

Submitted to the Division of Postgraduate Research of Kingston University,
London

**Investigating the metabolism of tartrazine by the human gut
microbiome**

This thesis is submitted in fulfilment of the requirements of Kingston University
for the degree of: **Masters by research**

Faculty of Science, Engineering and Computing

By: Gemma Hoynes

December 2021

Acknowledgements

First and foremost, to the dream team that is Dr. Ali Ryan and Dr. Elena Polycarpou, I don't know if I'll ever be able to express my gratitude and appreciation for honestly being the most extraordinary supervisors and friends that I've had the pleasure of working with. Thank you for providing an endless supply of patience, support, knowledge and inspiration, I couldn't have asked for two better mentors.

To all my colleagues in Biotech and the IRTL, thank you for always coming to my aid whenever I needed advice, emergency reagents, lab access, or just a good laugh. When the lab would start to feel lonely, I could always count on you for your energy and enthusiasm.

I am especially grateful to Rosalind Percival, who's understanding, flexibility and accommodations helped ease the burden during some particularly challenging times.

My husband, who is undoubtably one of the biggest distractions in my life. Thank you for picking me up when I was down, supporting me when I couldn't, encouraging me when I wanted to give up, and most importantly for cooking me meals with actual nutrition. You've made this entire journey significantly better and dare I say it, enjoyable.

I am forever grateful to my family, who have provided an incredible network of encouragement, support and understanding throughout my entire educational journey.

A huge apology to all of my friends and family for consistently using my thesis to excuse all of my last minute rainchecks, empty responses and overall absence these past two years. Please clear your calendars!

An obligatory shout out to my pup Harley. What would usually start out as a guilt provoked, obligatory walk would always turn out to be a welcome reprise from a long study session.

And as a final note, I will forever be indebted to this project, which brought to light two illnesses I had unknowingly suffered with for most of my life. I would like to give a special thanks to Dr Ana, who gave me the tools and support necessary for managing both conditions and to endure accompanying challenges, all while continuing to live a normal life. To conclude, I can say with confidence that by undertaking this postgraduate degree, I have ultimately become the healthiest, happiest version of myself, and for that I will forever be grateful.

Table of contents

Acknowledgements	2
Table of contents	3
List of figures.....	5
List of tables.....	7
Abstract.....	9
Chapter 1 Introduction.....	10
Chapter 2 Materials and methods	25
2.1 Chemicals and reagents.....	25
2.2 Bacterial Strains and growth conditions	25
2.3 Genomic DNA quantification	27
2.4 Bioinformatics.....	27
2.5 Gene amplification.....	27
2.5.1 Agarose gel analysis of PCR products.....	29
2.5.2 Purification and quantification of PCR products	29
2.6 Cloning of putative azoreductase gene	29
2.7 Transformation.....	31
2.7.1 Analysing successful transformants.....	31
2.7.2 Plasmid extraction.....	33
2.7.3 DNA sequencing.....	33
2.8 Transformation.....	33
2.9 Recombinant protein expression.....	33
2.9.1 Small scale recombinant protein expression.....	34
2.9.2 Large scale recombinant protein expression.....	35
2.9.3 SDS-PAGE analysis of all recombinant proteins	35

Chapter 3 Cloning of putative azoreductases genes from the genomic DNA of O. splanchnicus.....	37
3.1 Introduction.....	37
3.2 Results.....	40
3.2.1 Bioinformatics.....	40
3.2.2 Amplification of putative azoreductase genes	40
3.2.3 Cloning and transformation of putative azoreductase gene	42
3.3 Discussion.....	45
Chapter 4 Optimising the accumulation of soluble recombinant protein by performing expression trials.....	47
4.1 Introduction.....	47
4.2 Results.....	48
4.2.1 Expression of putative azoreductase enzymes using the pET102 and pET100 expression vectors.....	48
4.2.2 Comparison on the efficacy of the pET102 and pET100 expression vectors in obtaining soluble protein.....	48
4.2.3 Identifying the optimum length of expression time	50
4.2.4 Identifying the optimal growth medium type	54
4.2.5 Identifying the optimal IPTG concentration	54
4.2.6 Identifying the optimal expression temperature	55
4.2.7 Determining the overall environment best suited for the production of soluble recombinant protein	55
4.2.8 Scale up of protein expression.....	56
4.3 Discussion.....	57
Chapter 5 Final discussion and conclusion.....	60
References.....	69

List of figures

Figure 1.1 An example of host xenobiotic metabolism via hepatic biotransformation. ..	12
Figure 1.2 The methods of xenobiotic metabolism undertaken by the human gut microbiome	13
Figure 1.3 Mechanisms of gut microbiome modulation of host drug metabolism and response.....	16
Figure 1.4 The enzymatic reduction of the azo dye amaranth by an azoreductase enzyme, taken from Misal & Gawai, (2018).....	17
Figure 3.1 PCR amplification of putative azoreductase genes from <i>O. splanchnicus</i> using varying annealing temperatures and pET102 designed primers.	41
Figure 3.2 PCR amplification of putative azoreductase genes from <i>O. splanchnicus</i> using varying annealing temperatures and pET100 designed primers.	41
Figure 3.3 Colony screening for the successful transformation of ligated PCR product and pET102 vector into <i>E. coli</i>. One Shot® TOP10 cells.	43
Figure 3.4 Colony screening for the successful transformation of ligated PCR product and pET100 vector into <i>E. coli</i>. One Shot® TOP10 cells.	43
Figure 3.5 Alignment of the forward and reverse deduced amino acid sequences of a putative azoreductase from <i>O. splanchnicus</i>, against the original protein sequence obtained from the NCBI database.....	44
Figure 4.1 SDS-PAGE analysis of protein expressed from the pET100 and pET102 vectors.	49
Figure 4.2 SDS-PAGE analysis of protein expressed in TB media from the pET100 and pET102 vectors.....	51
Figure 4.3 SDS-PAGE analysis of the effect of varying growth medias, IPTG concentrations and induction temperatures on the expression of soluble protein from the pET102 expression vector.....	53

Figure 4.4 SDS-PAGE analysis of protein expression from the pET102 vector, scaled up to 100ml with the application of specific expression parameters.56

Figure 5.1 Major different human and murine intestinal genera. Only genera are shown that showed consistent differences in relative abundance between humans and mice...65

Figure 5.2 A heatmap demonstrating the abundance of (A) *O. splanchnicus* and (B) other identified gut bacterial strains, in faecal samples that were collected from 16 healthy adults. Each strain is graded by it's ability in reducing the azo dye Brilliant Black.....66

List of tables

Table 1.1 The names, codes, colour and chemical structure of the azo food dyes that are currently approved for use in the EU and the USA and the clinical manifestations found to have been associated with each dye.....	20
Table 1.2 The 22 human gut bacterial strains isolated from stool samples that demonstrated an ability to metabolise azo dyes.....	22
Table 1.3 The 7 azo dye excipients that were incubated with and reduced by the bacterial strains listed in Table 1.2.....	22
Table 2.1 Growth media	26
Table 2.2 Primers used for the amplification of putative <i>O. splanchnicus</i> 22-5 azoreductase gene.....	28
Table 2.3 PCR reaction mix	28
Table 2.4 Thermocycling parameters used for the amplification of <i>O. splanchnicus</i> 22-5	28
Table 2.5 The reagents used to prepare a 1x TBE buffer for agarose gel electrophoresis	30
Table 2.6 TOPO® cloning reaction mix. Adapted from (Thermo Fisher Scientific)	30
Table 2.7 The primers used to identify successful transformants	32
Table 2.8 Colony PCR reaction mix	32
Table 2.9 Thermocycling parameters used during colony PCR.....	32
Table 2.10 Materials used for SDS-PAGE and the staining and de-staining of SDS gels. Part adapted from (Holland, 2017)	36
Table 3.1 A compilation of the bacterial azoreductases that have been characterised to date.	39

Table 5.1 The amino acid sequences of the azoreductases that had been characterised to date, alongside their amino acid length and ID and accession numbers.61

Table 5.2 A hypothetical protein identified from the genome of *O. splanchnicus* using protein sequences from characterised azoreductases in a blastp search.61

Abstract

Azo dyes are a class of food dyes that are widely used in a variety of commercial industries. These synthetic colours are aromatic in nature and are characterised by one or more functional $-N=N-$ groups within its molecular structure. One example of an azo dye is tartrazine. It's lemon-yellow shade lends a bright, appealing colour to food products and pharmaceuticals, and is one of the most popular artificial dyes in the food trade. Tartrazine has also long been associated with negative health effects such as ADHD like symptoms, allergic responses and tissue damage. Posing a risk to human health has led to the prohibition and removal of many azo dyes from consumer products. At present, despite a long history of investigation, the toxicity of tartrazine has not yet been established and is still authorised for use in many countries.

Azoreductases are enzymes present in mammals, bacteria and yeast that are recognised for their ability to reduce azo-dyes which in turn generates metabolites such as aromatic amines, often with carcinogenic and toxic properties. It is well known that the human gut microbiota play an integral part in the reduction, activation and detoxification of xenobiotics, however the majority of these metabolic pathways remain uncharacterised.

This study aims to help advance these efforts, by selecting a common gut bacterial strain that has already demonstrated azoreductase capabilities, yet the enzymes responsible remain uncharacterised. To identify the genes that code for an azoreductase enzyme, the full genomic DNA of *Odoribacter splanchnicus*, a bacterial strain common to the human gut, was selected and a blastp search was performed, using a library of already characterised azoreductase protein sequences. This generated one protein sequence result which bore 51% sequence similarity to AzoC, an azoreductase from the anaerobic bacteria *Clostridium perfringens*. To amplify this gene, PCR conditions were optimised by applying a range of annealing temperatures. The temperature at which gene amplification was highest allowed for the gene of interest to be cloned into a choice of two pET TOPO[®] vectors which was then transformed into *E. coli* BL21 Star[™] (DE3) cells. Varying combinations of environmental conditions were applied during protein expression trials to optimise the expression of soluble protein. Time constraints prevented further experiments which would aid in the identification and subsequent characterisation of a putative azoreductase enzyme from the genome of *O. splanchnicus*. However, this study highlights the importance in uncovering the enzymatic properties of our gut microbiome, particularly in species that have not yet been identified.

Chapter 1 Introduction

Microbial life is believed to have been the very first life on Earth, originating over 3.5 billion years ago (Schopf *et al.*, 2002; Knoll, 2015). Their adaptable nature has allowed them to colonise all available ecological niches, transforming our once inhabitable planet into one that could support the evolution of multi-cellular life (Stal, 2007). Since then, microbes have continued to co-evolve with eukaryotic life in a way that allows them to inhabit, interact with and influence the development of humans and other complex organisms (McFall-Ngai *et al.*, 2013). The human body and its many microbial inhabitants share a fine-tuned, symbiotic relationship that has developed over the course of 15 million years (Moeller *et al.*, 2016) where microbes form specialised communities that thrive in specific regions of the body, the largest of which resides in the intestine, with a population of up to 100 trillion (10^{14}) microbes (Whitman, Coleman and Wiebe, 1998). This particular community, often referred to as the *human gut microbiome* (Liang, Leung, Guan and Au, 2018) is known to play a role in many vital human processes including metabolism (Li *et al.*, 2008), digestion (Bergman, 1990) and immunity (Mazmanian, Liu, Tzianabos and Kasper, 2005).

Colonisation of the human gut begins at birth upon neonatal exposure to host maternal gut microbes (Korpela, 2021). This event is crucial in the early development of the immune system (Gensollen, Iyer, Kasper and Blumberg, 2016) where, for example, the establishment of commensal bacterial species can direct the maturation of T cells (Lathrop *et al.*, 2011) and hinder the formation of invasive pathogenic communities, thereby preventing infection (Van der Waaij, Berghuis-de Vries and Lekkerkerk-van der Wees, 1971).

Once established, the gut microbiome contributes to digestion by metabolising undigested compounds such as starch that is resistant to amylase digestion, oligosaccharides and sugar alcohols (Ramakrishna, 2013). In turn, many of these bacterial strains synthesise essential nutrients that the human body itself is unable to produce such as vitamin K which is essential for normal blood coagulation, and for supporting bone and cardiovascular health (Walther, Karl, Booth and Boyaval, 2013). Other microbial metabolites include short chain fatty acids that provide energy to the host (Smith *et al.*, 2013), indoles and neurotransmitters that are thought to influence brain development and function (Heijtz *et al.*, 2011), and bile acid metabolites which are involved in microbe-microbe and host-microbe interactions (Wahlström, Sayin, Marschall and Bäckhed, 2016).

The composition of an individual's microbiome is as unique as their fingerprint (Cryan and O'Mahony, 2011) and studies examining this diversity have identified implications on host health resulting from a disruption to or a reduction of harboured microbial communities, which is also known as gut dysbiosis (Kriss *et al.*, 2018). For example, the treatment of an acquired infection with antibiotics, can reduce microbial diversity, and in particular commensal bacterial strains, resulting in the potential increase of pathogenic bacterial strains including/especially those with already acquired antibiotic resistance (Stecher *et al.*, 2010).

The gut microbiome is frequently referred to as an essential or metabolic organ (Possemiers, Bolca, Verstraete and Heyerick, 2011), collectively encoding for an estimated 2 – 4 million microbial genes (Qin *et al.*, 2010) and although its metabolic functions have yet to be fully established, it is widely accepted that its metabolic capacity exceeds that of its mammalian counterpart, the liver, (Possemiers, Bolca, Verstraete and Heyerick, 2011). Both the gut microbiome and the mammalian liver demonstrate an ability to reduce a wide array of xenobiotics *in vivo*, inclusive of pharmaceuticals, cosmetics and environmental pollutants (Patterson, Gonzalez and Idle, 2010), however, the methods used to reduce these compounds differ between hepatic and microbial enzymes (Koppel, Maini Rekdal and Balskus, 2017). The liver employs 'Phase I' enzymes to increase polarity of xenobiotics, followed by the addition of a functional chemical moiety by 'Phase II' enzymes to promote excretion (Figure 1.1) (Dancygier, 2010). However, microbial enzymes employ reductive and hydrolytic reactions which can essentially reverse this entire process, thereby encouraging the toxicity of these drugs (Figure 1.2) (Tuteja and Ferguson, 2019).

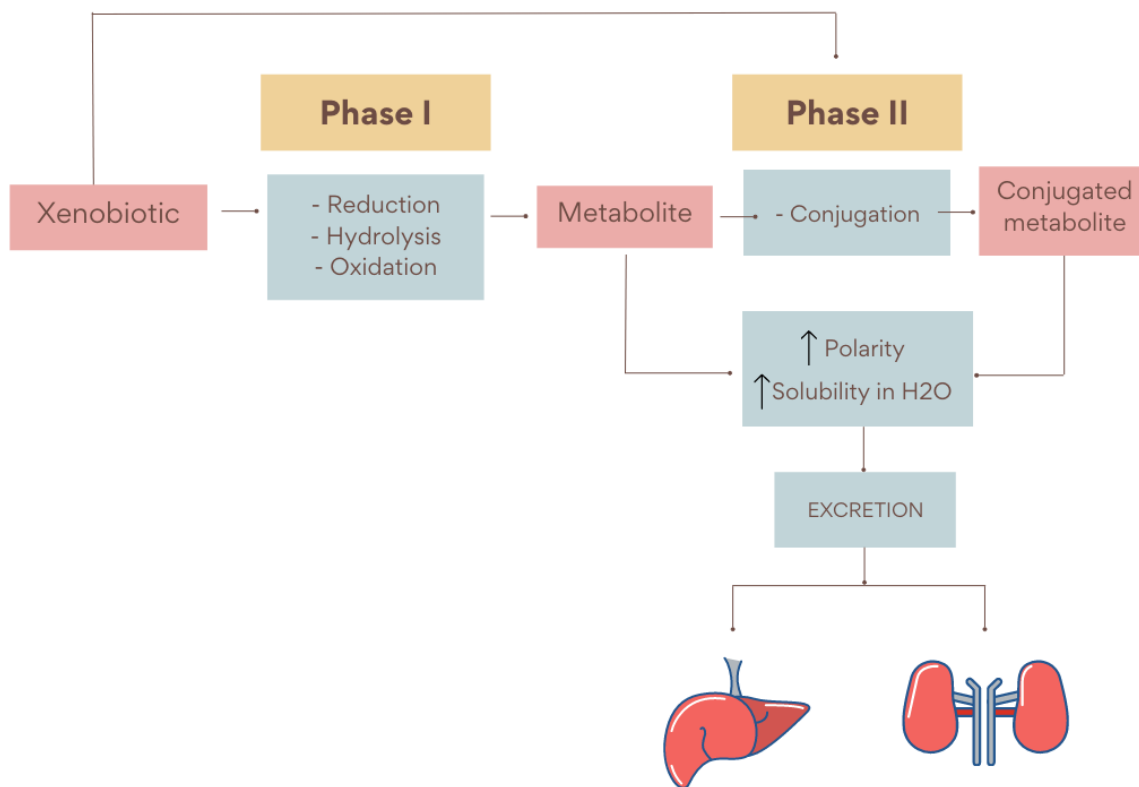


Figure 1.1 An example of host xenobiotic metabolism via hepatic biotransformation.

An example of host xenobiotic metabolism via hepatic biotransformation.

First, Phase I enzymes are employed to reduce foreign compounds, followed by conjugation of metabolites by Phase II enzymes. This increases their polarity and water solubility which subsequently promotes their excretion.

Figure is adapted from Dancygier (2010).

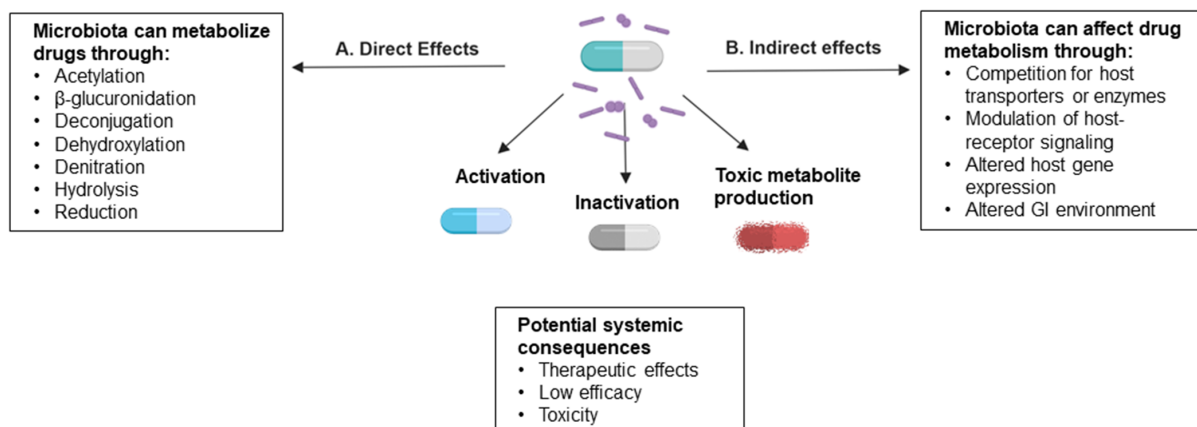


Figure 1.6 The methods of xenobiotic metabolism undertaken by the human gut microbiome

(A) Gut microbial enzymes can metabolise drugs directly by employing enzymatic reactions such as acetylation, azoreduction, β -glucuronidation, deconjugation, dihydroxylation, denitration, hydrolysis and reduction, thereby altering the pharmaceutical effect of the drug.

(B) the gut microbiome can also indirectly influence host response to xenobiotics by competing with host enzymes and receptors, interfering with host signalling pathways and gastrointestinal environmental factors. Both direct and indirect metabolic pathways result in the activation/inactivation of therapeutic agents and the potential generation of toxic metabolites which can alter systemic delivery of the drug and promote subsequent toxicity.

Figure is from Tuteja and Ferguson (2019).

A 1983 study highlights an important example of this, where Dobkin *et al.*, (1983) discovered that the metabolism and subsequent inactivation of the cardiac glycoside drug, digoxin, was dictated by the composition of an individual's gut microbiome and in particular, by the organism *Eubacterium lentum* (now known as *Eggerthella lentum*) (Figure 1.3 A). Similarly, Chen *et al.*, (2019) identified the ability of the gut microbial strain *Morganella morganii*, to metabolise L-phenylalanine into phenethylamine, a compound potentially fatal to those taking monoamine oxidase inhibitor drugs.

Azo dyes are another example of a xenobiotic compound that the gut microbiome can efficiently reduce (Roxon and Ryan, 1967). They are classified by their complex aromatic structure alongside the presence of one or more azo bonds ($-N=N-$) (Fatima, Farooq, Lindström and Saeed, 2017), which serve to strengthen its composition and provide stability (Lade, Waghmode, Kadam and Govindwar, 2012; Chengalroyen and Dabbs, 2012) These dyes are widely used across many industrial sectors including cosmetics, food, pharmaceuticals and textiles where their stable structure allows them to withstand heat, sunlight and varying pH, imparting a brighter, more permanent colour and requiring lower production costs (Llamas, Garrido, Nezio and Band, 2009; Przysaś, Zabłocka-Godlewska and Grabińska-Sota, 2011). The economic advantages of selecting synthetic dyes over their natural counterparts, saw a sustained rise in their popularity, dating back to the industrial revolution (Javaid and Qazi, 2019) and continuing today, where they account for more than 70% of the 900,000 tons of synthetic dye produced annually (Sarkar *et al.*, 2017).

The metabolism of azo dyes has been extensively studied as far back as 1911, where Sisley and Porcher (1911) first observed the reduction of the azo dye Orange I to aromatic amines; sulphanic acid and 4-amino-1-naphthol, but only when administered orally as opposed to subcutaneously. Similarly, Colebrook (1936) could demonstrate the antibacterial action of the azo prodrug prontosil against streptococcal infections *in vivo* (Figure 1.3 B). However the same effect could only be replicated *in vitro* when using its active metabolic product, sulphanilamide as opposed to the parent compound (Tréfouël, Nitti and Bovet, 1935). In response to these findings, Colebrook explained that the reduction of azo drugs *in vivo* first required enzymatic cleavage of their azo bonds, in order to release their metabolic products. Further investigation into the metabolic pathway of azo compounds revealed that the intestinal microflora was responsible for their degradation (Radmoski and Deichmann, 1956) and in 1967, following initial studies into the metabolism of tartrazine (Daniel, 1962), Roxon, Ryan and Wright (1967) finally discovered an enzyme capable of its reduction, which they

identified as an nicotinamide adenine dinucleotide (phosphate) (NAD(P)H) dependent, soluble flavoprotein bound to either a flavin mononucleotide (FMN) or a flavin adenine dinucleotide (FAD).

