 70kDa. This could be showing the formation of a Cfd1 & Nbp35 complex in both strains. Aside from the presence of cfd1, there is a large mark in the SufCB & Nbp35 lane of the BL21 samples that is not present in the BL21-Codonplus sample. It is possible that there is SufCB present in this sample.

Overall, this shows that BL21 is better at producing SufCB and worse at producing Cfd1 when compared to BL21-Codonplus. The blot in *Figure 18*, paired with the gels from *Figure 16* and *Figure 17* shows that the co-expression of Nbp35 seems to have reduced the levels of SufCB but increased the levels of Cfd1 that is produced in the BL21 samples.

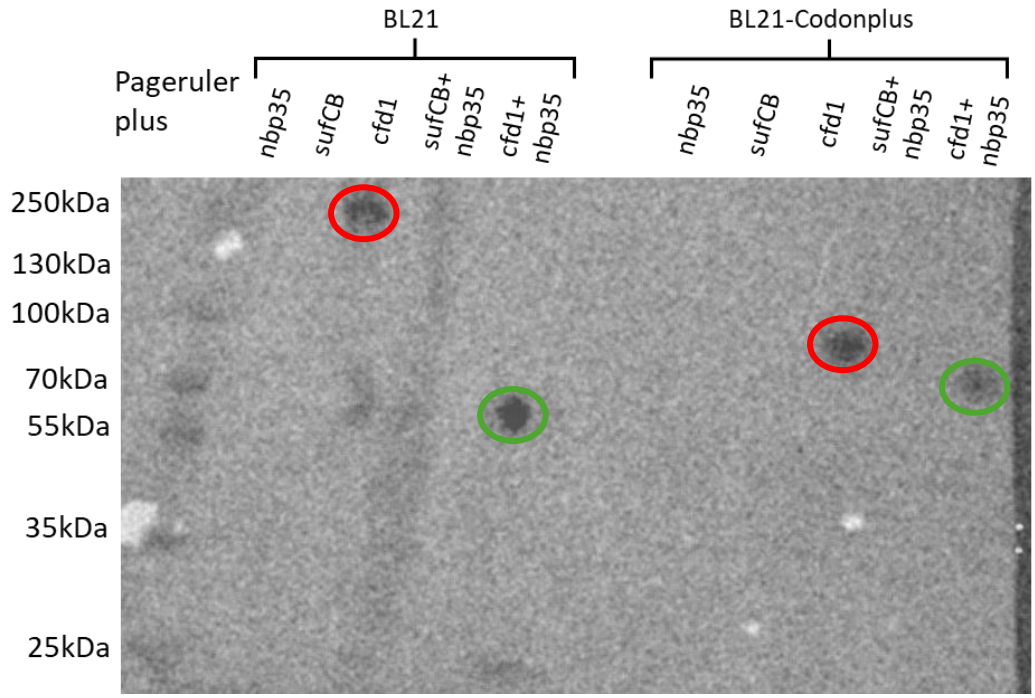

Figure 18: S-tagged Western Blot comparing the protein content of BL21 and BL21-Codonplus. From left to right lanes show different samples, pageruler plus protein ladder, the following samples for both BL21 and BL21-Codonplus respectively. Empty pDuet, lone Nbp35, lone SufCB, lone Cfd1, SufCB&Nbp35, Cfd1&Nbp35. These samples were loaded as mentioned in the methods section. Green circles show the Cfd1&Nbp35 samples in both BL21 and BL21-Codonplus at around 70kDa. Red circles show the lone Cfd1 samples in both BL21 and BL21-Codonplus at their weights of 250kDa and 90kDa respectively.

Immunoprecipitation of proteins from transformed BL21:

Next, we wanted to test whether SufCB and Cfd1 would be co-precipitated with Nbp35 if we were to isolate just the Nbp35 via its His-tag. This way, I would be testing if there are strong protein-protein interactions between Nbp35 and SufCB or Cfd1. By isolating only the Nbp35 protein, if SufCB or Cfd1 were present in the samples, it would show that they were 'pulled down' with the Nbp35 by bonding as expected.