Today, these flavoproteins belong to a broader class of enzymes known as azoreductases, which have been identified in an ever-growing number of bacterial species (Misal and Gawai, 2018) as well as in humans and plants (Ryan *et al.*, 2014). They are characterized by their ability to metabolize azo compounds using a bi-bi ping pong mechanism where NAD(P)H is used to reduce the enzyme bound flavin; FMN or FAD ((Mueller and Miller, 1949. This allows the azoreductase to bind to an azo-dye, which receives a hydride from the reduced FMN, converting it into a hydrazine. In order to fully cleave the bond between azo groups, this cycle is repeated twice which in turn, converts both hydrazines into two amines, as illustrated in (Figure 1.4) (Misal and Gawai, 2018).

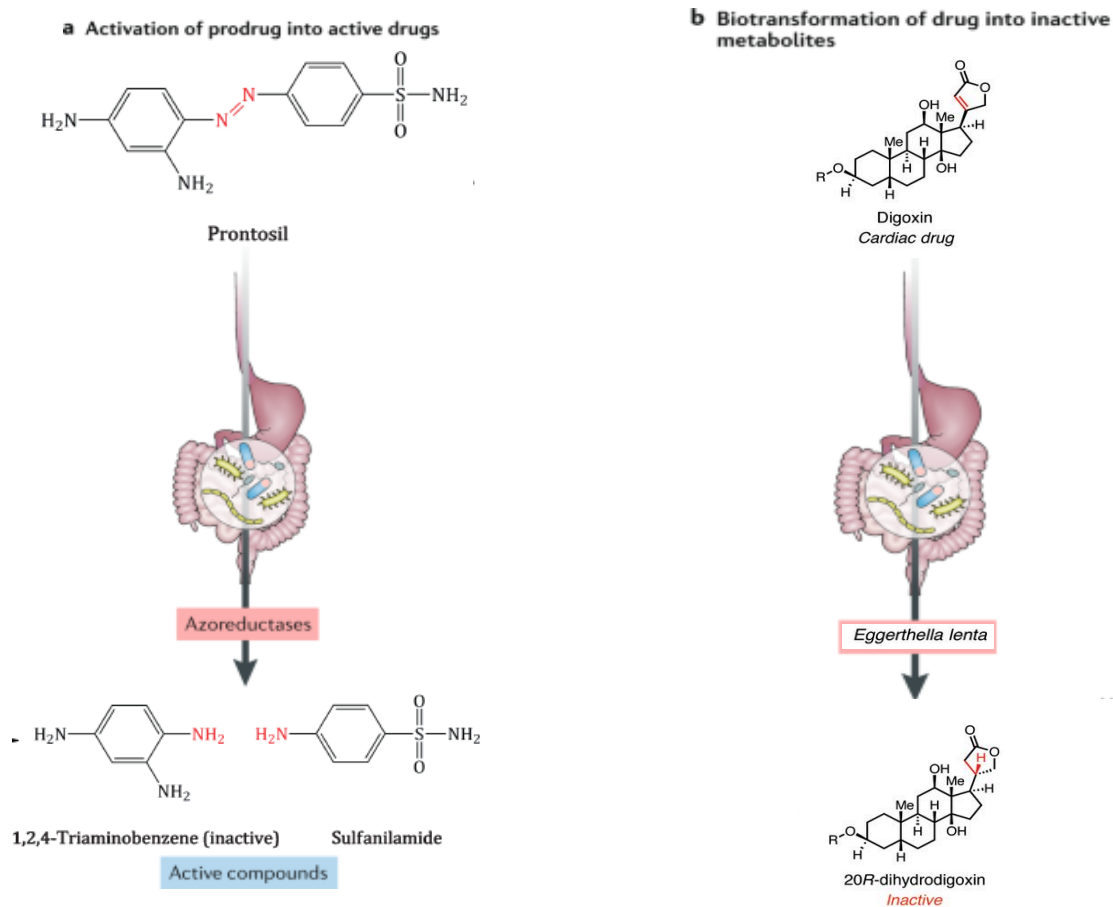


Figure 1.10 Mechanisms of gut microbiome modulation of host drug metabolism and response.

(A) Activation is the conversion of a prodrug into its bioactive form. For example, biotransformation of the inactive prodrug prontosil produces the inactive metabolites triaminobenzene (inactive) and sulfanilamide, which is the active form of the prodrug used to treat infection (B) Inactivation is the conversion of an active drug into a bioactive metabolite with reduced activity. For example, the cardiac drug digoxin is converted into dihydrodigoxin by the gut bacterial strain *E. lenta*.

Figure is adapted from Scher et al., (2020).

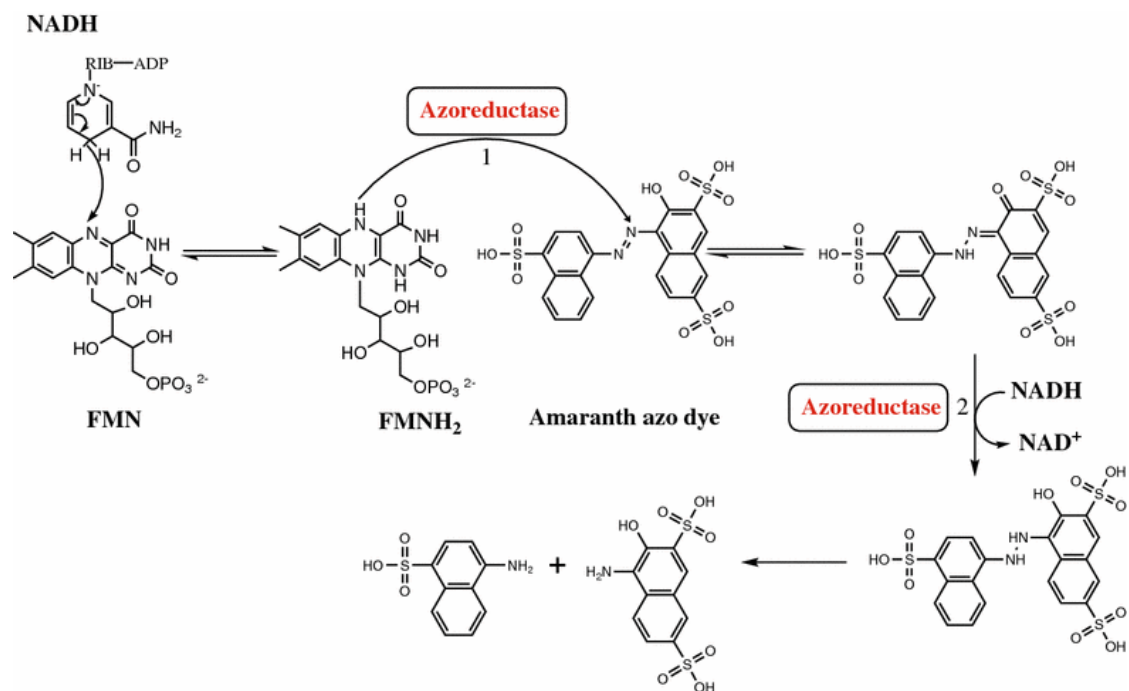


Figure 1.15 The enzymatic reduction of the azo dye amaranth by an azoreductase enzyme, taken from Misal & Gawai, (2018).

The first cycle involves the oxidation of NADH to reduce the enzyme-attached co-factor FMN. This allows for the transfer of a hydride to the azo dye which then binds to the azoreductase. The cycle is repeated twice in order to cleave the double bond between azo groups in the dye compound.

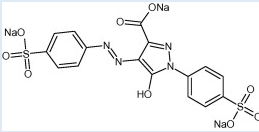
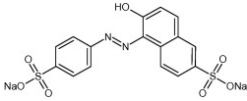
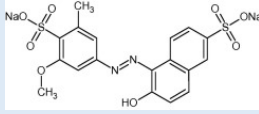
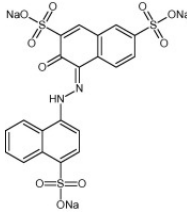
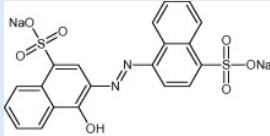
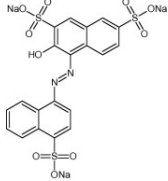
Azoreductase activity has been observed in a number of human gut microbial species including *Clostridium*, *Pseudomonas*, *Bacillus*, *Enterococcus*, *Eubacterium* and *Escherichia* (Misal & Gawai, 2018), however very few of these enzymes have been characterised (Morrison, Wright and John, 2012) and despite extensive research, their primary function has yet to be identified (Ryan *et al.*, 2010). This is especially concerning as humans are consistently exposed to thousands of azo compounds contained in food, pharmaceuticals, textiles and cosmetics, and over 500 of these contain aromatic amines. The particular use of synthetic azo dyes as food colourings has long been implicated with the potentially toxic and carcinogenic properties of their amine metabolites (Rimington & Hemmings, 1938; Franco *et al.*, 2018). Table 1.1 describes the effects exhibited by a number of ingested azo dyes, that have been identified as hazardous to human health. In response to this, many countries have introduced legislation aiming to recognise, control and where necessary, ban the use of specific azo colourings. One notable example of this, is the ‘Southampton study’, funded by the UK Food Standard Agency (FSA) and undertaken by McCann *et al.*, (2007) which focused solely on the consumption of the well-known ‘Southampton 6’ food colourings, five of which are azo dyes; Sunset yellow [E110], Carmoisine [E122], Tartrazine [E102], Ponceau 4R [E124], and Allura red AC [E129]). Quinoline yellow [E110] is a quinophthalone dye (EFSA, 2009).

The study suggested an association between the consumption of these dyes and an increase in hyperactivity and ADHD symptoms in children, however the study itself faced serious criticisms. The foods administered contained varying levels of food dyes and other ingredients capable of eliciting an allergic response, for example sulphites and monosodium glutamate. Additionally, the parents were left responsible for monitoring and reporting any noticeable side effects. Owing to the poor design of the study, these results were difficult to replicate and as a result, the use of these six synthetic colours was restricted as opposed to being prohibited as is done in a number of other countries (Table 1.1). (Coulter and Blackburn, 2018). The list of restricted and banned food dyes tends to vary worldwide, with many guidelines appearing to contradict each other, evidencing our lack of understanding regarding the mechanism of azo dye toxicity. For example, studies on the oral administration of the red azo dye; amaranth (FD&C Red No. 2 or E123) found that it could induce allergic reactions, hyperactivity and notable DNA damage in organ tissues (Gičević, Hindija and Karačić, 2019; Šuleková *et al.*, 2017). In response to these findings, the use of amaranth was prohibited in the USA, yet its consumption is still approved for use in the UK and Australia

(Bakthavachalu, Kannan and Qoronfleh, 2020). A more well-known example is tartrazine (E102 or FD&C Yellow 5), a synthetic yellow azo dye often found in foods such as flavoured crisps, sweets, sauces and cakes, in toiletries such as soap and shampoos and in pharmaceutical products including medicinal drinks, vitamins and the protective casing of some medications (Amin, Abdel Hameid and Abd Elsttar, 2010). Despite its ban in the UK, tartrazine remains to be one of the most widely used azo dyes in the global food industry (Esmaeili *et al.*, 2016; Kobylewski and Jacobson, 2015). Where tartrazine is authorised for human consumption, its outlined acceptable daily intake (ADI) is 0-7.5 mg/ kg.bw/day (Walton *et al.*, 1999). In response to its popularity, a considerable amount of research assessing the toxicological nature of tartrazine has identified an array of adverse effects associated with its consumption. Similar to amaranth, these effects include hyperactivity and ADHD symptoms in children (McCann *et al.*, 2007); allergic reactions such as asthma, hives and eczema (Moutinho, Bertges and Assis, 2007); and tissue damage in the liver and kidney, the last of which is suggested to be a result of the improper reduction of this dye (Amin *et al.*, 2010; Himri *et al.*, 2011).

Table 1.1 The names, codes, colour and chemical structure of the azo food dyes that are currently approved for use in the EU and the USA and the clinical manifestations found to have been associated with each dye.

Table was adapted from Feketea and Tsabouri, 2017; Bakthavachalu, Kannan and Qoronfleh, 2020; Yamjala, Nainar and Ramiseti, 2016; Corradini, 2019; Rovina, Siddiquee and Shaarani, 2016.

<i>Colour Name, Code (EU/USA)</i>	<i>Colour description</i>	<i>Approval in EU/USA</i>	<i>Chemical structure</i>	<i>Use</i>	<i>Associated clinical manifestations</i>
Tartrazine, E102, Yellow No. 5.	Lemon yellow	Banned in Norway and Austria, Permitted in the USA.		Confectionary, cheese, wine, baked goods, soft drinks, pharmaceuticals, cosmetics.	Asthma, hives, eczema, and ADHD symptoms.
Sunset yellow, E110, Yellow No. 6	Orange-Yellow	Banned in Norway, Approved in the rest of EU and USA.		Baked goods, soft drinks, confectionary, processed meats, pharmaceuticals.	Asthma, hives, eczema, angioedema and ADHD symptoms.
Allura red AC, E129, Red No. 40	Red	Banned in USA, Denmark, Belgium, Switzerland. Approved in rest of EU		Soft drinks, confectionary, meat products, pharmaceuticals, soaps, cosmetics, tattoo inks.	Hives, asthma and ADHD symptoms.
Amaranth, E123, Red No. 2	Red	Banned in USA, Austria and Norway. Approved in the rest of EU.		Soft and alcoholic drinks, cake mixes, confectionary, and fish roe	Hives, asthma, ADHD symptoms, headaches, damage to human lymphocyte cells.
Carmoisine, E122.	Red	Banned in USA, Sweden, Norway and Austria. Approved in rest of EU.		Confectionary, marzipan, soft drinks, mouthwash, cosmetics, pharmaceuticals.	Hives, asthma and ADHD symptoms.
Ponceau 4R, E124.	Red	Banned in USA, Norway. Approved in rest of EU.		Soft drinks, jams, confectionary, canned fruits, processed meats.	Hives, allergic reactions in those with asthma or aspirin sensitivity

Despite these extensive studies, the enzymes responsible for the reduction of azo dyes like tartrazine, have not yet been characterised. However a very recent study by Zou *et al.*, (2020) elegantly describes the identification of 22 human gut bacterial strains (Table 1.2) that were found to exhibit strong azoreductase activity against azo dyes similar to tartrazine, including Allura red (E129 or Red No. 40) and Sunset yellow FCF (E110 or Yellow No. 6) (Table 1.3). As the 2020 study did not go as far as to characterise these azoreductases, this project will focus on the identification of the genes that code for putative azoreductase enzymes from the genomic DNA of one the bacterial strains listed in Table 1.2. One strain in particular, *Odoribacter splanchnicus* has generated a lot of interest lately. In an effort to establish the role of *O. splanchnicus* in the gut microbiome, it's presence is suggested to have an influence on some intestinal disorders, such as Crohn's disease (Morgan *et al.*, 2012) and more intriguingly, on some neurological disorders such as ADHD (Wan *et al.*, 2020). As the enzymatic reduction of the azo dye tartrazine by *O. splanchnicus* has not yet been studied, both were chosen as the focus of this project.

Table 1.2 The 22 human gut bacterial strains isolated from stool samples that demonstrated an ability to metabolise azo dyes.

Each isolate was assigned an ID number and once sequenced, was categorised into the correct phylum. Adapted from Zou *et al.*, (2020).

<i>Phylum</i>	<i>Species</i>	<i>Isolate ID</i>
<i>Actinobacteria</i>	<i>Bifidobacterium adolescentis</i>	ADE1
	<i>Bifidobacterium bifidum</i>	ADE2
	<i>Bifidobacterium longum subsp. longum</i>	ADE3
	<i>Bifidobacterium pseudocatenulatum</i>	ADE4
	<i>Collinsella aerofaciens</i>	ADE5
	<i>Collinsella aerofaciens</i>	ADE6
<i>Bacteroidetes</i>	<i>Bacteroides fragilis</i>	ADE7
	<i>Bacteroides ovatus</i>	ADE23
	<i>Bacteroides uniformis</i>	ADE9
	<i>Bacteroides uniformis</i>	ADE10
	<i>Bacteroides vulgatus</i>	ADE11
	<i>Odoribacter splanchnicus</i>	ADE12
<i>Firmicutes</i>	<i>Blautia sp.</i>	ADE13
	<i>Clostridium hylemonae</i>	ADE14
	<i>Clostridium ramosum</i>	ADE26
	<i>Dorea longicatena</i>	ADE16
	<i>Dorea longicatena</i>	ADE17
	<i>Eisenbergiella tayi</i>	ADE18
	<i>Eubacterium cylindroides</i>	ADE19
	<i>Faecalicoccus pleomorphus</i>	ADE20
	<i>Fusicatenibacter saccharivorans</i>	ADE21
	<i>Ruminococcus faecis</i>	ADE22

Table 1.3 The 7 azo dye excipients that were incubated with and reduced by the bacterial strains listed in Table 1.2.

<i>Colour Name, Code (USA)</i>
<i>D&C Brown No. 1</i>
<i>D&C Orange No. 4</i>
<i>D&C Red No. 33</i>
<i>D&C Red No. 6</i>
<i>FD&C Red No. 4</i>
<i>FD&C Red No. 40</i>
<i>FD&C Yellow No. 6</i>

Originally isolated from a human abdominal abscess in 1971 by Werner and Reichertz, *O. splanchnicus* is now a commonly recognised inhabitant of the healthy, human gut microbiome (Göker *et al.*, 2011). It is a Gram-negative, rod shaped, strictly anaerobic bacteria, and the majority of isolated strains highlight its preferential colonization of the human intestine (Pike, 2019; Nagai *et al.*, 2010) and the appendix (Werner and Reichertz, 1975). However, more recent studies have identified isolates of *O. splanchnicus* that are capable of thriving in a variety of extreme environmental conditions, including regions of high salinity (Gulf of Mexico) (Escobedo-Hinojosa and Pardo-López, 2017), extreme cold (Antarctica) (Brambilla *et al.*, 2001) and heavy metal contamination (arsenic-rich rice fields and waste battery dump soil) (Baba, Kimura, Asakawa and Watanabe, 2013; Ogbulie, Okore and Ejele, 2020).

A number of secondary metabolic products generated by *O. splanchnicus* in the human gut, have already been examined in an effort to determine their effects on host health. The SCFAs, sphingolipids and amino acids produced by *O. splanchnicus* have demonstrated a range of beneficial effects within the intestinal environment, such as maintaining mucosal homeostasis (Hiippala *et al.*, 2020), suppressing pro-inflammatory cytokines (Parada Venegas *et al.*, 2019), inducing cytotoxicity in cancerous cells (Xing *et al.*, 2021) and even strongly inhibiting the growth of the potentially opportunistic *Candida albicans* (Seelbinder *et al.*, 2020). However, *O. splanchnicus* can also generate hydrogen sulfide and indoles, which have been strongly implicated with colorectal cancer (Wolf *et al.*, 2020), and negative emotional responses (Jaglin *et al.*, 2018) respectively.

Examining the quantitative role of *O. splanchnicus* in a variety of diverse intestinal microbial communities also yielded contradictory results. Low levels of *O. splanchnicus* have been associated with the presence of disorders such as ileal Crohn's disease, inflammatory bowel disease (IBD) and cystic fibrosis (Morgan *et al.*, 2012; Burke *et al.*, 2017). However, abnormally high levels of *O. splanchnicus* have been observed in individuals with ADHD, a disorder associated with deficiencies in dopamine and noradrenaline (Wan *et al.*, 2020). Interestingly, the findings from Wan's study may reinforce the hypothesis that altered levels of *O. splanchnicus* may contribute to abnormal dopamine metabolism (Quagliariello *et al.*, 2018; Maini Rekdal *et al.*, 2019).

These recent studies illustrate a growing interest in *O. splanchnicus* and the efforts undertaken in order to better understand its function(s) in the gut. However as of now, the full metabolic capabilities of this strain continue to remain largely unknown. The genome of *O. splanchnicus* was sequenced in 2011 by Göker *et al.*, 2011, but many of its annotated enzymes have yet to be characterised. It is important to shed some light on the role this microbial species plays in the gut and identify how its enzymes and subsequent metabolites can mediate xenobiotic metabolism. An understanding of these metabolic processes could allow for more accurately designed risk assessments of the foreign agents that humans are constantly exposed to. Likewise, these pathways could also be exploited to promote the microbial reduction and excretion of hazardous compounds already in the body. They could even to design microbe specific inhibitors of signalling pathways or enzymes that could control the population of opportunistic pathogens or prevent the toxic side effects often accompanying prodrugs (Morais, de Rosso, Estadella and Pisani, 2016; Bhatt *et al.*, 2020).

This project aims to identify and characterise putative azoreductases from the gut microbial strain *O. splanchnicus* to allow us a clearer insight into the metabolic ability of the gut microbiome in response to azo compounds, specifically the food dye tartrazine, which is still widely consumed across the world. It would also allow us to further hypothesise on the primary anaerobic function of microbial azoreductases in the human gut by identifying the interactions between these enzymes and synthetic food additives that could prove detrimental to human health.

Chapter 2 Materials and methods

2.1 Chemicals and reagents

The chemicals and reagents used in this study were provided by Sigma-Aldrich unless otherwise stated. DNA and plasmid purification and extraction kits were obtained from New England Biolabs (NEB). Champion™ pET100 Directional TOPO™ Expression Kit and BL21 Star™ (DE3) One Shot™ Chemically Competent *E. coli* were supplied by Invitrogen, Thermo Fisher Scientific. DNA primers were manufactured and supplied by Merck.

2.2 Bacterial Strains and growth conditions

The genomic DNA of the bacterial strain *O. splanchnicus* 22-5 was kindly provided by Dr. Emma Allen-Vercoe, University of Guelph, Canada.

E. coli TOP10 and BL21 Star™ (DE3) cells were provided as part of the Champion™ pET100 and the Champion™ pET102 Directional TOPO™ Expression Kits from Invitrogen, Thermo Fisher Scientific.

Growth medias used throughout this study were prepared according to the manufacturer's guidelines and sterilised following standard sterilisation parameters of 121°C for 15 minutes. Details of growth medias are provided in Table 2.1. Overnight bacterial cultures were prepared in 50mL Falcons containing 5mL of Luria Bertani (LB) broth and 100 µg/mL of ampicillin (Fisher Scientific). Cultures were then incubated overnight at 37°C with a speed of 120 rpm in an INFORS AG Incubator Shaker CH4103. For selection of *E. coli* BL21 Star™ (DE3) cells transformed with a pET TOPO® vector, cells were incubated on LB agar plates containing 100 µg/ml of ampicillin.

Table 2.1 Growth media

Table illustrating the types of growth media used, their components and concentrations, and the company each media was supplied by.

Medium	Components	Company
Luria Bertani Broth	10 g/L Tryptone 5 g/L Yeast Extract 5 g/L NaCl	Fisher Bioreagents
Luria Bertani Agar	10 g/L Tryptone 5 g/L Yeast Extract 5 g/L NaCl 15 g/L Agar	Fisher Bioreagents
Terrific Broth	12 g/L Tryptone 24 g/L Yeast Extract 12.5 g/L K ₂ HPO ₄ 2.3 g/L KH ₂ PO ₄	Fisher Bioreagents
2x YT	16 g/L Tryptone 10 g/L Yeast Extract 5 g/L NaCl	Sigma-Aldrich
S.O.C. Medium	20 g/L Tryptone 5 g/L Yeast Extract 0.58 g/L NaCl 2.5 mM KCl 10 mM MgCl ₂ 10 mM MgSO ₄ 20 mM Glucose	Invitrogen, Thermo Fisher Scientific

2.3 Genomic DNA quantification

The concentration, yield and purity of the genomic DNA from *O. splanchnicus* was determined using NanoVue™ Plus spectrophotometers (VWR).

2.4 Bioinformatics

A BLASTp analysis was performed using a set of characterised azoreductase gene sequences against the genomes of all strains of *O. splanchnicus* 22-5 in the NCBI database. Two parameters most commonly used when selecting the most appropriate gene sequences, are the expectation or E-value and the sequence identity percentage, the former of which measures the significance of the match while the latter measures the percentage of identical residues between the two proteins (Rost, 1999). Homologous sequences with the lowest e-value, that were above a cut off of 30% sequence identity, were selected and primers were designed based on these sequences. Primer design requirements outlined in the Champion™ Expression kit manuals were adhered to as well as using the Integrated DNA Technologies Oligoanalyzer tool (Owczarzy *et al*, 2008) to analyse GC content, predicted melting temperatures and hairpin formation of the chosen primers. All primer sequences are listed in Table 2.2.

2.5 Gene amplification

Using the primers designed in Section 2.4, polymerase chain reaction (PCR) was performed using the genomic DNA of *O. splanchnicus* 22-5 as template to amplify the sequences of putative azoreductase genes. Table 2.3 lists the components used in each PCR reaction. The parameters for thermal cycling were specified in the manual for the Q5 High-Fidelity 2X Master Mix (NEB) with the exception of the annealing temperature which was identified by carrying out PCR with annealing temperatures ranging from 55°C - 60°C. Thermocycling parameters are displayed in Table 2.4.

Table 2.2 Primers used for the amplification of putative *O. splanchnicus* 22-5 azoreductase gene

The sequences of both the forward and reverse primers used to amplify putative azoreductase genes from the genome of *O. splanchnicus* 22-5. Sequence design differs slightly according to requirements outlined with each expression kit (Thermo Fisher Scientific).

Expression Kit (Vector ID)	Bacterial Strain	Primer	Sequence
pET102/D-TOPO®	<i>O. splanchnicus</i> 22-5	2019 O.spp FWD	CACCATGAAAGTCGTCTTAATAAATGG
		2019 O.spp REV	TCGTATAAAATTAGTCATGACCC
pET100/D-TOPO®	<i>O. splanchnicus</i> 22-5	2020 O.spp FWD	CACCATGAAAGTCGTCTTAATAAA
		2020 O.spp REV	TCATCGTATAAAATTAGCATGACCCG

Table 2.3 PCR reaction mix

The components of a 50µL PCR reaction mixture and the volumes and concentrations of each.

Component	Volume (µl)	Final Concentration
Genomic DNA	1	60-100ng
Forward Primer	5	0.5µM
Reverse Primer	5	0.5µM
Q5 High-Fidelity 2X Master Mix	25	1x
H ₂ O	14	-

Table 2.4 Thermocycling parameters used for the amplification of *O. splanchnicus* 22-5

Thermocycling parameters for each stage of PCR that were followed to amplify putative azoreductase gene of *O. splanchnicus* 22-5.

Step	Temperature (°C)	Time	No. of cycles
Initial denaturation	98°C	30 seconds	1
Denaturation	98°C	10 seconds	35
Annealing	55°C - 60°C	30 seconds	
Extension	72°C	30 seconds	

2.5.1 Agarose gel analysis of PCR products

Analysis of PCR products was carried out using gel electrophoresis on an agarose gel consisting of 1% w/v agarose (Fisher Biologics) in 1x Tris Borate EDTA (TBE) buffer, pH 8 (Table 2.5) and SYBR safe (Life Technologies). Samples were prepared by adding Gel Loading Dye (6x) (NEB) to 10 μ L of PCR sample (5:1), reserving the remaining 40 μ L of PCR sample to be purified (Section 2.5.2). Ten microlitres of each sample was loaded into agarose gel wells accompanied by 6 μ L of Quick-Load 1 kb Plus DNA Ladder (NEB). Gels were run in a mini horizontal electrophoresis chamber (BIO-RAD) with 1x TBE as the running buffer. Electrophoresis was performed for 45 minutes at 80 volts (BIO-RAD PowerPac basic). Gels were analysed in a Molecular Imager[®] Gel Doc[™] XR System (BIO-RAD) and DNA ladder was referenced to estimate the size (or molecular weight) of DNA molecules that appeared in/as a band in the gel.

2.5.2 Purification and quantification of PCR products

PCR samples were purified using the Monarch[®] PCR & DNA Cleanup Kit (NEB) according to the manufacturer's instructions. The concentration and purity of DNA in each sample was measured as described in Section 2.3.

2.6 Cloning of putative azoreductase gene

Purified PCR products were TOPO[®] Cloned into a pET TOPO[®] vector following the procedure recommended by the supplier. Table 2.6 lists the components of each cloning reaction and their concentrations. Each reaction was performed using a 1:1 molar ratio of PCR product:TOPO[®] vector. Samples were incubated for 5 minutes at room temperature.

Table 2.5 The reagents used to prepare a 1x TBE buffer for agarose gel electrophoresis

The components of the 1x TBE buffer used in the preparation of a 1% w/v agarose gel and as a running buffer during electrophoresis, the concentration of each reagent and the company they were supplied by.

<i>Buffer</i>	<i>Reagent</i>	<i>Concentration</i>
<i>1x TBE pH 8</i>	<i>Tris Base</i>	<i>130 mM</i>
	<i>Boric Acid</i>	<i>45 mM</i>
	<i>EDTA (disodium salt)</i>	<i>2.5 mM</i>

Table 2.6 TOPO® cloning reaction mix. Adapted from (Thermo Fisher Scientific)

The reagents used in a 6µL TOPO® cloning reaction mix, their components and concentrations, and the company they sourced from.

<i>Reagent</i>	<i>Components</i>	<i>Concentration</i>	<i>Volume</i>	<i>Manufacturer</i>
<i>Purified PCR product</i>	-	<i>3 ng/µL</i>	<i>1 µL</i>	-
<i>Salt solution</i>	<i>NaCl</i>	<i>1.2 M</i>	<i>1 µL</i>	<i>Invitrogen, Thermo Fisher Scientific</i>
	<i>MgCl₂</i>	<i>0.06 M</i>		
<i>Sterile H₂O</i>	-	-	<i>3 µL</i>	-
<i>TOPO® Vector</i>	<i>Plasmid DNA</i>	<i>15-20 ng/µL</i>	<i>1 µL</i>	<i>Invitrogen, Thermo Fisher Scientific</i>
	<i>Glycerol</i>	<i>50%</i>		
	<i>Tris-HCl</i>	<i>50 mM</i>		
	<i>EDTA</i>	<i>1 mM</i>		
	<i>DTT</i>	<i>2 mM</i>		
	<i>Triton X-100</i>	<i>0.1%</i>		
	<i>BSA</i>	<i>100 µg/mL</i>		
<i>Bromophenol blue</i>	<i>30 µM</i>			

2.7 Transformation

All ligation samples (see Section 2.6) were transformed into competent *E. coli*. One Shot® TOP10 cells using the heat shock method outlined in the supplier's protocol (Thermo Fisher Scientific). Three microlitres of a TOPO® Cloning reaction was added to a vial of chemically competent *E. coli* TOP10 cells that had been defrosted on ice. Transformation mixtures were then incubated on ice for 15 minutes. A negative sample was always prepared, containing *E. coli* cells only. Cells were then heat shocked for 30 seconds at 42°C. Following this, cells were recovered by adding 250µl of warmed S.O.C medium and incubating for an hour at 37°C, shaking horizontally at 200 RPM in an INFORS AG Incubator Shaker CH4103. Two hundred microlitres of each transformation sample was spread onto a pre-warmed selective LB agar plate (Section 2.2) and incubated at 37°C overnight. For each bacterial strain, four colonies were picked and each immersed in 10µl of elution buffer from the Monarch® PCR & DNA Cleanup Kit (NEB). Colony samples were then stored at 4°C to be used either as a template for colony PCR (Section 2.7.1) or for plasmid extraction (Section 2.7.2).

2.7.1 Analysing successful transformants

Colony PCR was performed with 2.5µl of resuspended colony sample (Section 2.7) to identify positive clones. The primers used were designed and supplied by Thermo Fisher Scientific (Table 2.7). PCR reaction mix components and their volumes are listed in Table 2.8 and thermocycling conditions are described in Table 2.9. PCR products were then analysed as described in Section 2.5.1. Colonies that contained the pET TOPO® vector with gene insert were grown in 5ml of LB Broth inoculated 100 µg/mL ampicillin, until early/mid stationary phase was reached. Eight hundred and fifty microlitres of each culture sample was diluted with 150µL of Sterile 99% Glycerol (Alfa Aesar) to give a final concentration of 15% (v/v). Cultures were stored at -80°C.

Table 2.7 The primers used to identify successful transformants

The forward and reverse primers used to confirm that the pET TOPO® vector contains the gene insert. Forward primers were chosen according to the vector that was used in the transformation process. All primers were supplied by Thermo Fisher Scientific.

<i>Vector</i>	<i>Primer</i>	<i>Sequence</i>	<i>Concentration</i>
<i>pET100/D-TOPO®</i>	<i>T7</i>	<i>TAATACGACTCACTATAGGG</i>	<i>10µM</i>
<i>pET102/D-TOPO®</i>	<i>TrxFus Forward</i>	<i>TTCCTCGACGCTAACCTG</i>	
<i>pET100/D-TOPO®</i> <i>and</i> <i>pET102/D-TOPO®</i>	<i>T7 Reverse</i>	<i>TAGTTATTGCTCAGCGGTGG</i>	

Table 2.8 Colony PCR reaction mix

The components of a 25µL colony PCR reaction mix and the volume and concentration of each.

<i>Component</i>	<i>Volume (µL)</i>	<i>Concentration</i>
<i>Template DNA</i>	<i>2.5</i>	<i>Unknown</i>
<i>Forward Primer</i>	<i>1.25</i>	<i>10µM</i>
<i>Reverse Primer</i>	<i>1.25</i>	<i>10µM</i>
<i>Q5 High-Fidelity 2X Master Mix</i>	<i>12.5</i>	<i>1x</i>
<i>H2O</i>	<i>7.5</i>	<i>-</i>

Table 2.9 Thermocycling parameters used during colony PCR

Thermocycling parameters for each stage of PCR that were followed to confirm colonies had been successfully transformed with pET TOPO® containing the gene of interest.

<i>Step</i>	<i>Temperature (°C)</i>	<i>Time</i>	<i>No. of cycles</i>
<i>Initial denaturation</i>	<i>98°C</i>	<i>30 seconds</i>	<i>1</i>
<i>Denaturation</i>	<i>98°C</i>	<i>10 seconds</i>	<i>25</i>
<i>Annealing</i>	<i>58°C</i>	<i>30 seconds</i>	
<i>Extension</i>	<i>72°C</i>	<i>60 seconds</i>	

2.7.2 Plasmid extraction

Overnight cultures inoculated with successfully transformed *E. coli*. One Shot® TOP10 cells (Section 2.7.1) were prepared as described in Section 2.2. Plasmids were purified from overnight cultures using the Monarch® Plasmid Miniprep Kit (NEB) and following the instructions supplied with the kit. Plasmid samples were then analysed as described in Section 2.3.

2.7.3 DNA sequencing

Purified plasmid samples combined with corresponding primers used in Table 2.7 were sent to Genewiz for DNA sequence analysis. Sequencing results of each bacterial strain were then aligned against the original gene sequence that the primers had been designed to amplify (Figure 3.5). ClustalW multiple alignment (Hall, 1999) was performed using BioEdit, on sequencing results to confirm that the plasmid contained the correct sequence. Jalview (Waterhouse *et al.*, 2009) was used to generate visual images of multiple alignment sequence results.

2.8 Transformation

Purified plasmid template from Section 2.7.2 was transformed into BL21 Star™ (DE3) cells following the relevant transformation protocol outlined in the supplier's manual. The protocol was slightly modified so that the transformation reaction could be plated instead of inoculating into medium for subsequent expression. In order to do this, the method outlined in section 2.7 was adhered to once BL21 had undergone the heat shock process. This allowed successful clones to be cultured and stored as described in section 2.7.1. Glycerol stocks could then be used for repeated protein expression trials.

2.9 Recombinant protein expression

Starter cultures to be used for all protein expression trials were inoculated with transformed BL21 Star™ (DE3) cells (Section 2.8) and then prepared as described in Section 2.2.

2.9.1 Small scale recombinant protein expression

Starter culture was expanded by using 500 μ L of starter culture (Section 2.9) to inoculate 5ml of LB broth containing 100 μ g/mL ampicillin in three sterile 50mL Falcon tubes. These were then incubated with shaking at 37°C with a speed of 180 rpm in an INFORS AG Incubator Shaker CH4103 until an OD₆₀₀ of between 0.5 – 0.8 was reached. To induce protein expression, isopropyl β - d-1-thiogalactopyranoside (IPTG) was then added, with each falcon containing a different IPTG concentration (0.1mM, 0.5mM or 1M respectively) to determine optimal IPTG induction concentration.

After induction, a 500 μ l aliquot was removed from each culture every hour for 4 hours and then once more after overnight incubation. This was used to determine the optimal induction time. To collect the cells, each aliquot was immediately centrifuged at 16,000 x g in a microcentrifuge for 30 seconds at 4°C, the supernatant discarded, and the remaining cell pellets stored at -20°C until ready to lyse.

The above method was repeated, using 500 μ l of starter culture to inoculate all growth media listed in Table 2.1, to also determine optimal growth media.

After centrifugation, cells were lysed by resuspending pellets in 300 μ l of BugBuster[®] Protein Extraction Reagent (Millipore). The cell lysate was subsequently centrifuged for 20 minutes at 16,000 x g, 4°C to pellet the insoluble proteins. Fifty microlitres of the supernatant was then transferred to a fresh tube stored on ice and mixed together with an equal amount of 2x Laemmli sample buffer (LSB) (Table 2.10). Insoluble pellets were dissolved in 300 μ L of 1x LSB. All samples were then boiled at 95°C for 5 minutes to denature proteins before analysing via SDS-PAGE (Section 2.9.3).

2.9.2 Large scale recombinant protein expression

Once optimal growth media, induction time and IPTG concentration had been determined and soluble protein was successfully acquired, recombinant protein expression was then carried out using larger culture volumes. The method employed was similar to that in Section 2.9.1 with the following modifications: 100mL of growth media containing 100 µg/mL ampicillin was inoculated with 1ml of starter culture in a bevelled Erlenmeyer flask. Incubation parameters for starter expansion remained the same. Protein expression was induced using the optimal IPTG concentration, time and temperature measurements that had been determined in section 2.9.1.

Cells were collected by centrifuging the full volume of bacterial culture using the RC6 plus centrifuge (Sorvall) at 6000 x g for 20 minutes at 4°C. Collected cells were then lysed with 3 ml of BugBuster® Protein Extraction Reagent (Millipore), followed by a 20 minute period of centrifugation at 16,000 x g, 4°C, and the soluble fraction was separated from the insoluble fraction and prepared for analysis as described in Section 2.9.3.

2.9.3 SDS-PAGE analysis of all recombinant proteins

All soluble and insoluble fractions obtained from the protein expression trials (Sections 2.9.1 and 2.9.2) were analysed by SDS-PAGE on a 12% SDS gel (Table 2.10).

Ten microlitres of the soluble fraction and 5 µl of the insoluble fraction were loaded onto an SDS-PAGE gel. 5 µl of Thermo Scientific™ PageRuler™ Prestained Protein Ladder, 10 to 180 kDa was also loaded onto each gel. Electrophoresis was applied by running gels in a MiniProtean Tetra system (BIO-RAD) with a 1x-SDS running buffer, applying 60 V for 30 minutes and then 180 V for 45 minutes.

Gels were then stained in 100 ml of Coomassie staining buffer (Table 2.10) by heating for 10 second intervals until just boiled and then incubating for 30 minutes with rocking at 120 RPM in room temperature. Destaining was carried out by discarding staining buffer and adding 125 ml of destaining buffer, heating for 10 second intervals until just boiled and incubating with rocking at room temperature for 45 minutes. Destaining was repeated once more. All materials used in SDS-PAGE are specified in Table 2.10.

Table 2.10 Materials used for SDS-PAGE and the staining and de-staining of SDS gels. Part adapted from (Holland, 2017)

The materials used to prepare, run, stain and destain a 12% SDS gel, the reagents and their concentrations and the companies that supplied them.

<i>Material</i>	<i>Reagent</i>	<i>Concentration</i>
<i>Resolving gel</i>	<i>Acrylamide:bis acrylamide (29:1)</i>	<i>12% v/v</i>
	<i>SDS</i>	<i>0.125% w/v</i>
	<i>Ammonium persulphate</i>	<i>0.05% w/v</i>
	<i>Tetramethylethylenediamine (TEMED)</i>	<i>0.002% v/v</i>
	<i>Tris HCl</i>	<i>390 mM</i>
<i>Stacking gel</i>	<i>Acrylamide:bis acrylamide (29:1)</i>	<i>6% v/v</i>
	<i>SDS</i>	<i>0.125% w/v</i>
	<i>Ammonium persulphate</i>	<i>0.05% w/v</i>
	<i>Tetramethylethylenediamine (TEMED)</i>	<i>0.002% v/v</i>
	<i>Tris -HCl</i>	<i>116.6 mM</i>
<i>2x Laemmli sample buffer (LSB)</i>	<i>Tris -HCl</i>	<i>277.8 mM</i>
	<i>Glycerol</i>	<i>22.2% v/v</i>
	<i>SDS</i>	<i>2.2% w/v</i>
	<i>Bromophenol blue</i>	<i>0.01% w/v</i>
	<i>Beta-mercaptoethanol</i>	<i>10% v/v</i>
<i>SDS running Buffer</i>	<i>Tris- HCl</i>	<i>0.25 M</i>
	<i>Glycine</i>	<i>1.95 M</i>
	<i>SDS</i>	<i>1% w/v</i>
<i>Staining buffer</i>	<i>Coomassie brilliant blue</i>	<i>0.024% w/v</i>
	<i>Methanol</i>	<i>50% v/v</i>
	<i>Acetic Acid</i>	<i>10% v/v</i>
	<i>H₂O</i>	<i>40% v/v</i>
<i>Destaining buffer</i>	<i>Methanol</i>	<i>10% v/v</i>
	<i>Acetic Acid</i>	<i>10% v/v</i>
	<i>H₂O</i>	<i>80% v/v</i>

Chapter 3 Cloning of putative azoreductases genes from the genomic DNA of *O. splanchnicus*.

3.1 Introduction

The identification and characterisation of gut bacterial enzymes allows for further insight into their function, metabolic pathways and subsequent products, which allow us to understand the type of relationship microorganisms share with their host. It is widely accepted that the mammalian intestinal environment has a heavy influence on the metabolism of many xenobiotic compounds (Pfeiffer and Sonnenburg, 2011). Harboured here are a considerable number of enzymes, encoded by microbial genes so diverse and widespread that they outnumber our own by a factor of over 100 (Qin *et al.*, 2010). One such enzyme that catalyses the reduction of the azo bond, the gut bacterial azoreductase, has long been recognised and exploited in the design of various pro-drugs (Colebrook, 1936). However, since its discovery, knowledge on the genomics, evolutionary history and function of azoreductases has advanced slowly, evident by the limited number of enzymes that have been isolated and characterised to date (Table 3.1). Even fewer again have been characterised specifically from the human gut microbiota.

Owing to its natural residence in the human gut (Werner and Reichertz, 1975), we did not have the facilities to anaerobically culture *O. splanchnicus*. Thus, in order to identify putative azoreductases from its genome, a homology-based genome-wide bioinformatics approach was used, followed by a combination of molecular cloning techniques to further isolate and amplify genes. Finally, an appropriate cloning kit enabled the insertion of genetic material into a suitable vector that could be transformed into a host cell selected for its ability to efficiently express the protein of interest.