Immunoprecipitation is a technique that involves the isolation of a specific protein from a complex mixture. The protein is isolated using specific antibodies that bind to a tag on the target protein, in our case, the His-tag, and S-tag sections on the proteins. Immunoprecipitation was used because it is an effective way of isolating the target proteins. After this immunoprecipitation the samples were ran on an SDS-PAGE gel. These gels were then used to produce western blots. Lone Cfd1 was not included on these gels.

Figure 19 shows the His-tagged western blot. This blot shows the presence of Nbp35 in the samples. There are bands in the Nbp35, SufCB & Nbp35, and Cfd1 & Nbp35 samples showing the presence of the Nbp35 protein at around 45kDa. The expected weight of the Nbp35 protein is 35.228kDa. This slight difference in weight can be due to small changes such as the addition of the His-tag and potentially some post translational modifications. It can therefore be concluded that these bands show the presence of Nbp35. These bands are not present in the empty pDuet or the SufCB samples as expected. These results are as expected, and prove that the Nbp35 was isolated correctly in the samples that it was present in.

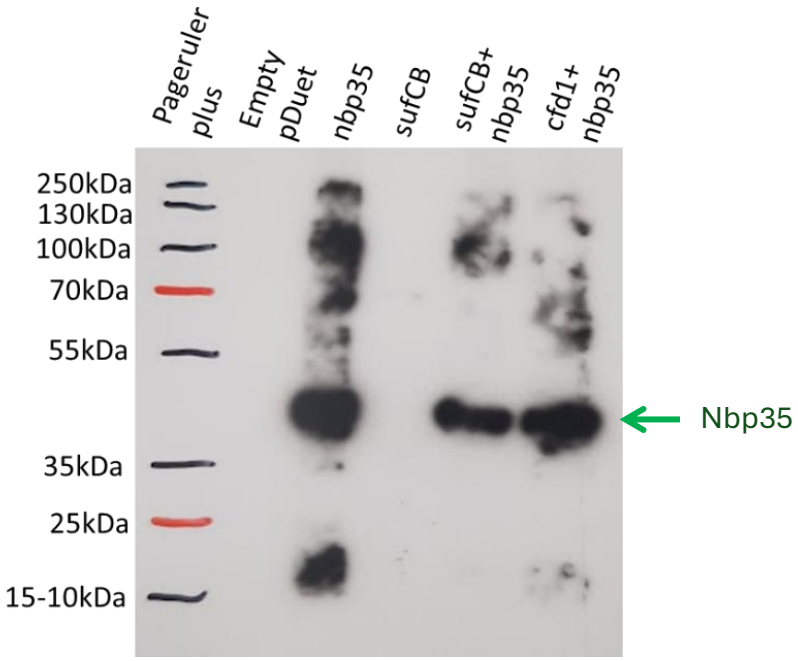

Figure 19: His-tagged Western blot of the His-tag purified proteins from the BL21 transformation. 20 second exposure. From left to right lanes show different samples, pageruler plus protein ladder, empty pDuet, lone Nbp35, lone SufCB, SufCB&Nbp35, Cfd1&Nbp35. These samples were loaded as mentioned in the methods section. Green arrow shows nbp35 bands at around 40kDa.

Figure 20 shows the presence of both Cfd1 and SufCB on the S-tagged western blot. Though this blot is not as clear as Figure 19, there are still visible bands. There is a band in the Cfd1 & Nbp35 sample at around 40kDa band from the Pageruler plus ladder. This band shows the presence of Cfd1, which has an expected weight of 31.916kDa. Thought this band is slightly higher, showing a slightly heavier protein, it still shows the presence of Cfd1 in this sample. The extra weight is likely due to post translational modifications to the protein. Unfortunately, due to the lack of a lone Cfd1 sample we cannot determine if the presence of Cfd1 is due to the formation of a complex with Nbp35 or if it is unexpectedly purified during the His-tag purification.

As well as this band, in Figure 20, In both the SufCB and SufCB & Nbp35 samples, there is a dark band at around 90kDa. This band being present in the lone SufCB sample means that despite

Nbp35 not being present in the sample, the SufCB protein was somehow pulled down in the His-tag purification, despite not having a His-tag. The equivalent band that is present in the SufCB & Nbp35 sample could be due to a complex forming, however there is no way to prove that, as it has been shown that the SufCB will stay despite the purification.

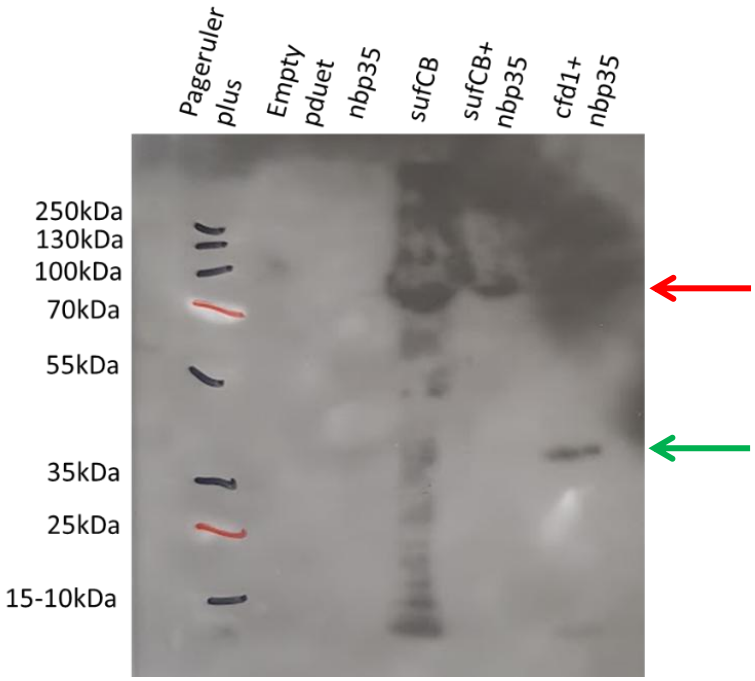

Figure 20: S-tagged Western blot of the His-tag purified proteins from the BL21 transformation. From left to right lanes show different samples, pageruler plus protein ladder, empty pDuet, lone Nbp35, lone SufCB, SufCB&Nbp35, Cfd1&Nbp35. These samples were loaded as mentioned in the methods section. Green arrow shows the Cfd1 band at around 40kDa. Red arrow shows the SufCB band at around 90kDa in both the lone SufCB and SufCB&Nbp35 lanes.

Discussion:

Summary of results:

In a previous project, others set their goal to modify the iron-sulphur cluster maturation machinery of yeast with specific proteins from the protozoa *Blastocystis*. Egan, 2019, had shown that the SufCB protein can be inserted and produced in *E.coli*. Using the co-immunoprecipitation of recombinant proteins expressed in E. coli, we were unable to confirm the observation that was originally produced with the yeast proteins that *Blastocystis* sufCB and yeast Nbp35 physically interact (Egan R., 2019). This could potentially show that SufCB does not act as a Cfd1 replacement in *Blastocystis*.

Insertion of the coding DNA into pDuet plasmid:

Our first section of this project involved the insertion of *NBP35*, *SufCB,* and *CFD1* coding DNA into pDuet plasmids for transformation into *E.coli.* pDuet is a type of expression vector that is used to co-express two proteins simultaneously in *E.coli*. pDuet contains two separate cloning sites each under the control of a T7 promoter. This is used for BL21 *E.coli* as this strain has an inducible T7 RNA polymerase system that can work with these T7 promoters (Sambrook et al., 2001). T7 polymerase is integrated into the E. coli genome under the control of an IPTG-repressible promoter. This means that the expression of genes, which is controlled by T7 promoters, can be induced through the addition of IPTG. As well as this, our chosen pDuet plasmid had an ampicillin resistant gene present to allow for the specific selection of bacterial cells. It should also be mentioned that the pDuet plasmid that we used contained coding DNA for both a polyhistidine tag (His-tag) and an s-tag on the *NBP35* and *SufCB*/*CFD1* sites respectively. These tags were attached to the proteins when they were produced, thus allowing the proteins to be easily selected for when analysing. Using this pDuet plasmid has allowed us to study the interactions that occur between our proteins and their coding DNA when present within the cell.

When inserting our coding DNA into the pDuet plasmid, it was required that we insert the *NBP35* coding DNA last. This is due to the presence of a NdeI site inside the *NBP35* gene. This NdeI site is used to open the pDuet plasmid for the *SufCB* and *CFD1* genes. Therefore, if we were to insert the *NBP35* gene first, then it would be cut during the insertion of *SufCB* or *CFD1*.