The Champion™ pET Directional TOPO™ Expression Kit employs a cloning strategy that eliminates the use of restriction enzymes and ligases. Instead, the enzyme Vaccinia virus-isolated topoisomerase I is exploited to first digest the primers on the amplified gene of interest at specifically designed overhangs before finally ligating the gene to the vector (Shuman, 1994). The overhangs are specifically designed as outlined by the manual supplied by Thermofisher (pBAD/Thio His TOPO manual, 2010), to ensure directional cloning of the gene in the correct orientation, with a success rate of over 90%. (Studier, 1991). The kit is

specifically designed to ensure a more rapid process of cloning putative azoreductase genes, with the entire process amounting to less than 5 minutes, all the while maintaining its ability to yield a high number of successful clones, establishing this kit as an attractive, highly efficient choice for this procedure.

Table 3.1 A compilation of the bacterial azoreductases that have been characterised to date.Adapted from Misal *et al.*, 2018.

Organism	Aerobic/ Anaerobic	Co-Factor	Molecular mass (kDa)
<i>Pseudomonas</i> KF46	Aerobic	NAD(P)H	32
<i>Escherichia coli</i>	Aerobic	FMN/NADH	23
<i>Bacillus</i> sp. OY1-2	Aerobic	NAD(P)H	20
<i>Xenophilus azovorans</i> KF46F	Aerobic	NAD(P)H	30
<i>Pigmentiphaga kullae</i> K24	Aerobic	NAD(P)H	20.6
<i>Rhodobacter sphaeroides</i> AS1.1737	Aerobic	NADH/NADPH	18.7
<i>Enterococcus faecalis</i>	Aerobic	FMN/NADH	23
<i>Pseudomonas aeruginosa</i>	Aerobic	NADH	29
<i>Staphylococcus aureus</i>	Aerobic	FMN/NADPH	85 (tetramer)
<i>Bacillus</i> sp.	Aerobic	NADH	23
<i>Pseudomonas aeruginosa</i> PAO1	Aerobic	NAD(P)H	23 (tetramer)
<i>Pigmentiphaga kullae</i> K24	Aerobic	NADPH	22
<i>Geobacillus stearothermophilus</i>	Aerobic	FMN/NADH	23
<i>Bacillus badius</i>	Aerobic	NADH	43
<i>Bacillus lentus</i> BI377	Aerobic	NAD(P)H	32
<i>Aquiflexum</i> sp.	Aerobic	NADH/NADPH	80
<i>Pseudomonas putida</i> MET94	Aerobic	NAD(P)H	21
<i>Brevibacillus laterosporus</i> TISTR1911	Aerobic	FMN/NADH	23
<i>Lysinibacillus sphaericus</i>	Aerobic	NADH	29
<i>Shewanella xiamenensis</i> BC0	Aerobic	NAD(P)H	26
<i>Halomonas elongata</i>	Aerobic	NADH	22
<i>Rhodococcus opacus</i> ICP	Aerobic	NAD(P)H	25
<i>Clostridium perfringens</i>	Anaerobic	FMN/NADH	90.4(monomer)
<i>Shewanella oneidensis</i> MR-1	Anaerobic	NADH	27

3.2 Results

3.2.1 Bioinformatics

The protein sequence of the identified bacterial azoreductase AzoC from *C. perfringens* was blasted against the genome sequence of *O. splanchnicus* DSM 20712 strain using NCBI's Basic local alignment search tool (BLAST).

This generated one protein result with 51% sequence identity (accession number WP_013611494). The nucleotide sequence of this result was obtained, upon which the primers for PCR amplification were designed upon.

3.2.2 Amplification of putative azoreductase genes

Using the genomic DNA of *O. splanchnicus* 22-5 strain and its corresponding designed primers (Table 2.2), putative azoreductase genes were amplified by PCR. Annealing temperature trials were screened to determine the optimum temperature so that primers may bind specifically to the template strand.

The first annealing temperature trial using primers designed for annealing with pET102 (Figure 3.1) was successful, where annealing temperatures of 60°C (lane 2) and 55°C (lane 3) produced significant PCR product with the expected molecular weight of 0.66 kB.

Similarly, the primers designed to anneal to the pET100 vector, did so under all applied annealing temperatures (Figure 3.2 A) however, bands were noticeably blurry so, PCR with an annealing temperature of 58°C was repeated to confirm successful amplification (Figure 3.2 B).

All PCR products were then purified and quantified prior to cloning into a pET TOPO[®] vector.

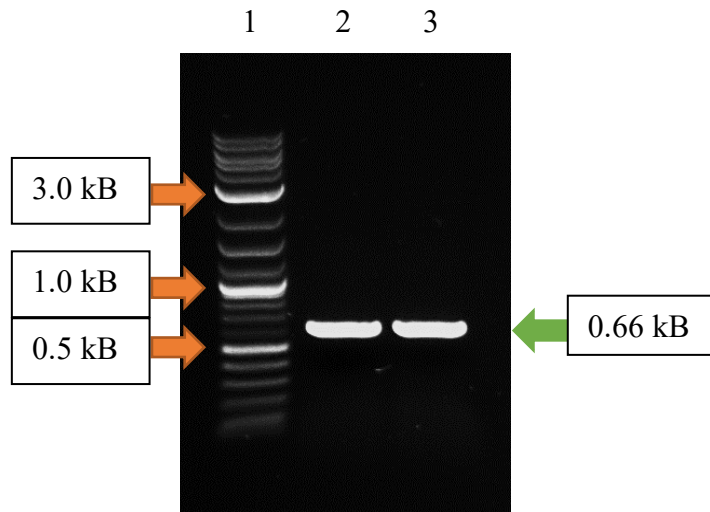


Figure 3.2 PCR amplification of putative azoreductase genes from *O. splanchnicus* using varying annealing temperatures and pET102 designed primers.

1% w/v agarose gel displays PCR products in lane 2 (60°C annealing temperature) and lane 3 (55°C annealing temperature) with an expected size of 0.66 kB. 1 kb Plus DNA Ladder was used as a reference (lane 1).

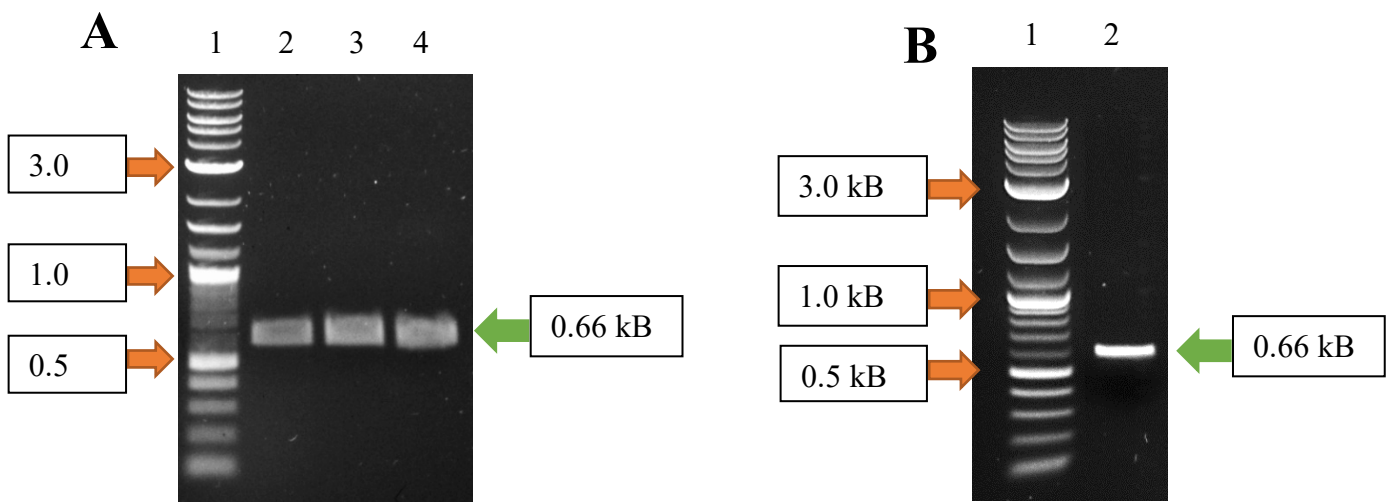


Figure 3.7 PCR amplification of putative azoreductase genes from *O. splanchnicus* using varying annealing temperatures and pET100 designed primers.

1% w/v agarose gels display PCR products with an expected size of 0.66 kB. 1 kb Plus DNA Ladder was used as a reference (lane 1). **(A)** Annealing temperature trials yielded PCR product at all temperatures. Annealing temperatures of 60°C (lane 2), 58°C (lane 3) and 55°C (lane 4) were applied. **(B)** Confirmation that an annealing temperature of 58°C (lane 2), successfully produced PCR product of expected size during PCR amplification.

3.2.3 Cloning and transformation of putative azoreductase gene

Successfully purified PCR products were inserted into either a pET102 or pET100 vector before being transformed into competent *E. coli*. One Shot® TOP10 cells. A negative transformation sample where both vector and PCR product were omitted, was also prepared and used as a control. LB agar plates prepared with 100 µg/mL ampicillin were used to incubate each transformation overnight to allow for selection of positive transformants. Colonies were picked to undergo colony PCR to confirm successful transformation and products were analysed on a 1% w/v agarose gel. PCR amplicons with an expected size of 0.9 kB were identified in lane 4 and 5 of Figure 3.3 and in lane 2 of Figure 3.4, confirming effective cloning of PCR product into the pET102 vector and pET100 vector respectively. Successfully transformed *E. coli*. One Shot® TOP10 cells were cultured overnight, from which plasmids were purified and sequenced. The resulting nucleotide sequences provided by Genewiz were aligned against the original gene sequence (Table 2.2) using the Clustal W software provided by BioEdit (Hall, 1999). The resulting alignment revealed correct orientation of insert and no point mutations (Figure 3.5), demonstrating that the cloning procedure was successful.

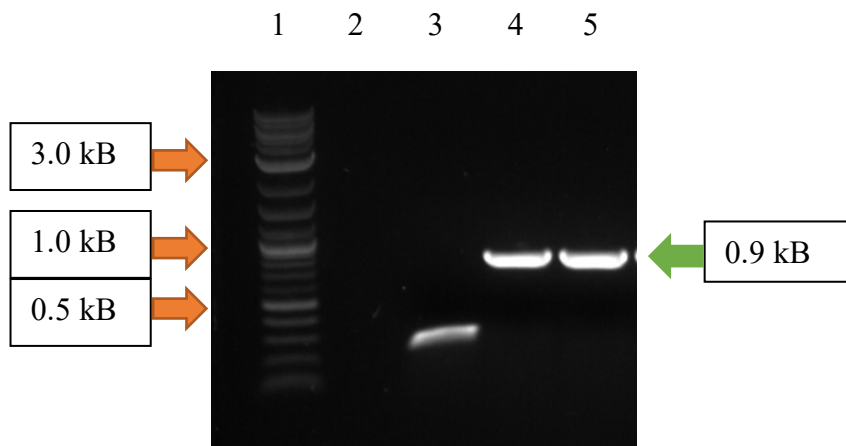


Figure 3.12 Colony screening for the successful transformation of ligated PCR product and pET102 vector into *E. coli*. One Shot® TOP10 cells.

1% w/v agarose gels indicate successful ligation and subsequent transformation in lane 4 and lane 5, where PCR product with an expected size of 0.9 kb can be observed. 1 kb Plus DNA Ladder was used as a reference (lane 1).

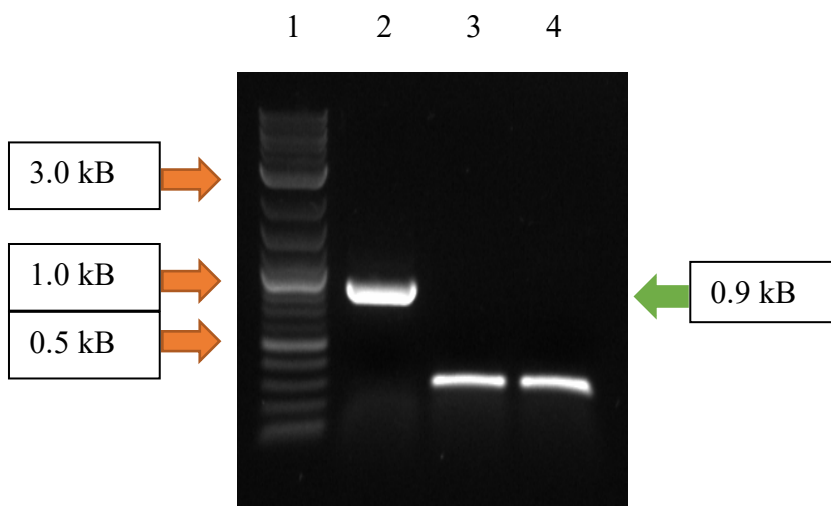


Figure 3.17 Colony screening for the successful transformation of ligated PCR product and pET100 vector into *E. coli*. One Shot® TOP10 cells.

1% w/v agarose gels indicate successful ligation and subsequent transformation in lane 1, where PCR product with an expected size of 0.9 kb can be observed. 1 kb Plus DNA Ladder was used as a reference (lane 1).

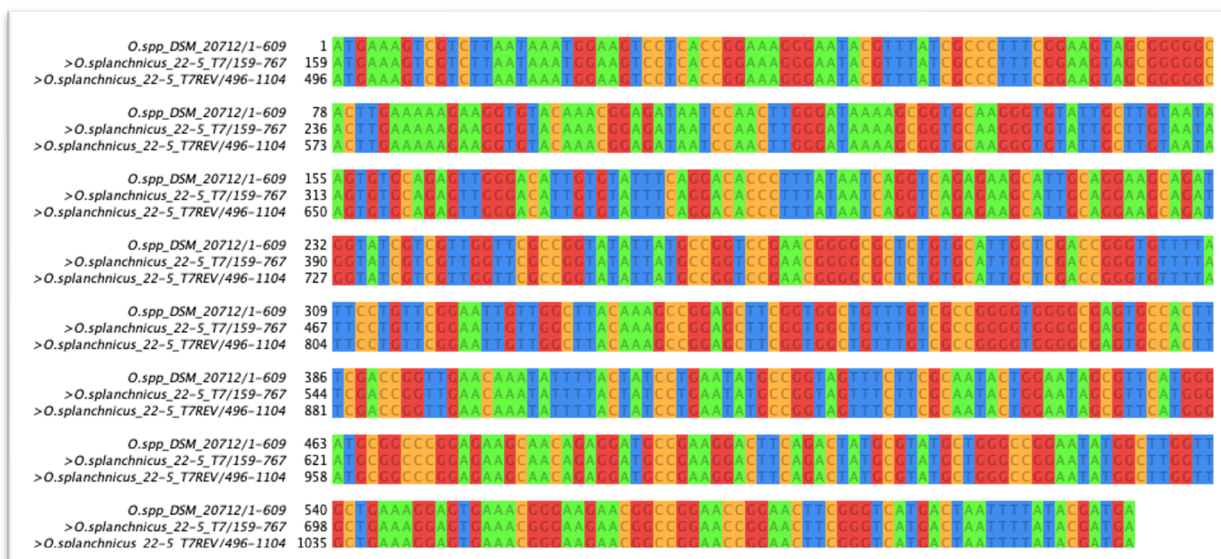


Figure 3.22 Alignment of the forward (T7) and reverse (T7REV) nucleotide sequences of a putative azoreductase gene cloned from the *O. splanchnicus* 22-5 strain, against the original genomic sequence of *O. splanchnicus* DSM 20712. The forward and reverse nucleotide sequences were provided by Genewiz. The genomic sequence of *O. splanchnicus* DSM 20712 was obtained from the NCBI database.

3.3 Discussion

As previously mentioned, the total number of characterised azoreductases is very minute. From the enzymes listed in Table 3.1, only three have been fully characterised from bacterial strains originating in the GI tract; AzoR from *E. coli* (Nakanishi *et al*, 2001), AzoA from *Enterococcus faecalis* (Chen *et al*, 2004) and AzoC from *C. perfringens* (Morrison *et al*, 2012). As the bacterium *O. splanchnicus* had similarly been isolated from the human intestine, it is not surprising to observe a sequence similarity of 51% when the protein sequence of AzoC is aligned against the genomic DNA of the DSM 20712 strain of *O. splanchnicus* as both strains are strictly anaerobic. Azoreductases are often classified into two major groups based on their oxygen tolerance (Rafii and Coleman 1999), whereby it is hypothesised that the sequence and structure of these enzymes could be influenced by a bacterium's sensitivity to oxygen (Misal and Gawai, 2018)

PCR amplification is a powerful molecular technique that is still widely used for the detection and quantification of DNA fragments (Garibyan and Avashia, 2013; Mullis, 1990). As its success depends on the specific binding between primers and its targeted DNA sequence, the design of suitable primers and optimization of the annealing temperature (Ta) are the two most crucial factors to consider before performing this technique (WU *et al.*, 1991). An unsuitable Ta can result in poor yield of PCR product, either due to amplification of non-specific DNA fragments where the Ta is too low or a decrease in purity of product resulting from a Ta that is too high (Rychlik, Spencer and Rhoads, 1990).

The primers used in this study were designed upon the gene hypothesised to code for an azoreductase from the genomic DNA of *O. splanchnicus* DSM 20712. The Ta was optimised by applying temperatures of between 55 °C and 65 °C to identify the temperature at which the highest amplification was achieved. The combination of primers designed to anneal to either the pET100 or the pET102 vector (Table 2.7) with an annealing temperature of 58°C (Figure 3.2B) and 60°C (Figure 3.1) respectively, successfully amplified a gene from the genomic DNA of an unknown strain of *O. splanchnicus*, generating a PCR product of 0.66 kB. As such, these results outline the optimal PCR conditions required specifically for amplification of a putative azoreductase gene from *O. splanchnicus*.

Subsequent cloning of this gene into expression vector pET100 or pET102 and its transformation into the host *E. coli*. One Shot® TOP10 cell was equally successful, where the resulting construct was identified as a 0.9 kB amplicon under the application of confirmational colony PCR.

Finally, this construct was purified and its nucleotide chain/structure was sequenced in both a forward and reverse direction by Genewiz. A ClustalW alignment of the translated amino acid sequence revealed an absence of point mutations and further confirmed the uptake of the target gene sequence in the correct orientation, demonstrating the success of the selected pET directional TOPO cloning kit and of the modified amplification and cloning procedures.

The next chapter will detail the methodology undertaken in order to optimise expression of this protein.

Chapter 4 Optimising the accumulation of soluble recombinant protein by performing expression trials.

4.1 Introduction

In order to verify that the gene that was amplified and cloned could heterologously express an enzyme with azoreductase activity, a recombinant protein expression system was designed. Once optimised, this system can be a highly efficient, inexpensive method of acquiring recombinant protein, however it is not without its challenges. The most important aims include achieving a high protein yield, and obtaining an amount of soluble protein suitable for its subsequent purification and characterisation. Several factors can limit the production of both, such as choice of expression vector, host cell line and culture conditions (Baneyx and Mujacic, 2004).

E. coli is the most well characterised and widely used expression host for the overexpression of heterologous proteins (Samuelson, 2010). Its popularity is defined by the ease and speed at which this bacterium can be cultured and expressed, alongside the wide range of genetic modifications it offers, and the low costs that are involved. (Francis and Page, 2010). The most attractive advantage this expression system offers is the variety of modified strains and complementary expression vectors & fusion tags that are specifically designed to increase the potentiality of successful soluble protein expression (Costa, Almeida, Castro and Domingues, 2014).

For this study, the DE3 strain of *E. coli* BL21 was selected, accompanied by the pET directional TOPO expression kit, an equally popular expression vector system designed to improve the synthesis and accumulation of soluble recombinant protein (Sokolosky and Szoka, 2013). In this study, both the pET100 and pET102 vectors were specifically selected to trial the expression of the purified, putative azoreductase gene. Additionally, the 6x His tag and the His Patch thioredoxin (TRX) protein were utilised to further increase soluble protein accumulation. The TRX protein in particular, has demonstrated great success in improving protein solubility, sometimes maintaining up to 250 fold of soluble protein in *E. coli*'s cytosolic environment, achieved by allowing for the proper formation of disulphide bonds in newly expressed proteins (Lee *et al.*, 2008).

4.2 Results

4.2.1 Expression of putative azoreductase enzymes using the pET102 and pET100 expression vectors.

Purified plasmid that yielded the desired gene sequence in Section 3.2.3 was transformed into BL21 Star™ (DE3) cells from which a starter culture was prepared, scaled up, induced with IPTG and incubated in a growth media of choice, selected from Table 2.1. IPTG concentration, induction times and temperature also varied, and are explored in more detail below. The cell lysates obtained from each test culture were then analysed via SDS-PAGE to evaluate the proportion of protein present in both the soluble and insoluble fractions.

4.2.2 Comparison on the efficacy of the pET102 and pET100 expression vectors in obtaining soluble protein.

Upon induction of cultures, a high level of expression was observed in *E. coli* BL21 (DE3) cells transformed with either pET vector. This is represented by prominent bands at roughly 24 kDa for cultures containing the pET100 vector, while cultures containing the pET102 vector were represented by bands with an increased molecular mass of 37 kDa, in order to account for the addition of the TRX tag. (Figure 4.1A and 4.1B.).

Recombinant protein expressed from the pET102 vector was present in both the insoluble and soluble fractions of most expression cultures. Protein in the latter fraction however, is noticeably lower, illustrated by fainter bands at the parallel 37 kDa weight marker.

Inversely, all samples that utilised the pET100 expression vector displayed an absence of bands of the appropriate weight in the lanes containing soluble fractions. Thus, the recombinant protein in these samples was primarily insoluble.

As the pET100 vector failed to express soluble protein, Sections 4.2.3 – 4.2.6 will analyse the effects of the remaining environmental factors exclusively on samples that utilise the pET102 expression vector, unless otherwise stated. Expression of soluble protein in these samples was affirmed by bands located at the 37 kDa weight marker outlined in the relevant SDS-PAGE figures, indicating the presence of the protein of interest.