Unfortunately, later on in our research we found that we were having problems with the co-expresses sufCB and *nbp35*. SDS-PAGE and western blotting results were showing either a decreased amount of, or complete lack of, sufCB in the sufCB & *nbp35* samples. Due to this problem occurring, we decided to use the lone sufCB sample and mix it with the lone *nbp35* sample before immunoprecipitation. Unfortunately, this meant that in this case the use of pDuet to co-express the two proteins was not utilised. However, it should be noted that, when comparing the lone *cfd1* to the *cfd1* and *nbp35* sample, the latter was proving to produce more *cfd1,* and so we used this as opposed to the lone *cfd1*. In future studies, this should be looked into more, as it could tell us more about how *cfd1* and sufCB interact with *nbp35* in native conditions.

Comparing BL21 and Bl21-Codonplus:

The next section of this project was determining if Bl21-Codonplus was worth using instead of regular Bl21. BL21 is a standard strain of *E.coli* that is commonly used for protein expression. Bl21-Codonplus is a specifically engineered version of BL21 designed to overcome codon bias. BL21-Codonplus contains additional tRNAs that decode rare codons. This change makes the Bl21-Codonplus strain more efficient at expressing certain proteins, especially those that originate from different organisms, with different codon usage (Rosano et al., 2014). Because our initial experiments had shown relatively low expression levels of yeast proteins, we hypothesised that BL21-Codonplus might increase the expression levels for these genes. Our target genes could contain codons that are rarely transcribed in *E.coli* but transcribed more in Bl21-Codonplus.

We decided to test whether Bl21-Codonplus was actually more efficient or not at producing both *Blastocystis* SufCB and Yeast nbp35 and cfd1. Unfortunately, our results were not as conclusive as we had expected. It appeared that Bl21-Codonplus was less effective at producing SufCB than regular Bl21 was. In the future, more time should be spent analysing the efficiency of using Bl21-Codonplus instead of Bl21. Due to our results not matching our original hypothesis, more conclusive results should be gathered to truly determine which strain is better used when studying these proteins in *E.coli*. This could lead to vastly different results, as the production of these recombinant proteins would be more accurate to their native efficiency.

Co-Immunoprecipitation:

The final section of this thesis involves the co-immunoprecipitation of nbp35 to detect the binding of SufCB and Cfd1 to Nbp35. Immunoprecipitation is a technique used to purify targeted proteins from mixtures utilising antibodies that bind to the proteins (Aebersold & Mann., 2003). Co-immunoprecipitation is a version of immunoprecipitation that is used to analyse interactions between proteins by utilising the co-isolation of proteins when only one is purified (Iqbal et al., 2018).

In our experiment, we utilised the histidine tag on Nbp35 to pull it down. This has allowed us to detect if SufCB or Cfd1 are forming complexes with Nbp35 as we would expect. The results we collected showed that the SufCB protein was immunoprecipitated when no Nbp35 was present in the sample. This is unexpected and should not have happened as the SufCB protein did not contain a His-tag binding site.

One potential reason for this immunoprecipitation could be the effect of immobilised nickel ions on the nickel column. It is possible that the nickel ions have become immobilised on the nickel columns using chelators and that these chelators are acting on the iron that is present in the iron-sulphur clusters. Nickel affinity chromatography is a common method used to purify His-tagged proteins. It works by immobilising nickel ions in a column via chelators, nitrilotriacetic acid (NTA)

in our case. These chelators form stable complexes with the nickel ions, thus allowing the His-tagged protein to bind to the column and be eluted (Zheng et al., 1993). However, it could be possible that the iron from the iron sulphur clusters of SufCB is forming a complex with the nickel, thus eluting the SufCB protein without the need for a His-tag.

In the future, others should look further into this to determine if the iron-sulphur clusters of other proteins are affected this same way. As well as this, we were unable to compare the Cfd1 and Nbp35 sample with a lone cfd1 sample. This means that we could not determine if the Cfd1 is being co-immunoprecipitated with the Nbp35 or if it is immunoprecipitating on its own, like the lone SufCB sample. In future experiments, a lone Cfd1 sample should be used to ensure that the expected complex is being formed between Cfd1 and Nbp35.

Conclusion:

Overall, the work outlined in this thesis has assisted in gathering information relating to the SufCB protein in *Blastocystis* and how it acts when inserted into *E.coli*. As well as this, we have shown that SufCB and cfd1 coding DNA can be inserted into plasmids and co-expressed in *E.coli* but this can cause issues and that expressing the proteins separately may be more effective. Overall, these findings will assist future attempts to further this area of research.

This work leaves lots of opportunities for further research around the *Blastocystis* protein SufCB and its integration into other organisms. There is potential for others to research how other protozoan SufCB proteins interact with Nbp35 and if it is any different from the results we have shown in this thesis. It would also be interesting to delve further into the interactions between *Blastocystis* sufCB and its native conditions within the organism.

References:

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. Nature, 422(6928), 198–207.

Alhebshi, A., Sideri, T.C., Holland, S.L., et al. (2012). The essential iron-sulfur protein Rli1 is an important target accounting for inhibition of cell growth by reactive oxygen species. Mol. Biol. Cell 23, 3582–3590.

Bai, Y., Chen, T., Happe, T., Lu, Y., & Sawyer, A. (2018). Iron-sulphur cluster biogenesis via the SUF pathway. *Metallomics : integrated biometal science*, *10*(8), 1038–1052.

Bedekovics, T., Li, H., Gajdos, G.B., et al. (2011). Bedekovics et al. 2011 - Leucine Biosynthesis Regulates Cytoplasmic Iron-Sulfur Enzyme Biogenesis in an Atm1p-independent Manner.pdf. J. Biol. Chem. 286, 40878–40888.

Blahut, M., Sanchez, E., Fisher, C. E., & Outten, F. W. (2020). Fe-S cluster biogenesis by the bacterial Suf pathway. *Biochimica et biophysica acta. Molecular cell research*, *1867*(11), 118829.

Boyd, E.S., Thomas, K.M., Dai, Y., et al. (2014). Interplay between Oxygen and Fe-S Cluster Biogenesis: Insights from the Suf Pathway. Biochemistry 53, 5834–5847.

Bych, K., Netz, D.J.A., Vigani, G., et al. (2008). The essential cytosolic iron-sulfur protein Nbp35 acts without Cfd1 partner in the green lineage. J. Biol. Chem. 283, 35797–35804.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. 2009. BLAST+: architecture and applications.

Champoux JJ (2001). "DNA topoisomerases: structure, function, and mechanism". *Annual Review of Biochemistry*. **70**: 369–413

Cho, J. S., Kim, G. B., Eun, H., Moon, C. W., & Lee, S. Y. (2022). Designing Microbial Cell Factories for the Production of Chemicals. JACS Au, 2(8), 1781–1799.

Djaman, O., Outten, F. W., & Imlay, J. A. (2004). Repair of oxidized iron-sulfur clusters in Escherichia coli. *The Journal of biological chemistry*, *279*(43), 44590–44599.

Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J. L., Vázquez, E., & Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals. *Microbial cell factories*, *8*, 17.

Frazzon, J., & Dean, D. R. (2003). Formation of iron-sulfur clusters in bacteria: an emerging field in bioinorganic chemistry. *Current opinion in chemical biology*, *7*(2), 166–173.

Gentekaki, E., Curtis, B. A., Stairs, C. W., Klimeš, V., Eliáš, M., Salas-Leiva, D. E., Herman, E. K., Eme, L., Arias, M. C., Henrissat, B., Hilliou, F., Klute, M. J., Suga, H., Malik, S. B., Pightling, A. W., Kolisko, M., Rachubinski, R. A., Schlacht, A., Soanes, D. M., Tsaousis, A. D., ... Roger, A. J. (2017). Extreme genome diversity in the hyper-prevalent parasitic eukaryote Blastocystis. *PLoS biology*, *15*(9), e2003769.

Hagman, A., Säll, T., Compagno, C., & Piskur, J. (2013). Yeast "make-accumulate-consume" life strategy evolved as a multi-step process that predates the whole genome duplication. *PloS one*, *8*(7), e68734.

Hinton, T. V., Batelu, S., Gleason, N., & Stemmler, T. L. (2022). Molecular characteristics of proteins within the mitochondrial Fe-S cluster assembly complex. *Micron (Oxford, England : 1993)*, *153*, 103181.