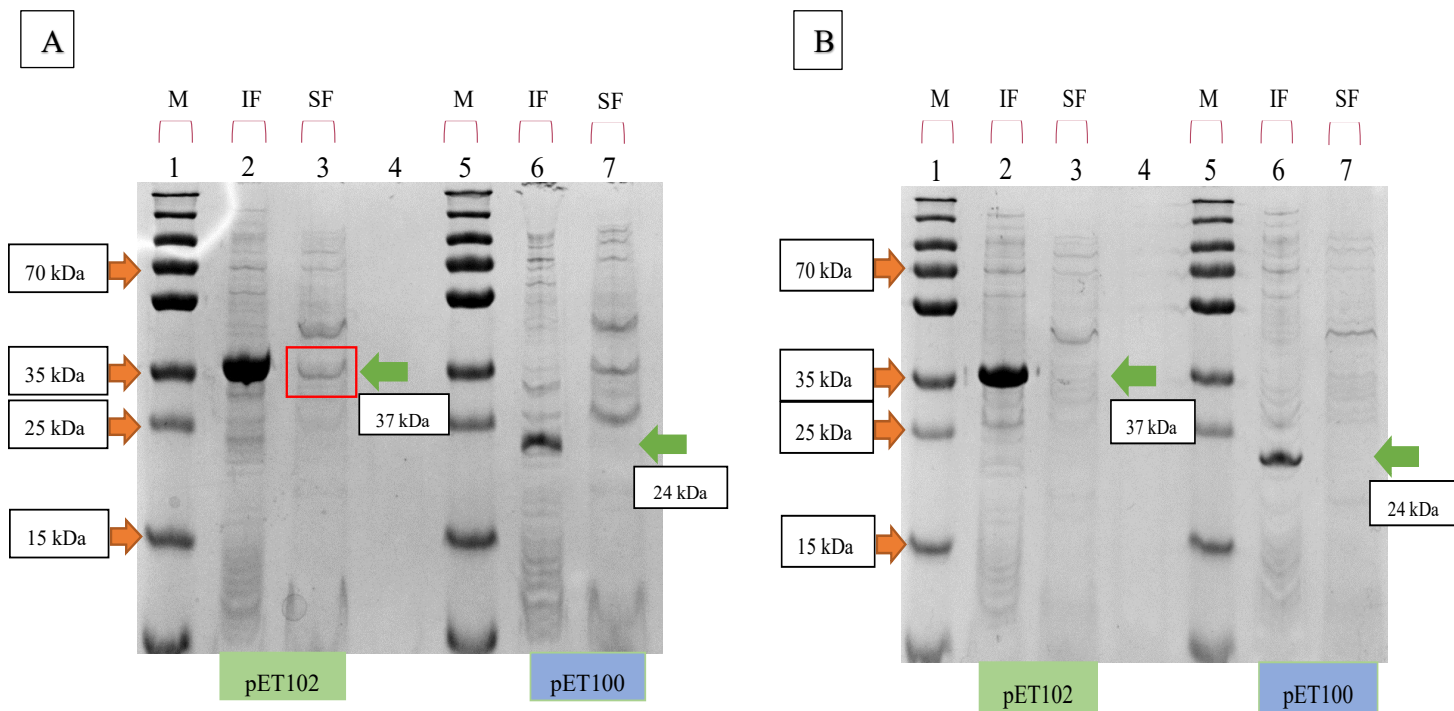


Figure 4.2 SDS-PAGE analysis of protein expressed from the pET100 and pET102 vectors.

Expression was carried out with (A) LB or (B) LB Sorbitol with an IPTG concentration of 0.2mM and an overnight induction at 16°C.

Lane M: Protein molecular weight standards. Insoluble fractions (IF) were loaded into lanes 2 and 6 of both gels , while the soluble fractions (SF) are in lanes 3 and 7. The protein of interest is visible, with a MW of 37kDa when expressed by the pET102 vector and a MW of 24kDa expressed from the pET100 vector.

4.2.3 Identifying the optimum length of expression time

A time course analysis was performed by removing 500µl aliquots of cell culture from IPTG induced samples after 1, 2, 3 and 4 hours of incubation, with a final aliquot removed after a period of overnight incubation.

In samples inclusive of both pET expression vectors, a steady increase in protein expression was observed in line with an increase in incubation time. This can be interpreted by the increased intensity and thickness of corresponding bands in Figures 4.2A and 4.2B. Samples that were allowed to incubate overnight, demonstrated an increase in soluble product (Figures 4.3B, 4.3D, 4.3F) except for overnight samples incubated at 16°C (Figure 4.1A).

Within the 1 - 4 hour period post induction, the length of incubation did not affect the amount of soluble protein product expressed from the pET102 vector (Figures 4.2A & 4.2B). As such, further analysis of these samples will now be focused solely on those that were incubated for three hours, to allow for a more efficient interpretation of results.

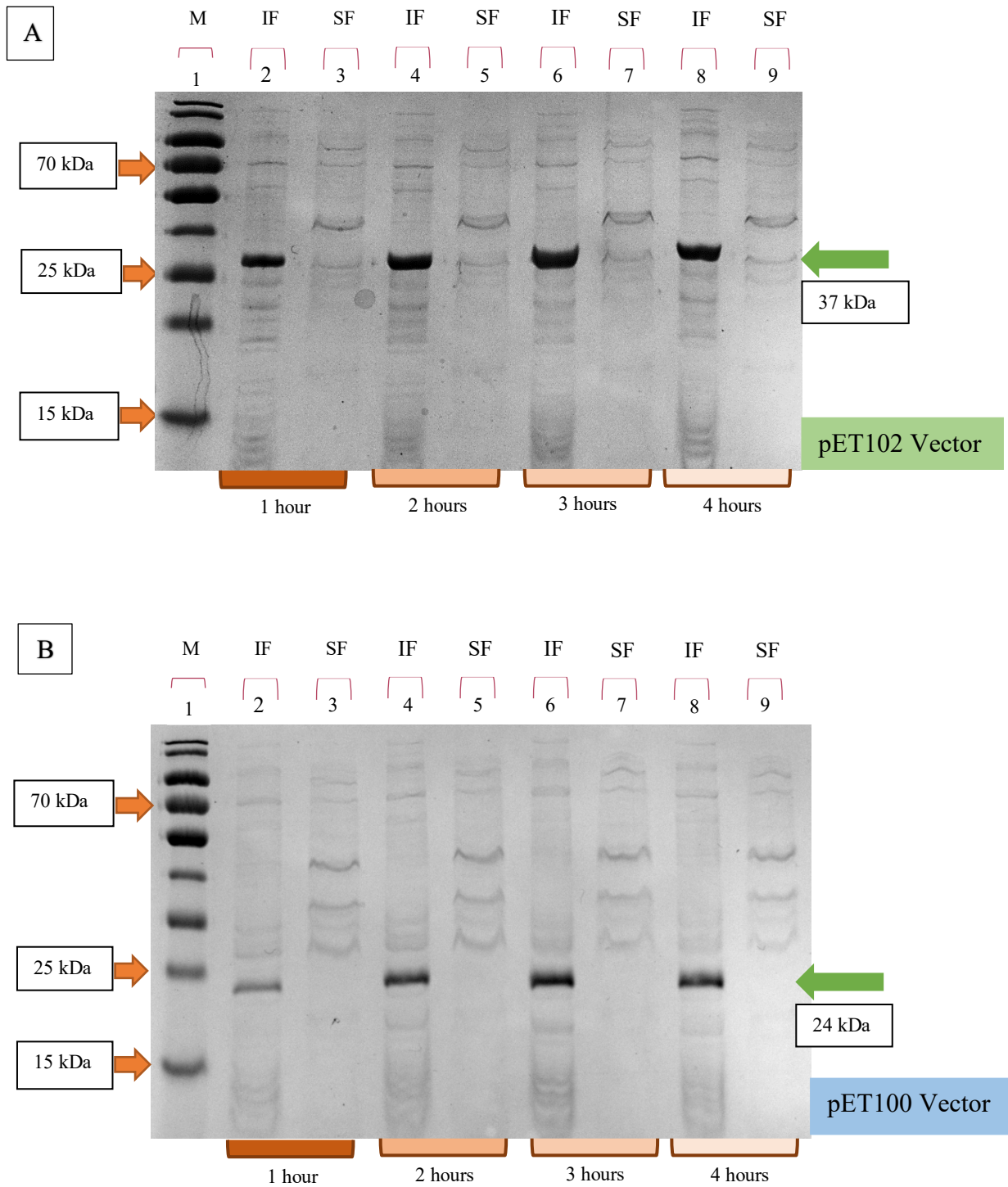
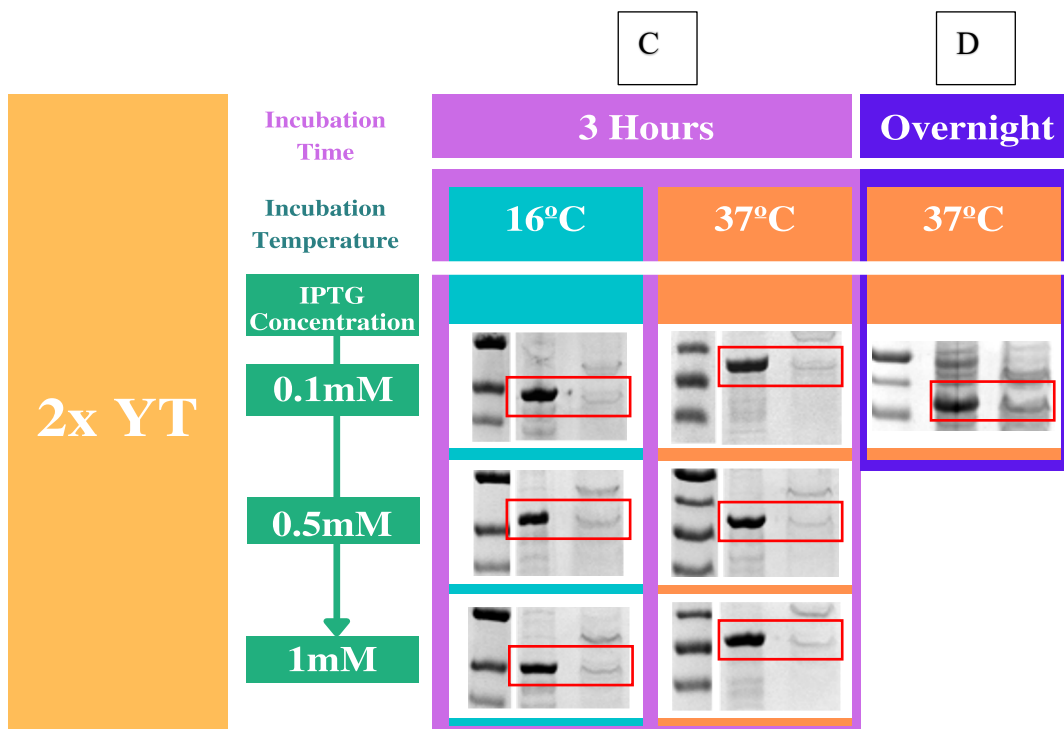
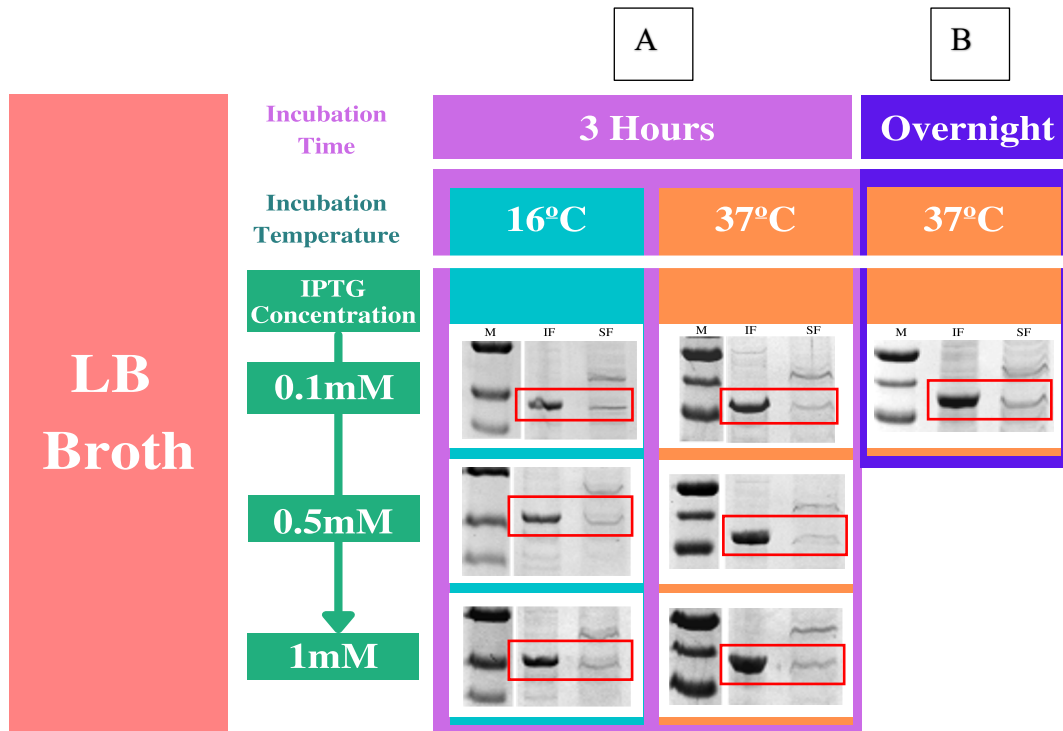


Figure 4.7 SDS-PAGE analysis of protein expressed in TB media from the pET100 and pET102 vectors. Expression was induced from either the (A) pET102 vector or (B) pET100 vector using an IPTG concentration of 0.1mM and an incubation temperature of 37°C. 500µl aliquots were removed after a 1 hour (lanes 2, 3), 2 hour (lanes 4, 5), 3 hour (lanes 6, 7) and 4 hour (lanes 8, 9) incubation periods. Lane M: Protein molecular weight standards. Insoluble fractions (IF) were loaded into lanes 2 and 6 of both gels, while the soluble fractions (SF) were in lanes 3 and 7. Soluble protein is visible only in figure A, with a MW of 37kDa as indicated by the green arrow.



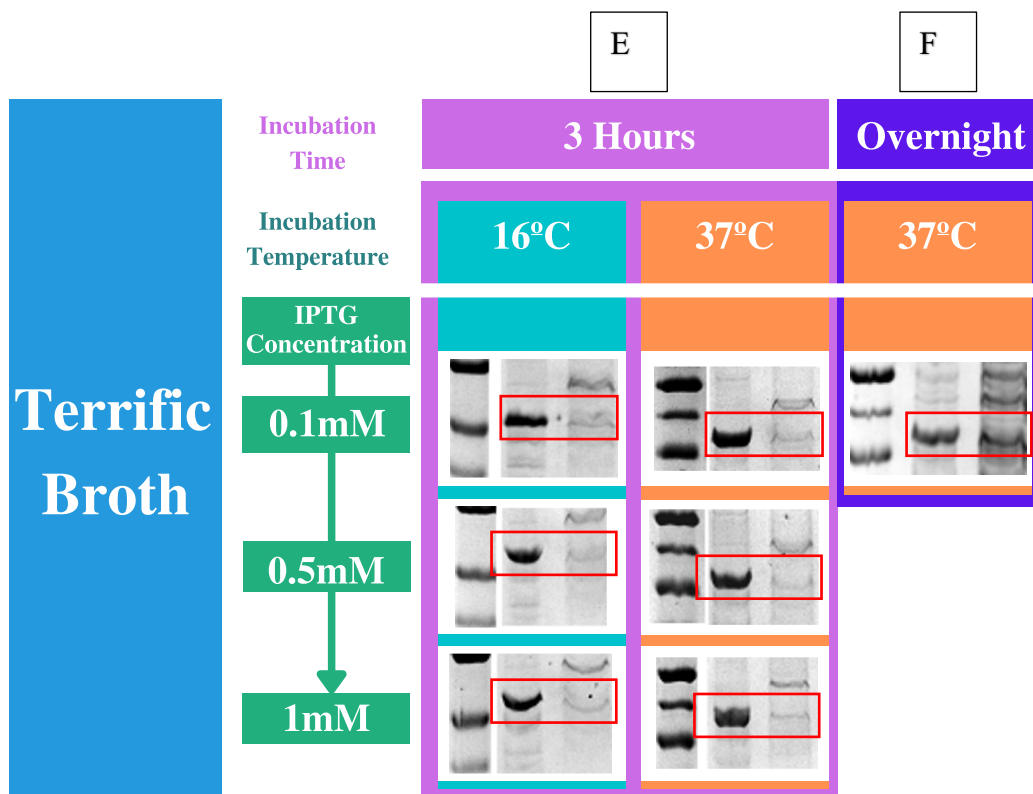


Figure 4.11 SDS-PAGE analysis of the effect of varying growth medias, IPTG concentrations and induction temperatures on the expression of soluble protein from the pET102 expression vector.

Concentrations of 0.2mM, 1mM and 2mM IPTG, combined with an incubation temperature of 16°C and 37°C were applied to samples in A, C and E, while B, D and F depicts samples induced with 0.2mM IPTG and an overnight incubation temperature of 37°C only.

To culture cells, the following growth medias were selected; LB broth (A, B), 2xYT (C, D) and TB (E, F).

4.2.4 Identifying the optimal growth medium type

E. coli BL21 cells were cultivated in one of the following four growth media; LB medium, Terrific broth, 2x YT medium or LB supplemented with 1M sorbitol 2.5 mM betaine medium (Oganesyan, Ankoudinova, Kim and Kim, 2007). Cultured cells were then subjected to lysis, followed by a period of centrifugation as outlined in Section 2.9.1 before analysing the insoluble and soluble fractions via SDS-PAGE. Soluble protein expression was observed in three out of the four media under a variety of incubation times, temperatures and IPTG concentrations as seen in Figure 4.3. However, similar to samples obtained using the pET100 vector, samples cultured in LB sorbitol yielded entirely insoluble recombinant protein, which is highlighted in Figure 4.1B.

Protein migrating to 37 kDa was observed in the soluble fractions obtained from the cultures of the three remaining growth media in Figure 4.3. Although the yield of protein could not be determined solely by the type of media used, samples obtained from three hour LB cultures (Figure 4.3A) have yielded almost identical quantities of soluble protein, regardless of other external factors. In contrast, there was a very slight reduction in band intensity amongst most 2xYT growth media samples incubated for three hours (Figure 4.3C). Overall however, a notable increase was observed in samples from all three, once an overnight period of incubation was applied (Figures 4.3B, 4.3D, & 4.3F).

4.2.5 Identifying the optimal IPTG concentration

IPTG concentrations of 0.2mM, 1mM or 2mM were used to induce cell cultures once an OD_{600} of between 0.6 – 0.8 had been reached. Again, analysing the SDS-PAGE results in Figure 4.3, recombinant protein of varying quantities, could be detected in the soluble fraction of samples induced with each concentration. The volume of IPTG did not appear to influence the amount of soluble protein produced, with the exception of samples that had been induced with 0.2mM and once again incubated overnight, as previously highlighted in Figure 4.3.

4.2.6 Identifying the optimal expression temperature

Incubation temperatures of both 37°C and 16°C were chosen to incubate samples after IPTG induction. Recombinant protein can be observed in the soluble fractions of most samples that have been incubated at either temperature (Figure 4.3).

Similar to IPTG concentration analysis, the level of soluble protein expression does not appear to be influenced? dictated by either temperature. The exceptions to this are samples that underwent an overnight incubation period. Those that did so in a 16°C environment (Figure 4.3A) produced marginally less soluble protein than those that were incubated at 37°C.

4.2.7 Determining the overall environment best suited for the production of soluble recombinant protein

Analysing the overall results achieved upon the application of expression parameters described above, it can be concluded that further trials are needed in order to accumulate enough soluble protein that would allow for its subsequent purification and characterisation.

Although the total soluble protein expressed in this study was insufficient, there were some expression parameters that appeared more effective in comparison to others. The application of an overnight incubation appeared to be the most effective length of time for the incubation of cell culture, when combined with each variety of growth media. Unfortunately, due to time constraints, this overnight incubation was only carried out at a temperature of 37 °C, excluding the application of a 16°C incubation temperature. The same can be said when analysing the most optimal IPTG concentration, where inducing cell cultures with a much lower concentration of 0.2mM appeared to yield the most soluble protein, especially when combined with an overnight incubation period. Once again however, this particular trial was not repeated with the implementation of additional IPTG concentrations. Overall, the combination of incubating cell culture induced with 0.2mM IPTG at 37 °C appeared to be the most favourable combination of expression parameters. The use of a richer growth media such as TB or 2x YT appeared marginally more effective compared LB broth and so a scale up of protein expression incorporating these factors was performed and the results detailed in the next section.

4.2.8 Scale up of protein expression.

Using the environmental parameters determined to favour the expression of soluble protein that are outlined above, a 100ml scale up was performed. At the time, samples from TB culture had not yet been analysed and owing to limited lab access time, a scale-up was instead performed with 2xYT media which had previously delivered similar results.

Samples were induced with 0.2mM of IPTG and incubated at 37°C overnight. Once cultures had been lysed and centrifuged, a two-fold dilution was performed to address the predicted increase in protein concentration that would affect the reliable interpretation of results. As expected, excess protein can clearly be observed in lanes 2-5 of Figure 4.4. Dilutions of 1 in 4 and 1 in 8, (lanes 6-9) successfully address this issue, however bands identified at the 37 kDa weight marker are noticeably fainter compared to those using the small-scale expression. Thus, the combination of environmental parameters hypothesised to aid in expression of soluble protein proved unsuccessful, for reasons that will be discussed in Chapter 5.