Huang, M., Bao, J., and Nielsen, J. (2014). Biopharmaceutical protein production by Saccharomyces cerevisiae: current state and future prospects. Pharm. Bioprocess. 2, 167–182.

Iqbal, H., Akins, D. R., & Kenedy, M. R. (2018). Co-immunoprecipitation for Identifying Protein-Protein Interactions in Borrelia burgdorferi. *Methods in molecular biology (Clifton, N.J.)*, *1690*, 47–55.

Johnson M. K. (1998). Iron-sulfur proteins: new roles for old clusters. *Current opinion in chemical biology*, *2*(2), 173–181.

Johnson, D. C., Dean, D. R., Smith, A. D., & Johnson, M. K. (2005). Structure, function, and formation of biological iron-sulfur clusters. *Annual review of biochemistry*, *74*, 247–281.

Lee, S. Y., & Kim, H. U. (2015). Systems strategies for developing industrial microbial strains. *Nature Biotechnology*, 33(10), 1061-1072.

Lill R. (2009). Function and biogenesis of iron-sulphur proteins. *Nature*, *460*(7257), 831–838.

Lill R. (2020). From the discovery to molecular understanding of cellular iron-sulfur protein biogenesis. *Biological chemistry*, *401*(6-7), 855–876.

Lill, R., & Mühlenhoff, U. (2006). Iron-sulfur protein biogenesis in eukaryotes: components and mechanisms. *Annual review of cell and developmental biology*, *22*, 457–486.

Lu, Y., Yeung, N., Sieracki, N., & Marshall, N. M. (2009). Design of functional metalloproteins. *Nature*, *460*(7257), 855–862.

MORTENSON, L. E., VALENTINE, R. C., & CARNAHAN, J. E. (1962). An electron transport factor from Clostridium pasteurianum. *Biochemical and biophysical research communications*, *7*, 448–452.

Netz, D.J.A.A., Pierik, A.J., Stümpfig, M., et al. (2007). The Cfd1–Nbp35 complex acts as a scaffold for iron-sulfur protein assembly in the yeast cytosol. Nat. Chem. Biol. 3, 278–286.

Nielsen J. (2013). Production of biopharmaceutical proteins by yeast: advances through metabolic engineering. *Bioengineered*, *4*(4), 207–211.

Nielsen, J., & Keasling, J. D. (2016). *Engineering Cellular Metabolism*. Cell, 164(6), 1185-1197.

Outten, F. W., Wood, M. J., Munoz, F. M., & Storz, G. (2003). The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in Escherichia coli. *The Journal of biological chemistry*, *278*(46), 45713–45719.

Ozer, H. K., Dlouhy, A. C., Thornton, J. D., Hu, J., Liu, Y., Barycki, J. J., Balk, J., & Outten, C. E. (2015). Cytosolic Fe-S Cluster Protein Maturation and Iron Regulation Are Independent of the Mitochondrial Erv1/Mia40 Import System. *The Journal of biological chemistry*, *290*(46), 27829–27840.

Pallesen, L.J., Solodovnikova, N., Sharma, A.K., et al. (2013). Interaction with Cfd1 increases the kinetic lability of FeS on the Nbp35 scaffold. J. Biol. Chem. 288, 23358–23367.

Parapouli, M., Vasileiadis, A., Afendra, A. S., & Hatziloukas, E. (2020). Saccharomyces cerevisiae and its industrial applications. AIMS microbiology, 6(1), 1–31.

Pray, L. (2008) Recombinant DNA technology and transgenic animals. Nature Education 1(1):51

Raj, K., Partow, S., Correia, K., Khusnutdinova, A. N., Yakunin, A. F., & Mahadevan, R. (2018). Biocatalytic production of adipic acid from glucose using engineered *Saccharomyces cerevisiae*. *Metabolic engineering communications*, *6*, 28–32.

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in microbiology*, *5*, 172.

Rouault, T. A., & Tong, W. H. (2008). Iron-sulfur cluster biogenesis and human disease. *Trends in genetics : TIG*, *24*(8), 398–407.

Saha, P. P., Vishwanathan, V., Bankapalli, K., & D'Silva, P. (2018). Iron-Sulfur Protein Assembly in Human Cells. *Reviews of physiology, biochemistry and pharmacology*, *174*, 25–65.