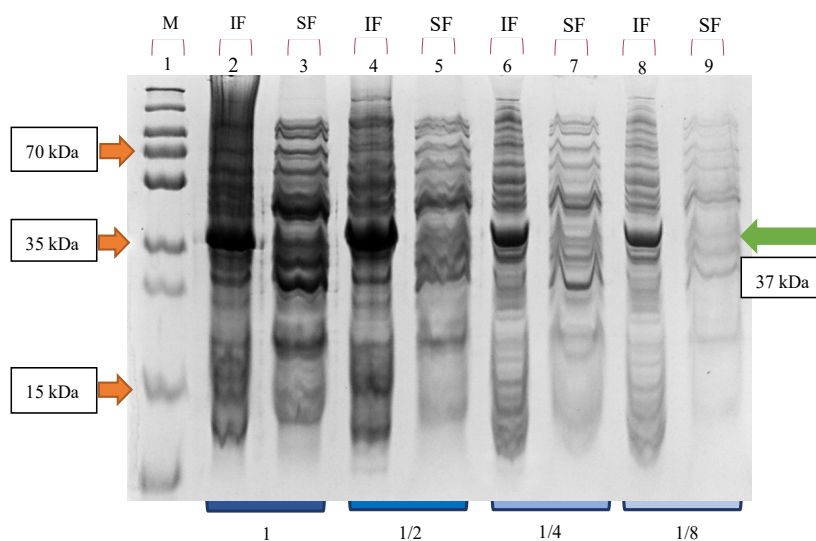


Figure 4.16 SDS-PAGE analysis of protein expression from the pET102 vector, scaled up to 100ml with the application of specific expression parameters.

Conditions determined to favour the greatest yield in soluble protein comprise of *E. coli* BL21 (DE3) cells cultured in 2xYT media and induced with 0.2mM of IPTG followed by a 37°C incubation overnight.

Lane M depicts the molecular weights of protein markers. Insoluble fractions (IF) and soluble fractions (SF) are indicated above the image. A two-fold dilution was performed on samples in lane 2 and 3 which resulted in samples diluted 1 in 2 (lanes 4 and 5), 1 in 4 (lanes 6 and 7) and 1 in 8 (lanes 8 and 9).

4.3 Discussion

Following the successful cloning of putative azoreductase genes from the genome sequence of the *O. splanchnicus* 22-5 strain into a BL21 (DE3) host cell line, a variety of molecular and environmental conditions were applied during overexpression of recombinant protein. This would allow for the identification and optimisation of the expression parameters necessary for procuring a satisfactory volume of soluble protein, allowing for its subsequent purification and characterisation. Solubility is usually the most common encountered obstacle when optimising heterologous expression systems for the production of recombinant protein. Obtaining soluble proteins is desirable as they are typically correctly folded, functional and much easier to purify than their insoluble counterparts (Costa, Almeida, Castro and Domingues, 2014)

In the past, recombinant expression of novel azoreductases has proven to be challenging, evident by the variation of environmental conditions reported to best suit accumulation of soluble protein. A standard expression protocol had originally been established for the successful expression and characterisation of azoreductases paAzoR1 (Wang *et al.*, 2007) and paAzoR2 and paAzoR3 (Ryan, 2016) from the bacterial genome of *Pseudomonas aeruginosa* PAO1 strain. An additional four azoreductases; PA2280, PA2580, PA1204 and PA0949, also from the same strain were later characterised by Crescente and colleagues (2016) who had utilized the same expression method, and had similarly adopted the pET expression vector and BL21 host cell line. This protocol detailed the use of LB supplemented with 1M sorbitol 2.5 mM betaine, an IPTG concentration of 0.5mM and an incubation time of 16 hours at 16°C. However Holland (2017), utilised the same expression techniques to characterise a further three azoreductases from this strain; PA1224, PA1225 and PA4945. Interestingly this protocol only proved successful in the overexpression of PA1224. Subsequent variations in culture conditions including the elimination of sorbitol and a reduced IPTG concentration and temperature were required to successfully express the latter enzymes and procure an acceptable volume of protein in the soluble fraction. As demonstrated here by Holland and Crescente, the identification of optimum expression parameters are crucial to the success in expressing proteins of interest.

Predictably, in this study, the volume of recombinant protein expressed and its degree of solubility was indeed dependent on molecular factors and culture conditions. In order to observe the influence of these elements, SDS-PAGE was performed on samples subjected to

varying expression parameters and ImageJ software was employed to observe protein levels in both soluble and insoluble fractions of expression cultures.

High level expression can be observed amongst all samples, however most of the protein was contained within the insoluble fraction.

A high level of expression in *E. coli* is a common cause of aggregating expressed protein into inclusion bodies with high incubation temperature, IPTG concentrations etc. reported to be the most common causes (Baneyx and Mujacic, 2004). A similar observation can be made in Figures 4.1 – 4.4, however the measure of soluble protein proportional to the total volume of protein expressed does not appear to be greatly affected when incubation time or temperature are reduced. Other factors influencing the formation of insoluble aggregates include poor post-translational mechanisms and molecular chaperones, and the highly reducing environment of the bacterial cytosol (Carrió, Cubarsi and Villaverde, 2000). SDS-PAGE gel image analysis demonstrates that the TRX tag of the pET102 vector rescues a small amount of soluble protein. Conversely, a complete absence of soluble protein is observed when expression is performed with the pET100 vector, whose structure lacks the TRX tag as evident in SDS-PAGE analysis, where it is identifiable as a smaller 24 kDa band in comparison to the 37 kDa band of the pET102 vector.

As discussed earlier in this chapter, the TRX tag is thought to aid protein solubility by maintaining a more stable cytoplasm environment, allowing proteins to fold correctly.

Hence, the use of the TRX tag in this study may indicate that the cytoplasm of BL21 DE3 cells is unsuitable for the formation of this target protein. Indeed, Derman *et al.*, (1993) details similar difficulties where specific proteins were unable to properly fold due to the naturally reducing environment of the cytoplasm in BL21 DE3 cells. Additionally, these studies observed an increase in soluble protein yield when using an altered form of this *E. coli* strain, where modifications such as mutations or inactivation of cytoplasmic reductases contribute to a more oxidative environment (Lamark *et al.*, 2001; Xiong, 2005).

Out of all the samples that utilised the pET102 vector, two in particular displayed a notable increase in soluble recombinant protein evident by bands at the 37 kDa marker. Interestingly, both samples had been induced with 0.2mM of IPTG and had undergone an overnight incubation of 37°C. The idea that these specific parameters could contribute to an

ideal/optimal expression environment is strengthened by the observed reoccurrence of results in two different growth media; 2xYT and TB, however the latter yielded slightly more protein.

Both media are rich in nutrients, but high concentrations of tryptone, yeast and potassium phosphate in TB help maintain a higher cell density for an increased period of time in comparison to LB (Tartoff and Hobbs, 1987) This would explain why the above parameters were not as successful when repeated with LB growth medium.

Where the supplementation of LB with sorbitol 1M, a proposed solubility enhancer (Blackwell and Horgan, 1991) had achieved success for Ryan *et al.*, (2010), Crescente *et al.*, (2016) and Holland (2017), it had, by contrast resulted in the production of entirely insoluble protein in this study.

Evaluating all other gradients of IPTG concentration, incubation temperature and time, an increase in each appeared to have a direct correlation with the volume of total protein expressed and all had procured some level of this protein in the soluble fraction, but notably less than the two samples discussed above. Thus, it would be reasonable to assume that the environment best suited for successful protein solubility mainly requires cultivation in a rich media, preferably TB or 2xYT, combined with an IPTG concentration of 0.2mM and an overnight incubation period at 37°C.

Due to time constraints, performing a scale up with TB was not possible, and was instead performed using 2xYT medium. Results were not as successful, an occurrence similarly experienced by Holland and by Peng and colleagues (2004) who notes that this most likely a result of insufficient oxygen levels influenced by an increase in cell density.

Similar to Holland's solution, the most appropriate follow up strategy would have warranted the repeat of a small scale protein expression trial using the same cultivation conditions outlined in Figure 4.4, only this time, a single 50ml 2x YT culture would be exchanged for duplicates of samples containing TB growth media. Unfortunately, time restraints prevented the development of any further laboratory work. Instead, using the results accomplished in Chapters 3 and 4, a hypothesis regarding the identity and function of protein of interest and of the bacterium *O. splanchnicus*, will be discussed in the following chapter.

Chapter 5 Final discussion and conclusion

The first aim of this project was to screen the genome of a representative gut bacterial strain, more specifically, that of *O. splanchnicus* in order to identify one or more putative azoreductase genes. To do this, the amino acid sequences of all characterised and accessible azoreductases were compiled and blasted against the genomic sequence of *O. splanchnicus*. (Table. 5.1). Considering the environmental preferences of *O. splanchnicus*, it made sense when AzoC, a well characterised azoreductase in the bacteria *C. perfringens* was the only protein to successfully align. Both *O. splanchnicus* and *C. perfringens* are anaerobic and are recognised as normal components of the human intestinal tract. Additionally, many reports detail the azoreductase ability of AzoC in reducing azo compounds including amaranth (Mcbain and Macfarlane, 1998), methyl red and most importantly tartrazine (Morrison, Wright and John, 2012; Morrison and John, 2016). Thus, the alignment of AzoC, yielded a potential candidate in the form of a 202 amino acid protein with 51% sequence identity (Table. 5.2). Generally, a cut-off above 30% sequence identity is recommended when identifying homologous proteins (Rost, 1999). However, although the sequence identity is lower than desired it should be noted that many characterised azoreductases share a low sequence homology (Misal and Gawai, 2018). In response to this, azoreductases have been subsequently categorised based on their oxygen tolerance, their expression (intra/extracellular) and their co-factor preference (Chen, Hopper and Cerniglia, 2005; Nakanishi, Yatome, Ishida and Kitade, 2001; (Bafana and Chakrabarti, 2008).

Recently, Zou and colleagues (2020) performed an identical blast search, aligning AzoC against the same strain of *O. splanchnicus*. Likewise, upon generating a protein with 51% sequence identity, they also hypothesised that the protein in question would possess azoreductase properties/qualities. As mentioned in Chapter 1, Zou and her colleagues further elaborated their investigation by demonstrating the enzymatic abilities of *O. splanchnicus*, in particular its ability to reduce a variety of azo dyes (Table 1.2 and 1.3). Similarly, these findings help reinforce the hypothesised azoreductase properties of the protein identified in this project.

Table 5.1 The amino acid sequences of the azoreductases that had been characterised to date, alongside their amino acid length and ID and accession numbers.

These were used in a Blastp search against the full genome sequence of *O. splanchnicus* in order to identify a hypothetical azoreductase protein with a similar sequence identity.

<i>Scientific Name</i>	<i>Protein description</i>	<i>Protein ID</i>	<i>Acc. Len</i>	<i>Accession</i>
<i>Enterococcus faecalis</i>	<i>Azoreductase</i>	<i>AzoA</i>	208 aa	AAR38851.1
<i>Shewanella oneidensis</i>	<i>FMN-dependent NADH-azoreductase</i>	<i>soAzoR</i>	198 aa	WP_011074045.1
<i>Enterococcus faecium</i>	<i>FMN-dependent NADH-azoreductase</i>	<i>AzoEfl</i>	206 aa	WP_125190802.1
<i>Clostridium perfringens B str. ATCC 3626</i>	<i>Azoreductase</i>	<i>AzoC</i>	206 aa	AGH15624.1
<i>Pseudomonas aeruginosa PAO1</i>	<i>FMN-dependent NADH-azoreductase 1</i>	<i>PaAzoR1</i>	212 aa	AAG04174.1
<i>Bacillus sp. B29</i>	<i>Azoreductase</i>	<i>AzrA</i>	208 aa	BAF02597.1
<i>Escherichia coli str. K-12 substr. MG1655</i>	<i>FMN dependent NADH:quinone oxidoreductase</i>	<i>AzoR</i>	201 aa	NP_415930.1

Table 5.2 A hypothetical protein identified from the genome of *O. splanchnicus* using protein sequences from characterised azoreductases in a blastp search.

The name, description and sequence similarities of the protein identified from *O. splanchnicus* using a blastp search with the azoreductase amino acid sequences listed in Table 5.1.

<i>Scientific Name</i>	<i>Protein description</i>	<i>Query Cover</i>	<i>E value</i>	<i>Per. ident</i>	<i>Acc. Len</i>	<i>Accession</i>
<i>Odoribacter splanchnicus</i>	<i>flavodoxin family protein [Odoribacter splanchnicus]</i>	100%	3.00E-73	50.72	202	<u>WP_013611494.1</u>

The second aim of this project involved selecting a cloning kit that would effectively clone the gene of interest and transform it into an appropriate/suitable host cell line. While the pET TOPO[®] vector successfully enabled cloning of this gene, as confirmed by sequencing results provided by Genewiz, it appears that the transformation process could be improved.

In this study, a number of molecular and culture conditions were modified to determine conditions optimal in generating recombinant protein in the soluble fraction.

Applying varying combinations of environmental factors during cell culture incubation, and protein expression, with a selection of expression vectors were effective in generating recombinant protein product. In particular, the combination of using a nutrient-rich media such as TB or 2xYT, inducing with 0.2mM IPTG, followed by an overnight incubation at 37°C appeared to be the most effective in yielding recombinant protein. Attempts to procure protein in the soluble fraction however, were not as successful.

The conditions commonly used to induce a high level of protein expression can often cause the expressed protein to become aggregated and subsequently insoluble (Chrnyk *et al.*, 1993). An increase in post expression temperature and inducer concentration are conditions often relied on to accelerate the translation and expression of a desired protein. This steep rise in expression however, predictably increases cell density while at the same time exhausting cell resources resulting in limited oxygen, a decrease in pH and ultimately generating damaged and misfolded proteins (Jia and Jeon., 2016; Carrió *et al.*, 2000). These issues can indeed be observed in Figure. 4.3, where increased inducer concentration and temperature yielded a high level of recombinant protein yet the majority was insoluble. Mild improvements can be seen when using rich media types as their high nutrient and phosphate levels help maintain pH levels and prolong cell growth (Tartoff and Hobbs, 1987). Likewise, a lower IPTG concentration also appears to slightly improve protein solubility as suggested by Baneyx and Mujacic, (2004).

It would have been interesting to observe the effects of combining an overnight incubation with a reduced temperature such as recommended by Mogk, Mayer and Deuerling, 2002. If a reduction in temperature failed to improve protein solubility, similar to what was observed in Figure 4.3 when it was combined with a 3 hour incubation period, then a number of additional modifications could be considered in future experiments. To combat the depletion of chaperones whose expression and activity is often reduced in colder temperatures, the use of the ArticExpress[™] expression host cell line could prove particularly useful. Ferrer, (2003)

describes the high folding abilities of the chaperonins it possesses, which are derived from the psychrophilic bacterium *Oleispira antarctica*.

Other host cell lines that could prove advantageous include *E. coli* BL21 *Rosetta-gami2(DE3)* cells which have been engineered to counter potential codon bias by introducing rare codons (Tegel, Tourle, Ottosson and Persson, 2010). Origami or SHuffle cells can also combat the issue of *E.coli*'s naturally oxidising cytoplasm that can otherwise prove unfavourable when expressing some heterologous proteins (Xiong, 2005; (Lobstein *et al.*, 2012). Another modification for future experiments would be the consideration of an alternate fusion tag such as the maltose-binding protein (MBP) or SUMO tags. Like the TRX tag used in this study, both tags are highly regarded for their effective solubilising capabilities. These tags all employ different mechanisms to improve protein solubility, for example the SUMO tag creates a detergent like effect on insoluble proteins (Butt, Edavettal, Hall and Mattern, 2005) while the MBP protein binds to and inactivates intramolecular proteins that are susceptible to aggregation (Raran-Kurussi, Keefe and Waugh, 2015).

Finally, a newly introduced method of protein expression known as microexpression, offers several attractive benefits. It allows many more combinations of expression parameters to be carried out in 500 μ L culture samples, without compromising time constraints or budgeted costs (Page *et al.*, 2004). Hence, microexpression would allow for rapid screening and the development of a more effective expression protocol.

It is clear that there is no universal protocol for expressing novel proteins. Instead, the most common method undertaken is the 'trial and error' approach, as described above. These methods can prove to be tedious, requiring a great degree of improvisation, patience, and often resulting in considerable delays and increased strain on costs and resources.

In conclusion, the results here demonstrate the difficulties encountered in increasing solubility of the protein of interest. Subsequent purification and characterisation could not be carried out, preventing the identification of this protein and its function, and as a result the second aim of this project could not be accomplished.

Irrespective of these uncompleted aims, it is still possible to further hypothesise on the classification and function of the protein that was successfully expressed in this study. An unpublished report by Gao and her colleagues describe identifying AzoC homologs in many species of gut bacteria belonging to/from all major phyla (Guo, Lee and Jeong, 2020). Likewise, from the protein sequence of AzoC, Zou *et al.*, (2020) identified a protein of

identical sequence similarity to the protein expressed in this project. More significantly, they were able to first demonstrate the azo reducing abilities of a strain of *O. splanchnicus* isolated from the human gut. While these results strengthen the hypothesised function of the protein in question, they also necessitate the need for further understanding the primary role of these gut microbial azoreductases.

It could simply be that these gut microbial strains have developed defensive enzymes/metabolic pathways upon exposure to environmental stressors such as ingested xenobiotics or to the toxins or antibiotics produced from neighbouring gut pathogens.

Furthermore, it is also possible that the potentially harmful by-products of azoreduction may be degraded even further before they reach the liver or kidney. Zhang and colleagues (2019) demonstrated an interesting example of this. They highlighted the ability of certain gut bacterial strains in metabolising carcinogenic heterocyclic aromatic amines that are normally obtained through diet. The resulting metabolites displayed significantly reduced cytotoxicity towards human intestinal cells.

To date, many studies concerning tartrazine consumption appear to omit important factors that would have otherwise created a more accurate representation of human conditions, allowing for better reproducibility of results. Reports of its genotoxic and cytotoxic effects were mostly observed when tartrazine had been administered *systemically*, a process that avoids exposure to the intestinal microflora (Zhang *et al.*, 2013). As tartrazine is widely consumed via food products, its metabolism is mainly concentrated in the digestive tract, and so the results of these reports are not an accurate portrayal of the metabolic route of tartrazine in humans. Studies where the dye had been orally administered suggest a potential to induce inflammation and DNA damage in hepatic and colonic cells. However in these cases, the dose of tartrazine used was well over its outlined ADI (Sasaki *et al.*, 2002; (Meyer *et al.*, 2017).

Another important consideration is that many of these studies were performed using murine models. This is significant because the gut microbiome of a mouse does not accurately represent that of a human, and so the use of murine models would likely impact the reproducibility of results. Hugenholtz and de Vos (2017) discuss this concern in great depth, broadly illustrating the differences in the abundance of gut bacterial genera in murine vs human (Figure 5.1). They also note that the murine gut microbiome can be heavily influenced

by many factors that differ greatly from our own such as environmental, diet and genetic factors (many mice strains are inbred and raised on a restricted diet in captivity).

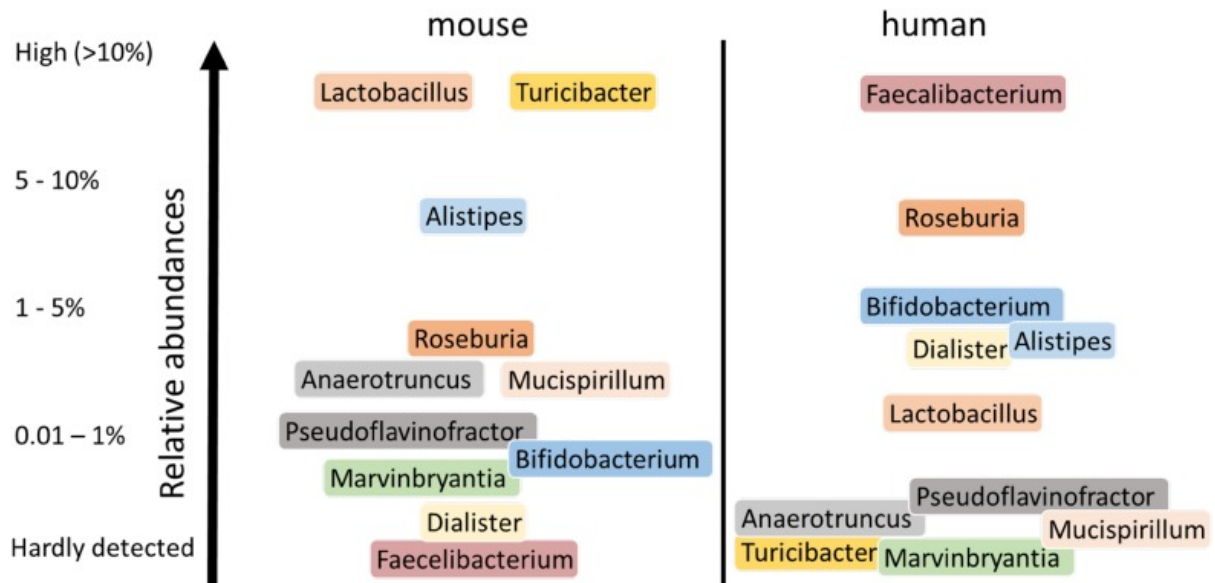


Figure 5.1 Major quantitative differences between human and murine intestinal genera.

A comparison between the metagenomic datasets of the human and murine intestinal microbiota. This figure illustrates strong similarities at the genus level but large deviations on a quantitative level. Figure is taken from (Hugenholtz and de Vos, 2017)

Even if these studies were repeated using a human microbiome model, it would still be difficult to accurately predict the metabolic pathway of tartrazine. The human digestive tract is extraordinarily complex, with its metabolic capability being heavily dictated by a) the intricate relationship shared with the unique composition of resident microflora it harbours and b) by individual human factors of the host such as diet, age and genetics.

Zahran, Ali-Tammam, Ali and Aziz, (2021) provide a good illustration of this in their very recently published pilot study. With the faecal samples of 16 healthy adults, they isolated bacterial species that demonstrated an ability to reduce the azo dye Brilliant Black. Then, for each sample the relative abundance of each species was calculated and its efficiency in reducing the dye. By doing this, they were able to show that the rate at which an azo compound is reduced can be dictated by the abundance of azo reducing species in the human gut. This in turn is inversely proportionate to the abundance of other species present. Species in the phylum Firmicutes demonstrated a higher azoreductase ability as opposed to the phylum Bacteroides, to which *O. splanchnicus* belongs, which reduced the azo dye at a much lower rate. As displayed in Figure 5.2, the bacterial population of each stool sample varied. Typically, samples with a low population of Firmicutes exhibited lower azo-reducing activity. These findings are extremely important as they clearly demonstrate that the composition of one's gut microbiome can influence the fate of xenobiotics, in this case those containing an azo compound such as pro-drugs like prontosil or food colourings like tartrazine.