Sambrook, Joseph & Russell, David W. (1957) & Cold Spring Harbor Laboratory. (2001). *Molecular cloning : a laboratory manual / Joseph Sambrook, David W. Russell*. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory

Sipos, K., Lange, H., Fekete, Z., Ullmann, P., Lill, R., & Kispal, G. (2002). Maturation of cytosolic iron-sulfur proteins requires glutathione. *The Journal of biological chemistry*, *277*(30), 26944–26949.

Stehling, O., Jeoung, J.H., Freibert, S.A., et al. (2018). Function and crystal structure of the dimeric P-loop ATPase CFD1 coordinating an exposed [4Fe-4S] cluster for transfer to apoproteins. Proc. Natl. Acad. Sci. U. S. A. 115, E9085–E9094.

Stensvold, C. R., & Clark, C. G. (2016). Current status of Blastocystis: A personal view. *Parasitology international*, *65*(6 Pt B), 763–771.

Stephanopoulos, G., & Aristidou, A. (1998). *Metabolic Engineering: Principles and Methodologies*. Academic Press.

Tan K. S. (2008). New insights on classification, identification, and clinical relevance of Blastocystis spp. *Clinical microbiology reviews*, *21*(4), 639–665.

Thompson, J.D. and Higgins, D.G. and T.J.Gibson. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice., Nucleic Acids Research

Tsaousis, A. D., Ollagnier de Choudens, S., Gentekaki, E., Long, S., Gaston, D., Stechmann, A., Vinella, D., Py, B., Fontecave, M., Barras, F., Lukeš, J., & Roger, A. J. (2012). Evolution of Fe/S cluster biogenesis in the anaerobic parasite Blastocystis. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(26), 10426–10431.

Tsaousis, A. D., Ollagnier de Choudens, S., Gentekaki, E., Long, S., Gaston, D., Stechmann, A., Vinella, D., Py, B., Fontecave, M., Barras, F., Lukeš, J., & Roger, A. J. (2012). Evolution of Fe/S

cluster biogenesis in the anaerobic parasite Blastocystis. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(26), 10426–10431.

Tsaousis, A.D., Gentekaki, E., Eme, L., et al. (2014). Evolution of the Cytosolic Iron-Sulfur Cluster Assembly Machinery in Blastocystis Species and Other Microbial Eukaryotes. Eukaryot. Cell 13, 143–153.

Vallières, C., Benoit, O., Guittet, O., Huang, M. E., Lepoivre, M., Golinelli-Cohen, M. P., & Vernis, L. (2024). Iron-sulfur protein odyssey: exploring their cluster functional versatility and challenging identification. *Metallomics : integrated biometal science*, *16*(5), mfae025.

Wächtershäuser G. (1992). Groundworks for an evolutionary biochemistry: the iron-sulphur world. *Progress in biophysics and molecular biology*, *58*(2), 85–201.

Wang, B., et al. (2017). Advanced microbial engineering: New concepts and applications. Nature Reviews Microbiology, 15(12), 735-750.

Wang, M.S., Hoegler, K.J., and Hecht, M.H. (2019). Unevolved De Novo Proteins Have Innate Tendencies to Bind Transition Metals. Life 9, 1–15.

Wang, Y., Qian, J., Shi, T., Wang, Y., Ding, Q., & Ye, C. (2024). Application of extremophile cell factories in industrial biotechnology. *Enzyme and microbial technology*, *175*, 110407.

Záhonová, K., Low, R. S., Warren, C. J., Cantoni, D., Herman, E. K., Yiangou, L., Ribeiro, C. A., Phanprasert, Y., Brown, I. R., Rueckert, S., Baker, N. L., Tachezy, J., Betts, E. L., Gentekaki, E., van der Giezen, M., Clark, C. G., Jackson, A. P., Dacks, J. B., & Tsaousis, A. D. (2023). Evolutionary analysis of cellular reduction and anaerobicity in the hyper-prevalent gut microbe Blastocystis. *Current biology : CB*, *33*(12), 2449–2464.

Zheng, L., White, R. H., Cash, V. L., Jack, R. F., & Dean, D. R. (1993). Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(7), 2754–2758.