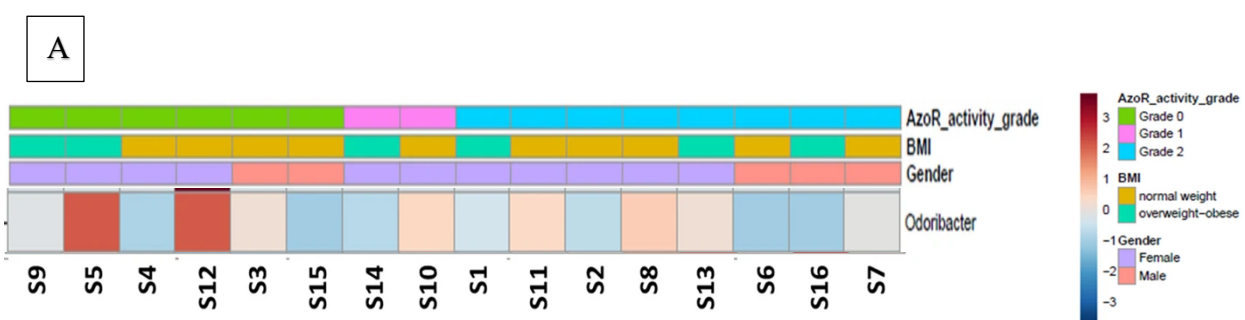
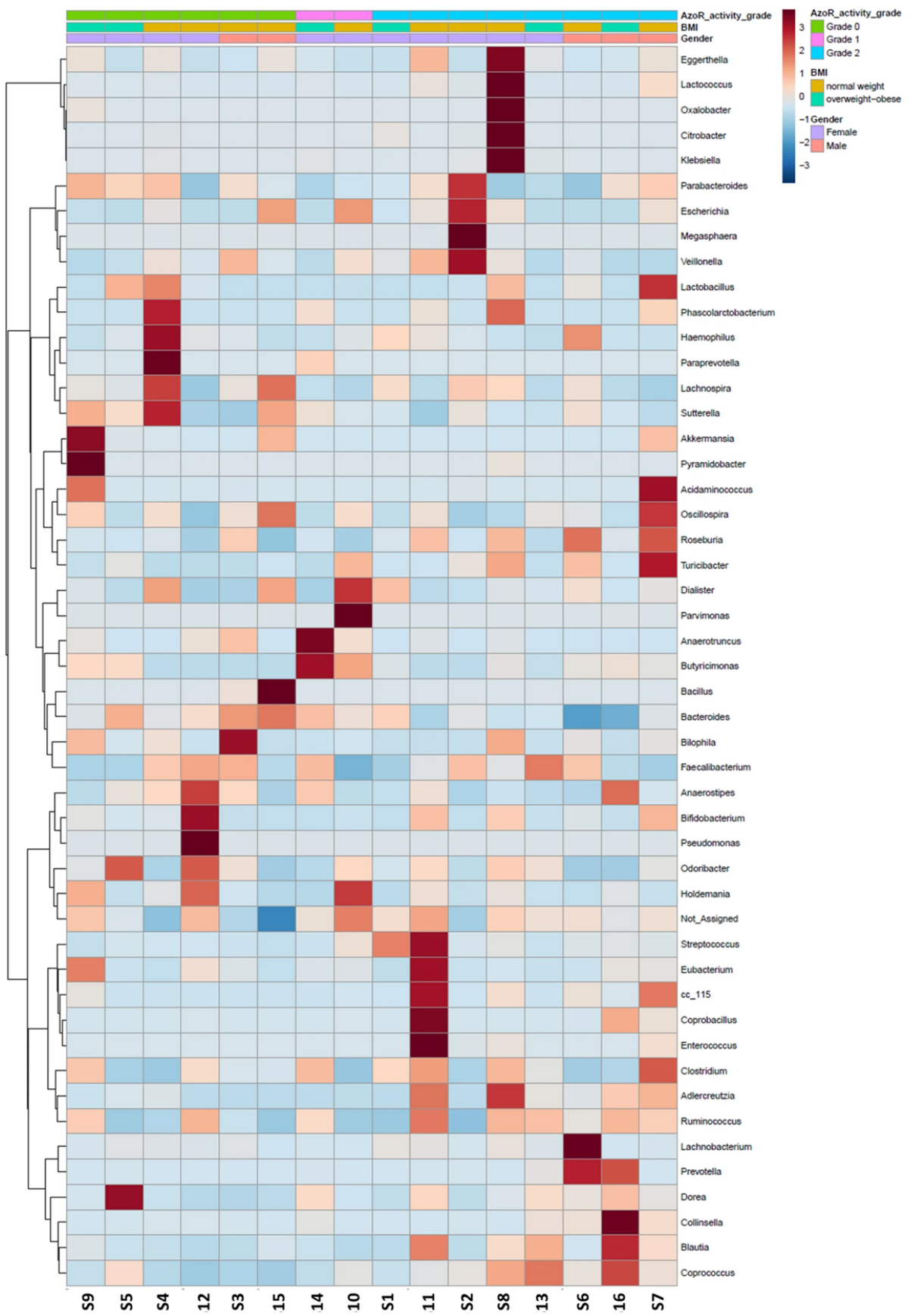


Figure 5.12 A heatmap demonstrating the abundance of (A) *O. splanchnicus* and (B) other identified gut bacterial strains, in faecal samples that were collected from 16 healthy adults. Each strain is graded by its ability in reducing the azo dye Brilliant Black.

The heatmap colour denotes the relative abundance of each strain across 16 collected faecal samples. The colours above the heatmap represent the total azo-reducing capacity of each species where Grade 0 indicates partial decolourisation and Grade 2 indicates near to or complete decolourisation. The sex and BMI classification of each subject is also represented in colour above. Adapted from Zahran *et al.*, 2021.

B



This study also allows us to further speculate on the azoreductase properties specific to *O. splanchnicus*. The first observation is that despite the localized diversity and small scale of the sample population (Egyptian participants that likely share a similar diet and environment), the abundance of *O. splanchnicus* in each individual stool sample is highly variable. This suggests that the composition of one's intestinal microflora could be highly sensitive and easily influenced by other factors such as antibiotic treatment and regular exposure to xenobiotics including azo compounds.

The second observation is that *O. splanchnicus* may possess better azo-reducing abilities in some individuals compared to others. Many factors may be responsible for influencing this. If the azoreductases in question functioned primarily as a defensive mechanism, then its expression and activity could be dependent on exposure to dangers such as oxidative stress or the toxins and antibiotics produced by neighbouring species.

To conclude, the metabolic capability of the gut microbiota has become an important issue of concern in recent years. These findings highlight the potential sensitivity of an individual's gut microbiome and suggest that exposure to environmental, xenobiotic or otherwise toxic conditions may disrupt the balance of the gut bacterial community which could inherently influence the fate of these substances. While it is imperative that a more comprehensive understanding into the functioning of the gut microbiome is needed, it is equally important to appreciate the individuality of each microbiome where variations in the abundance of species such as *O. splanchnicus* can dictate the metabolic pathway of everyday xenobiotic compounds such as tartrazine.

References

2010. *pBAD/Thio His TOPO manual*. [online] Available at: <https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FSLSG%2Fmanuals%2Fpettopo_man.pdf> [Accessed 1 December 2021].
- Amin, K., Abdel Hameid, H. and Abd Elsttar, A., 2010. Effect of food azo dyes tartrazine and carmoisine on biochemical parameters related to renal, hepatic function and oxidative stress biomarkers in young male rats. *Food and Chemical Toxicology*, 48(10), pp.2994-2999.
- Baba, R., Kimura, M., Asakawa, S. and Watanabe, T., 2013. Analysis of [FeFe]-hydrogenase genes for the elucidation of a hydrogen-producing bacterial community in paddy field soil. *FEMS Microbiology Letters*, 350(2), pp.249-256.
- Bafana, A. and Chakrabarti, T., 2008. Lateral gene transfer in phylogeny of azoreductase enzyme. *Computational Biology and Chemistry*, 32(3), pp.191-197.
- Bakthavachalu, P., Kannan, S. and Qoronfleh, M., 2020. Food Color and Autism: A Meta-Analysis. *Advances in Neurobiology*, pp.481-504.
- Baneyx, F. and Mujacic, M., 2004. Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology*, 22(11), pp.1399-1408.
- Bergman, E., 1990. Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. *Physiological Reviews*, 70(2), pp.567-590.
- Bhatt, A., Pellock, S., Biernat, K., Walton, W., Wallace, B., Creekmore, B., Letertre, M., Swann, J., Wilson, I., Roques, J., Darr, D., Bailey, S., Montgomery, S., Roach, J., Azcarate-Peril, M., Sartor, R., Gharaibeh, R., Bultman, S. and Redinbo, M., 2020. Targeted inhibition of gut bacterial β -glucuronidase activity enhances anticancer drug efficacy. *Proceedings of the National Academy of Sciences*, 117(13), pp.7374-7381.

- Bin, Y., Jiti, Z., Jing, W., Cuihong, D., Hongman, H., Zhiyong, S. and Yongming, B., 2004. Expression and characteristics of the gene encoding azoreductase from *Rhodobacter sphaeroides* AS1.1737. *FEMS Microbiology Letters*, 236(1), pp.129-136.
- Blackwell, J. and Horgan, R., 1991. A novel strategy for production of a highly expressed recombinant protein in an active form. *FEBS Letters*, 295(1-3), pp.10-12.
- Blümel, S. and Stolz, A., 2003. Cloning and characterization of the gene coding for the aerobic azoreductase from *Pigmentiphaga kullae* K24. *Applied Microbiology and Biotechnology*, 62(2-3), pp.186-190.
- Blümel, S., Knackmuss, H. and Stolz, A., 2002. Molecular Cloning and Characterization of the Gene Coding for the Aerobic Azoreductase from *Xenophilus azovorans* KF46F. *Applied and Environmental Microbiology*, 68(8), pp.3948-3955.
- Brambilla, E., Hippe, H., Hagelstein, A., Tindall, B. and Stackebrandt, E., 2001. 16S rDNA diversity of cultured and uncultured prokaryotes of a mat sample from Lake Fryxell, McMurdo Dry Valleys, Antarctica. *Extremophiles*, 5(1), pp.23-33.
- Burke, D., Fouhy, F., Harrison, M., Rea, M., Cotter, P., O'Sullivan, O., Stanton, C., Hill, C., Shanahan, F., Plant, B. and Ross, R., 2017. The altered gut microbiota in adults with cystic fibrosis. *BMC Microbiology*, 17(1).
- Butt, T., Edavettal, S., Hall, J. and Mattern, M., 2005. SUMO fusion technology for difficult-to-express proteins. *Protein Expression and Purification*, 43(1), pp.1-9.
- Carrió, M., Cubarsi, R. and Villaverde, A., 2000. Fine architecture of bacterial inclusion bodies. *FEBS Letters*, 471(1), pp.7-11.
- Chen, H., Hopper, S. and Cerniglia, C., 2005. Biochemical and molecular characterization of an azoreductase from *Staphylococcus aureus*, a tetrameric NADPH-dependent flavoprotein. *Microbiology*, 151(5), pp.1433-1441.

Chen, H., Nwe, P., Yang, Y., Rosen, C., Bielecka, A., Kuchroo, M., Cline, G., Kruse, A., Ring, A., Crawford, J. and Palm, N., 2019. A Forward Chemical Genetic Screen Reveals Gut Microbiota Metabolites That Modulate Host Physiology. *Cell*, 177(5), pp.1217-1231.e18.

Chen, H., Wang, R. and Cerniglia, C., 2004. Molecular cloning, overexpression, purification, and characterization of an aerobic FMN-dependent azoreductase from *Enterococcus faecalis*. *Protein Expression and Purification*, 34(2), pp.302-310.

Chengalroyen, M. and Dabbs, E., 2012. The microbial degradation of azo dyes: minireview. *World Journal of Microbiology and Biotechnology*, 29(3), pp.389-399.

Chrnyk, B., Evans, J., Lillquist, J., Young, P. and Wetzel, R., 1993. Inclusion body formation and protein stability in sequence variants of interleukin-1 beta. *Journal of Biological Chemistry*, 268(24), pp.18053-18061.

Colebrook, I., 1936. Treatment with prontosil of puerperal infections *1 due to hemolytic streptococci. *The Lancet*, 228(5910), pp.1319-1322.

Corradini, M., 2019. Synthetic Food Colors. *Encyclopedia of Food Chemistry*, pp.291-296.

Costa, S., Almeida, A., Castro, A. and Domingues, L., 2014. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. *Frontiers in Microbiology*, 5.

Coultate, T. and Blackburn, R., 2018. Food colorants: their past, present and future. *Coloration Technology*, 134(3), pp.165-186.

Crescente, V., Holland, S., Kashyap, S., Polycarpou, E., Sim, E. and Ryan, A., 2016. Identification of novel members of the bacterial azoreductase family in *Pseudomonas aeruginosa*. *Biochemical Journal*, 473(5), pp.549-558.

Cryan, J. and O'Mahony, S., 2011. The microbiome-gut-brain axis: from bowel to behavior. *Neurogastroenterology & Motility*, 23(3), pp.187-192.

Dancygier, H., 2010. Hepatic Biotransformation. *Clinical Hepatology*, pp.127-130.

Daniel, J., 1962. The excretion and metabolism of edible food colors. *Toxicology and Applied Pharmacology*, 4(5), pp.572-594.

Derman, A., Prinz, W., Belin, D. and Beckwith, J., 1993. Mutations that Allow Disulfide Bond Formation in the Cytoplasm of *Escherichia coli*. *Science*, 262(5140), pp.1744-1747.

Dobkin, J., Saha, J., Butler, V., Neu, H. and Lindenbaum, J., 1983. Digoxin-Inactivating Bacteria: Identification in Human Gut Flora. *Science*, 220(4594), pp.325-327.

EFSA Journal, 2009. Scientific Opinion on the re-evaluation of Quinoline Yellow (E 104) as a food additive. 7(11), p.1329.

Escobedo-Hinojosa, W. and Pardo-López, L., 2017. Analysis of bacterial metagenomes from the Southwestern Gulf of Mexico for pathogens detection. *Pathogens and Disease*, 75(5).

Esmaili, S., Ashrafi-Kooshk, M., Khaledian, K., Adibi, H., Rouhani, S. and Khodarahmi, R., 2016. Degradation products of the artificial azo dye, Allura red, inhibit esterase activity of carbonic anhydrase II: A basic in vitro study on the food safety of the colorant in terms of enzyme inhibition. *Food Chemistry*, 213, pp.494-504.

Fatima, M., Farooq, R., Lindström, R. and Saeed, M., 2017. A review on biocatalytic decomposition of azo dyes and electrons recovery. *Journal of Molecular Liquids*, 246, pp.275-281.

Feingold, B., 1976. Hyperkinesis and Learning Disabilities Linked to the Ingestion of Artificial Food Colors and Flavors. *Journal of Learning Disabilities*, 9(9), pp.551-559.

Feketea, G. and Tsaouri, S., 2017. Common food colorants and allergic reactions in children: Myth or reality?. *Food Chemistry*, 230, pp.578-588.

Ferrer, M., Chernikova, T., Yakimov, M., Golyshin, P. and Timmis, K., 2003. Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nature Biotechnology*, 21(11), pp.1266-1267.

Francis, D. and Page, R., 2010. Strategies to Optimize Protein Expression in *E. coli*. *Current Protocols in Protein Science*, 61(1).

Franco, J., da Silva, B., Dias, E., de Castro, A., Ramalho, T. and Zanoni, M., 2018. Influence of auxochrome group in disperse dyes bearing azo groups as chromophore center in the biotransformation and molecular docking prediction by reductase enzyme: Implications and assessment for environmental toxicity of xenobiotics. *Ecotoxicology and Environmental Safety*, 160, pp.114-126.

Gariyban, L. and Avashia, N., 2013. Polymerase Chain Reaction. *Journal of Investigative Dermatology*, 133(3), pp.1-4.

Gensollen, T., Iyer, S., Kasper, D. and Blumberg, R., 2016. How colonization by microbiota in early life shapes the immune system. *Science*, 352(6285), pp.539-544.

Gičević, A., Hindija, L. and Karačić, A., 2019. Toxicity of Azo Dyes in Pharmaceutical Industry. *IFMBE Proceedings*, pp.581-587.

Göker, M., Gronow, S., Zeytun, A., Nolan, M., Lucas, S., Lapidus, A., Hammon, N., Deshpande, S., Cheng, J., Pitluck, S., Liolios, K., Pagani, I., Ivanova, N., Mavromatis, K., Ovchinikova, G., Pati, A., Tapia, R., Han, C., Goodwin, L., Chen, A., Palaniappan, K., Land, M., Hauser, L., Jeffries, C., Brambilla, E., Rohde, M., Detter, J., Woyke, T., Bristow, J., Markowitz, V., Hugenholtz, P., Eisen, J., Kyrpides, N. and Klenk, H., 2011. Complete genome sequence of *Odoribacter splanchnicus* type strain (1651/6T). *Standards in Genomic Sciences*, 4(2), pp.200-209.

Guo, Y., Lee, H. and Jeong, H., 2020. Gut microbiota in reductive drug metabolism. *Progress in Molecular Biology and Translational Science*, pp.61-93.

Hall, T., 1999. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, pp.95-98.

Heijtz, R., Wang, S., Anuar, F., Qian, Y., Bjorkholm, B., Samuelsson, A., Hibberd, M., Forssberg, H. and Pettersson, S., 2011. Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences*, 108(7), pp.3047-3052.

Hiippala, K., Barreto, G., Burrello, C., Diaz-Basabe, A., Suutarinen, M., Kainulainen, V., Bowers, J., Lemmer, D., Engelthaler, D., Eklund, K., Facciotti, F. and Satokari, R., 2020. Novel *Odoribacter splanchnicus* Strain and Its Outer Membrane Vesicles Exert Immunoregulatory Effects in vitro. *Frontiers in Microbiology*, 11.

Himri, I., BELLAHCEN, S., Souna, F., Belmekki, F., Aziz, M., Bnouham, M., Zoheir, J., Berkhia, Z., Mekhfi, H. and Saalaoui, E., 2011. A 90-day oral toxicity study of tartrazine, a synthetic food dye, in wistar rats. *International Journal of Pharmacy and Pharmaceutical Sciences*, 3(3).

Holland, S., 2017. *Investigating Azoreductases and NAD(P)H dependent Quinone Oxidoreductases in Pseudomonas aeruginosa*. Ph.D. Kingston University London.

Hugenholtz, F. and de Vos, W., 2017. Mouse models for human intestinal microbiota research: a critical evaluation. *Cellular and Molecular Life Sciences*, 75(1), pp.149-160.

Hugenholtz, F. and de Vos, W., 2017. Mouse models for human intestinal microbiota research: a critical evaluation. *Cellular and Molecular Life Sciences*, 75(1), pp.149-160.

Jaglin, M., Rhimi, M., Philippe, C., Pons, N., Bruneau, A., Goustard, B., Daugé, V., Maguin, E., Naudon, L. and Rabot, S., 2018. Indole, a Signaling Molecule Produced by the Gut Microbiota, Negatively Impacts Emotional Behaviors in Rats. *Frontiers in Neuroscience*, 12.

Javaid, R. and Qazi, U., 2019. Catalytic Oxidation Process for the Degradation of Synthetic Dyes: An Overview. *International Journal of Environmental Research and Public Health*, 16(11), p.2066.

Jia, B. and Jeon, C., 2016. High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. *Open Biology*, 6(8), p.160196.

Knoll, A., 2015. Paleobiological Perspectives on Early Microbial Evolution. *Cold Spring Harbor Perspectives in Biology*, 7(7), p.a018093.

Kobylewski, S. and Jacobson, M., 2010. *Food dyes*. Washington, D.C.: Center for Science in the Public Interest.

Koppel, N., Maini Rekdal, V. and Balskus, E., 2017. Chemical transformation of xenobiotics by the human gut microbiota. *Science*, 356(6344).

Korpela, K., 2021. Impact of Delivery Mode on Infant Gut Microbiota. *Annals of Nutrition and Metabolism*, pp.1-9.

Kriss, M., Hazleton, K., Nusbacher, N., Martin, C. and Lozupone, C., 2018. Low diversity gut microbiota dysbiosis: drivers, functional implications and recovery. *Current Opinion in Microbiology*, 44, pp.34-40.

Lade, H., Waghmode, T., Kadam, A. and Govindwar, S., 2012. Enhanced biodegradation and detoxification of disperse azo dye Rubine GFL and textile industry effluent by defined fungal-bacterial consortium. *International Biodeterioration & Biodegradation*, 72, pp.94-107.

Lamark, T., Ingebrigtsen, M., Bjørnstad, C., Melkko, T., Mollnes, T. and Nielsen, E., 2001. Expression of Active Human C1 Inhibitor Serpin Domain in *Escherichia coli*. *Protein Expression and Purification*, 22(2), pp.349-358.

Lathrop, S., Bloom, S., Rao, S., Nutsch, K., Lio, C., Santacruz, N., Peterson, D., Stappenbeck, T. and Hsieh, C., 2011. Peripheral education of the immune system by colonic commensal microbiota. *Nature*, 478(7368), pp.250-254.

Lee, C., Kim, J., Lee, W., Nelson, K., Yoon, J. and Sedlak, D., 2008. Bactericidal Effect of Zero-Valent Iron Nanoparticles on *Escherichia coli*. *Environmental Science & Technology*, 42(13), pp.4927-4933.

Li, M., Wang, B., Zhang, M., Rantalainen, M., Wang, S., Zhou, H., Zhang, Y., Shen, J., Pang, X., Zhang, M., Wei, H., Chen, Y., Lu, H., Zuo, J., Su, M., Qiu, Y., Jia, W., Xiao, C., Smith, L., Yang, S., Holmes, E., Tang, H., Zhao, G., Nicholson, J., Li, L. and Zhao, L., 2008. Symbiotic gut microbes modulate human metabolic phenotypes. *Proceedings of the National Academy of Sciences*, 105(6), pp.2117-2122.

Liang, D., Leung, R., Guan, W. and Au, W., 2018. Involvement of gut microbiome in human health and disease: brief overview, knowledge gaps and research opportunities. *Gut Pathogens*, 10(1).

Llamas, N., Garrido, M., Nezio, M. and Band, B., 2009. Second order advantage in the determination of amaranth, sunset yellow FCF and tartrazine by UV-vis and multivariate curve resolution-alternating least squares. *Analytica Chimica Acta*, 655(1-2), pp.38-42.

Lobstein, J., Emrich, C., Jeans, C., Faulkner, M., Riggs, P. and Berkmen, M., 2012. SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial Cell Factories*, 11(1).

Maini Rekdal, V., Bess, E., Bisanz, J., Turnbaugh, P. and Balskus, E., 2019. Discovery and inhibition of an interspecies gut bacterial pathway for Levodopa metabolism. *Science*, 364(6445).

Matsumoto, K., Mukai, Y., Ogata, D., Shozui, F., Nduko, J., Taguchi, S. and Ooi, T., 2009. Characterization of thermostable FMN-dependent NADH azoreductase from the moderate thermophile *Geobacillus stearothermophilus*. *Applied Microbiology and Biotechnology*, 86(5), pp.1431-1438.

Mazmanian, S., Liu, C., Tzianabos, A. and Kasper, D., 2005. An Immunomodulatory Molecule of Symbiotic Bacteria Directs Maturation of the Host Immune System. *Cell*, 122(1), pp.107-118.

Mcbain, A. and Macfarlane, G., 1998. Ecological and physiological studies on large intestinal bacteria in relation to production of hydrolytic and reductive enzymes involved in formation of genotoxic metabolites. *Journal of Medical Microbiology*, 47(5), pp.407-416.

McCann, D., Barrett, A., Cooper, A., Crumpler, D., Dalen, L., Grimshaw, K., Kitchin, E., Lok, K., Porteous, L., Prince, E., Sonuga-Barke, E., Warner, J. and Stevenson, J., 2007. Food additives and hyperactive behaviour in 3-year-old and 8/9-year-old children in the community: a randomised, double-blinded, placebo-controlled trial. *The Lancet*, 370(9598), pp.1560-1567.

McFall-Ngai, M., Hadfield, M., Bosch, T., Carey, H., Domazet-Lošo, T., Douglas, A., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S., Hentschel, U., King, N., Kjelleberg, S., Knoll, A., Kremer, N., Mazmanian, S., Metcalf, J., Nealson, K., Pierce, N., Rawls, J., Reid, A., Ruby, E., Rumpho, M., Sanders, J., Tautz, D. and Wernegreen, J., 2013. Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences*, 110(9), pp.3229-3236.

Meyer, S., Probert, P., Lakey, A., Axon, A., Leitch, A., Williams, F., Jowsey, P., Blain, P., Kass, G. and Wright, M., 2017. Hepatic effects of tartrazine (E 102) after systemic exposure are independent of oestrogen receptor interactions in the mouse. *Toxicology Letters*, 273, pp.55-68.

Misal, S. and Gawai, K., 2018. Azoreductase: a key player of xenobiotic metabolism. *Bioresources and Bioprocessing*, 5(1).

Moeller, A., Caro-Quintero, A., Mjungu, D., Georgiev, A., Lonsdorf, E., Muller, M., Pusey, A., Peeters, M., Hahn, B. and Ochman, H., 2016. Cospeciation of gut microbiota with hominids. *Science*, 353(6297), pp.380-382.

Mogk, A., Mayer, M. and Deuerling, E., 2002. Mechanisms of Protein Folding: Molecular Chaperones and Their Application in Biotechnology. *ChemBioChem*, 3(9), pp.807-814.

Morais, C., de Rosso, V., Estadella, D. and Pisani, L., 2016. Anthocyanins as inflammatory modulators and the role of the gut microbiota. *The Journal of Nutritional Biochemistry*, 33, pp.1-7.

Morgan, X., Tickle, T., Sokol, H., Gevers, D., Devaney, K., Ward, D., Reyes, J., Shah, S., LeLeiko, N., Snapper, S., Bousvaros, A., Korzenik, J., Sands, B., Xavier, R. and Huttenhower, C., 2012. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biology*, 13(9), p.R79.

Morrison, J. and John, G., 2016. Growth and physiology of *Clostridium perfringens* wild-type and Δ azoC knockout: an azo dye exposure study. *Microbiology*, 162(2), pp.330-338.

Morrison, J., Wright, C. and John, G., 2012. Identification, Isolation and characterization of a novel azoreductase from *Clostridium perfringens*. *Anaerobe*, 18(2), pp.229-234.

Moutinho, I., Bertges, L. and Assis, R., 2007. Prolonged use of the food dye tartrazine (FD&C yellow n° 5) and its effects on the gastric mucosa of Wistar rats. *Brazilian Journal of Biology*, 67(1), pp.141-145.

Mueller, G. and Miller, J., 1949. The reductive cleavage of 4-dimethylaminoazobenzene by rat liver: the intracellular distribution of the enzyme system and its requirement for triphosphopyridine nucleotide. *Journal of Biological Chemistry*, 180(3), pp.1125-1136.

Mullis, K., 1990. The Unusual Origin of the Polymerase Chain Reaction. *Scientific American*, 262(4), pp.56-65.

Nachiyar, C. and Rajakumar, G., 2005. Purification and characterization of an oxygen insensitive azoreductase from *Pseudomonas aeruginosa*. *Enzyme and Microbial Technology*, 36(4), pp.503-509.

Nagai, F., Morotomi, M., Watanabe, Y., Sakon, H. and Tanaka, R., 2010. *Alistipes indistinctus* sp. nov. and *Odoribacter laneus* sp. nov., common members of the human intestinal microbiota isolated from faeces. *International Journal of Systematic and Evolutionary Microbiology*, 60(6), pp.1296-1302.

Nakanishi, M., Yatome, C., Ishida, N. and Kitade, Y., 2001. Putative ACP Phosphodiesterase Gene (acpD) Encodes an Azoreductase. *Journal of Biological Chemistry*, 276(49), pp.46394-46399.

Oganesyan, N., Ankoudinova, I., Kim, S. and Kim, R., 2007. Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization. *Protein Expression and Purification*, 52(2), pp.280-285.

Ogbulie, T., Okore, C. and Ejele, A., 2020. Comparative Heavy Metal Removal Efficiencies of Biosurfactants Produced by *Odoribacter Splanchnicus* DSM 20712, Bacterium Clone JX981747 and Soil Washing Agents.

Owczarzy, R., Tataurov, A., Wu, Y., Manthey, J., McQuisten, K., Almabrazi, H., Pedersen, K., Lin, Y., Garretson, J., McEntaggart, N., Sailor, C., Dawson, R. and Peek, A., 2008. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Research*, 36(Web Server), pp.W163-W169.

Page, R., Moy, K., Sims, E., Velasquez, J., McManus, B., Grittini, C., Clayton, T. and Stevens, R., 2004. Scalable high-throughput micro-expression device for recombinant proteins. *BioTechniques*, 37(3), pp.364-370.

Parada Venegas, D., De la Fuente, M., Landskron, G., González, M., Quera, R., Dijkstra, G., Harmsen, H., Faber, K. and Hermoso, M., 2019. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. *Frontiers in Immunology*, 10.

Patterson, A., Gonzalez, F. and Idle, J., 2010. Xenobiotic Metabolism: A View through the Metabolometer. *Chemical Research in Toxicology*, 23(5), pp.851-860.

Peng, L., Xu, Z., Fang, X., Wang, F. and Cen, P., 2004. High-level expression of soluble human β -defensin-2 in *Escherichia coli*. *Process Biochemistry*, 39(12), pp.2199-2205.

Pfeiffer, J. and Sonnenburg, J., 2011. The Intestinal Microbiota and Viral Susceptibility. *Frontiers in Microbiology*, 2.

Pike, L., 2019. *Characterising antibiotic susceptibility and resistance in human commensal gut bacteria*. PhD. Gonville and Caius College, University of Cambridge.

Possemiers, S., Bolca, S., Verstraete, W. and Heyerick, A., 2011. The intestinal microbiome: A separate organ inside the body with the metabolic potential to influence the bioactivity of botanicals. *Fitoterapia*, 82(1), pp.53-66.

Possemiers, S., Bolca, S., Verstraete, W. and Heyerick, A., 2011. The intestinal microbiome: A separate organ inside the body with the metabolic potential to influence the bioactivity of botanicals. *Fitoterapia*, 82(1), pp.53-66.

Przystaś, W., Zabłocka-Godlewska, E. and Grabińska-Sota, E., 2011. Biological Removal of Azo and Triphenylmethane Dyes and Toxicity of Process By-Products. *Water, Air, & Soil Pollution*, 223(4), pp.1581-1592.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, J., Brunak, S., Doré, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Bork, P., Ehrlich, S. and Wang, J., 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, 464(7285), pp.59-65.

Quagliariello, A., Del Chierico, F., Russo, A., Reddel, S., Conte, G., Lopetuso, L., Ianiro, G., Dallapiccola, B., Cardona, F., Gasbarrini, A. and Putignani, L., 2018. Gut Microbiota Profiling and Gut–Brain Crosstalk in Children Affected by Pediatric Acute-Onset Neuropsychiatric Syndrome and Pediatric Autoimmune Neuropsychiatric Disorders Associated With Streptococcal Infections. *Frontiers in Microbiology*, 9.

Radmoski, J. and Deichmann, W., 1956. Cathartic Action and Metabolism of Certain Coal Tar Food Dyes. *Journal of Pharmacology and Experimental Therapeutics November*, 118(3), pp.322-327.

Ramakrishna, B., 2013. Role of the gut microbiota in human nutrition and metabolism. *Journal of Gastroenterology and Hepatology*, 28, pp.9-17.

Raran-Kurussi, S., Keefe, K. and Waugh, D., 2015. Positional effects of fusion partners on the yield and solubility of MBP fusion proteins. *Protein Expression and Purification*, 110, pp.159-164.

Rimington, C. and Hemmings, A., 1938. Porphyrinuria Following Sulphanilamide:. *The Lancet*, 231(5979), pp.770-776.

Rost, B., 1999. Twilight zone of protein sequence alignments. *Protein Engineering, Design and Selection*, 12(2), pp.85-94.

Rovina, K., Siddiquee, S. and Shaarani, S., 2016. Extraction, Analytical and Advanced Methods for Detection of Allura Red AC (E129) in Food and Beverages Products. *Frontiers in Microbiology*, 7.

Roxon, J. and Ryan, A., 1967. Reduction of water-soluble azo dyes by intestinal bacteria. *Food and Cosmetics Toxicology*, 5, pp.367-369.

Roxon, J., Ryan, A. and Wright, S., 1967. Enzymatic reduction of tartrazine by *Proteus vulgaris* from rats. *Food and Cosmetics Toxicology*, 5, pp.645-656.

Ryan, A., 2016. Azoreductases in drug metabolism. *British Journal of Pharmacology*, 174(14), pp.2161-2173.

Ryan, A., Kaplan, E., Nebel, J., Polycarpou, E., Crescente, V., Lowe, E., Preston, G. and Sim, E., 2014. Identification of NAD(P)H Quinone Oxidoreductase Activity in Azoreductases from *P. aeruginosa*: Azoreductases and NAD(P)H Quinone Oxidoreductases Belong to the Same FMN-Dependent Superfamily of Enzymes. *PLoS ONE*, 9(6), p.e98551.

Ryan, A., Wang, C., Laurieri, N., Westwood, I. and Sim, E., 2010. Reaction mechanism of azoreductases suggests convergent evolution with quinone oxidoreductases. *Protein & Cell*, 1(8), pp.780-790.

Rychlik, W., Spencer, W. and Rhoads, R., 1990. Optimization of the annealing temperature for DNA amplification in vitro;. *Nucleic Acids Research*, 18(21), pp.6409-6412.

Samuelson, J., 2010. Recent Developments in Difficult Protein Expression: A Guide to *E. coli* Strains, Promoters, and Relevant Host Mutations. *Methods in Molecular Biology*, pp.195-209.

Sarkar, S., Banerjee, A., Halder, U., Biswas, R. and Bandopadhyay, R., 2017. Degradation of Synthetic Azo Dyes of Textile Industry: a Sustainable Approach Using Microbial Enzymes. *Water Conservation Science and Engineering*, 2(4), pp.121-131.

Sasaki, Y., Kawaguchi, S., Kamaya, A., Ohshita, M., Kabasawa, K., Iwama, K., Taniguchi, K. and Tsuda, S., 2002. The comet assay with 8 mouse organs: results with 39 currently used

food additives. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 519(1-2), pp.103-119.

Scher, J., Nayak, R., Ubeda, C., Turnbaugh, P. and Abramson, S., 2020.

Pharmacomicrobiomics in inflammatory arthritis: gut microbiome as modulator of therapeutic response. *Nature Reviews Rheumatology*, 16(5), pp.282-292.

Schopf, J., Kudryavtsev, A., Agresti, D., Wdowiak, T. and Czaja, A., 2002. Laser–Raman imagery of Earth's earliest fossils. *Nature*, [online] 416(6876), pp.73-76. Available at: <<https://www.nature.com/articles/416073a>>.

Seelbinder, B., Chen, J., Brunke, S., Vazquez-Urbe, R., Santhaman, R., Meyer, A., de Oliveira Lino, F., Chan, K., Loos, D., Imamovic, L., Tsang, C., Lam, R., Sridhar, S., Kang, K., Hube, B., Woo, P., Sommer, M. and Panagiotou, G., 2020. Antibiotics create a shift from mutualism to competition in human gut communities with a longer-lasting impact on fungi than bacteria. *Microbiome*, 8(1).

Shuman, S., 1994. Novel Approach to Molecular Cloning and Polynucleotide Synthesis Using Vaccinia DNA Topoisomerase. *J Biol Chem*, 269(51), pp.32678-84.

Sisley, P. and Porcher, C., 1911. Du sort des matières colorantes dans l'organisme animal. *Comptes Rendus de l'Académie des Sciences*, 152, pp.1062-1064.

Smith, P., Howitt, M., Panikov, N., Michaud, M., Gallini, C., Bohlooly-Y, M., Glickman, J. and Garrett, W., 2013. The Microbial Metabolites, Short-Chain Fatty Acids, Regulate Colonic T reg Cell Homeostasis. *Science*, 341(6145), pp.569-573.

Sokolosky, J. and Szoka, F., 2013. Periplasmic production via the pET expression system of soluble, bioactive human growth hormone. *Protein Expression and Purification*, 87(2), pp.129-135.

Stal, L., 2007. Cyanobacteria. In: J. Seckbach, ed., *Algae and cyanobacteria in extreme environments*. Dordrecht: Springer, pp.659-680.

Stecher, B., Chaffron, S., Käppeli, R., Hapfelmeier, S., Friedrich, S., Weber, T., Kirundi, J., Suar, M., McCoy, K., von Mering, C., Macpherson, A. and Hardt, W., 2010. Like Will to Like: Abundances of Closely Related Species Can Predict Susceptibility to Intestinal Colonization by Pathogenic and Commensal Bacteria. *PLoS Pathogens*, 6(1), p.e1000711.

Studier, F., 1991. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *Journal of Molecular Biology*, 219(1), pp.37-44.

Šuleková, M., Smrčová, M., Hudák, A., Heželová, M. and Fedorová, M., 2017. Organic Colouring Agents in the Pharmaceutical Industry. *Folia Veterinaria*, 61(3), pp.32-46.

Suzuki, Y., Yoda, T., Ruhul, A. and Sugiura, W., 2001. Molecular Cloning and Characterization of the Gene Coding for Azoreductase from *Bacillus sp.* OY1-2 Isolated from Soil. *Journal of Biological Chemistry*, 276(12), pp.9059-9065.

Tartoff, K. and Hobbs, C., 1987. Improved Media for Growing Plasmid and Cosmid Clones. *Bethesda Res. Lab*, 9(12).

Tegel, H., Tourle, S., Ottosson, J. and Persson, A., 2010. Increased levels of recombinant human proteins with the *Escherichia coli* strain Rosetta(DE3). *Protein Expression and Purification*, 69(2), pp.159-167.

Tegel, H., Tourle, S., Ottosson, J. and Persson, A., 2010. Increased levels of recombinant human proteins with the *Escherichia coli* strain Rosetta(DE3). *Protein Expression and Purification*, 69(2), pp.159-167.

Tréfouël, J., Nitti, F. and Bovet, D., 1935. Activité du p-aminophénylsulfamide sur les infections streptococciques expérimentales de la souris et du lapin. *Comptes Rendus des Seances de la Societe de Biologie, Paris*, 120, pp.756-758.

Tuteja, S. and Ferguson, J., 2019. Gut Microbiome and Response to Cardiovascular Drugs. *Circulation: Genomic and Precision Medicine*, 12(9).

- Van der Waaij, D., Berghuis-de Vries, J. and Lekkerkerk-van der Wees, J., 1971. Colonization resistance of the digestive tract in conventional and antibiotic-treated mice. *Journal of Hygiene*, 69(3), pp.405-411.
- Wahlström, A., Sayin, S., Marschall, H. and Bäckhed, F., 2016. Intestinal Crosstalk between Bile Acids and Microbiota and Its Impact on Host Metabolism. *Cell Metabolism*, 24(1), pp.41-50.
- Walther, B., Karl, J., Booth, S. and Boyaval, P., 2013. Menaquinones, Bacteria, and the Food Supply: The Relevance of Dairy and Fermented Food Products to Vitamin K Requirements. *Advances in Nutrition*, 4(4), pp.463-473.
- Walton, K., Walker, R., van de Sandt, J., Castell, J., Knapp, A., Kozianowski, G., Roberfroid, M. and Schilter, B., 1999. The application of in vitro data in the derivation of the Acceptable Daily Intake of food additives. *Food and Chemical Toxicology*, 37(12), pp.1175-1197.
- Wan, L., Ge, W., Zhang, S., Sun, Y., Wang, B. and Yang, G., 2020. Case-Control Study of the Effects of Gut Microbiota Composition on Neurotransmitter Metabolic Pathways in Children With Attention Deficit Hyperactivity Disorder. *Frontiers in Neuroscience*, 14.
- Wang, C., Hagemeyer, C., Rahman, N., Lowe, E., Noble, M., Coughtrie, M., Sim, E. and Westwood, I., 2007. Molecular Cloning, Characterisation and Ligand-bound Structure of an Azoreductase from *Pseudomonas aeruginosa*. *Journal of Molecular Biology*, 373(5), pp.1213-1228.
- Waterhouse, A., Procter, J., Martin, D., Clamp, M. and Barton, G., 2009. Jalview Version 2-- a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9), pp.1189-1191.
- Werner, H. and Reichertz, C., 1971. Buttersäurebildende Bacteroides-Kulturen [Butyric acid producing *bacteroides* cultures]. *Zentralblatt für Bakteriologie*, 217(2), pp.206-16.
- Werner, H. and Reichertz, C., 1975. [A new butyric acid-producing *bacteroides* species: *B. splanchnicus* n. sp. (author's transl)]. *Zentralblatt für Bakteriologie*, 231, pp.133-44.

- Whitman, W., Coleman, D. and Wiebe, W., 1998. Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences*, 95(12), pp.6578-6583.
- Wolf, P., Kolossov, V., Zhou, Z., Ly, L., Doden, H., Devendran, S., Breister, A., Lucio, L., Polak, P., Matatov, S., Anantharaman, K., Ridlon, J. and Gaskins, R., 2020. Abstract 3342: The colorectal cancer associated microbe *Odoribacter splanchnicus* produces genotoxic hydrogen sulfide via cysteine metabolism. *Cancer Research*, 80(16_Supplement), pp.3342-3342.
- Wu, D., Ugozzoli, L., Pal, B., Qian, J. and Wallace, R., 1991. The Effect of Temperature and Oligonucleotide Primer Length on the Specificity and Efficiency of Amplification by the Polymerase Chain Reaction. *DNA and Cell Biology*, 10(3), pp.233-238.
- Xing, C., Wang, M., Ajibade, A., Tan, P., Fu, C., Chen, L., Zhu, M., Hao, Z., Chu, J., Yu, X., Yin, B., Zhu, J., Shen, W., Duan, T., Wang, H. and Wang, R., 2021. Microbiota regulate innate immune signaling and protective immunity against cancer. *Cell Host & Microbe*, 29(6), pp.959-974.e7.
- Xiong, S., 2005. Solubility of disulfide-bonded proteins in the cytoplasm of *Escherichia coli* and its “oxidizing” mutant. *World Journal of Gastroenterology*, 11(7), p.1077.
- Yamjala, K., Nainar, M. and Ramiseti, N., 2016. Methods for the analysis of azo dyes employed in food industry – A review. *Food Chemistry*, 192, pp.813-824.
- Zahran, S., Ali-Tammam, M., Ali, A. and Aziz, R., 2021. Compositional variation of the human fecal microbiome in relation to azo-reducing activity: a pilot study. *Gut Pathogens*, 13(1).
- Zhang, J., Lacroix, C., Wortmann, E., Ruscheweyh, H., Sunagawa, S., Sturla, S. and Schwab, C., 2019. Gut microbial beta-glucuronidase and glycerol/diol dehydratase activity contribute to dietary heterocyclic amine biotransformation. *BMC Microbiology*, 19(1).
- Zhang, L., Huang, Y., Zhou, Y., Buckley, T. and Wang, H., 2013. Antibiotic Administration Routes Significantly Influence the Levels of Antibiotic Resistance in Gut Microbiota. *Antimicrobial Agents and Chemotherapy*, 57(8), pp.3659-3666.

Zimmermann, T., Kulla, H. and Leisinger, T., 1982. Properties of Purified Orange II Azoreductase, the Enzyme Initiating Azo Dye Degradation by *Pseudomonas* KF46. *European Journal of Biochemistry*, 129(1), pp.197-203.

Zou, L., Spanogiannopoulos, P., Pieper, L., Chien, H., Cai, W., Khuri, N., Pottel, J., Vora, B., Ni, Z., Tsakalozou, E., Zhang, W., Shoichet, B., Giacomini, K. and Turnbaugh, P., 2020. Bacterial metabolism rescues the inhibition of intestinal drug absorption by food and drug additives. *Proceedings of the National Academy of Sciences*, 117(27), pp.16009-16018.