Dissecting the indigo pathway

By

Tim Daykin B. App. Sci (Hons)

Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy

June 2011

Department of Biotechnology and Environmental Biology

RMIT University

Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

Tim Daykin

Date:

Acknowledgements

I would like to sincerely thank my supervisor Associate Professor Trevor Stevenson for his help and guidance throughout my candidature, especially in the last six months. I appreciate his candour, the thorough scrutiny of my data and his ongoing advice, all of which was critical to me submitting my thesis. Big thanks also go to Professor David Stalker for the chance to 'look after his baby'. This project was inspired by Dave and I appreciate the opportunity he gave me and the knowledge I gained from his know-how. Special gratitude also goes to Trevor and Dave for the constant light relief they provided from the stress of life in the lab and the perspective I gained from their experiences.

I would like to acknowledge Associate Professor Peter Smooker and Dr Manvendra Saxena for providing me with the Red disruption system that was crucial to my research. To Dr Greg Nugent, Dr Veera Chikkala and Kim Stevenson I appreciate the ear you provided when I needed someone to bounce ideas off. To all of my fellow scientists and students who called the plant biotech lab home over the year's thanks for all the support you afforded. Recognition also must to go to Natalie Kikidopoulos, Aya Taki, Shannon Fernandes and Dr Fiona Baird for putting up with my antics. Furthermore, thanks to Shannon for all the technical help she provided. In addition, thanks must go to Dr Lisa Dias and Dr Kylie White for helping me with formatting.

To the inhabitants of the 'Swamp': Teigs, Liam, Adam, Whit and my brother Sadie, thanks for the good times. I needed them and you guys provided them, it kept me sane. Finally to my family, especially my parents David and Eileen, the biggest thanks go to them, for their ongoing support and encouragement.

Thesis Abstract

Indigo is a vibrant blue pigment synthesised by plants and bacteria. This study examined the mechanisms by which plants generate indole, an intermediate common to both the indigo and tryptophan biosynthetic pathways. Cleavage of Indole-3-glycerol phosphate (I3GP) by the tryptophan synthase (TS) αβ enzyme complex forms indole for subsequent tryptophan biosynthesis. Specifically, indole synthesis is catalysed by the TSα sub-unit, encoded by the *trp*A gene, leading to this sub-unit being described as an 'indole synthase'. Typically, TSα sub-units only catalyse indole synthesis when bound to a specific TSβ sub-unit counterpart. However, there does exist in plants enzymes capable of catalysing indole synthesis independently. In *Zea mays*, there is an indole synthase (IS) enzyme ten times more efficient than the TSαβ enzyme complex at catalysing indole biosynthesis from I3GP.

When the maize indole synthase gene was expressed in a TSα deficient *E. coli* strain it complemented the *trp*A mutation, indicating that the enzyme encoded by this gene produced indole as an intermediate in tryptophan biosynthesis in the absence of a plant specific TSβ sub-unit. The removal of tryptophanase activity by targeted mutagenesis in a host *E. coli* strain eliminated the alternative route for tryptophan biosynthesis from indole. Complementation by the *Z. mays* IS of the *trp*A mutation suggests that the IS forms a functional enzyme complex with an *E. coli* TSβ sub-unit.

A *Polygonum tinctorium* cDNA library was screened via hybridisation with the *Z. mays* IS gene and by functional analysis. The unsuccessful hybridisation suggested the absence of a gene homologous to the maize IS gene within *P. tinctorium*, whilst failure of the cDNA library to rescue the *E. coli trp*A mutant strain indicated the non-existance of a functional clone.

Table of Contents

Abbreviations

BASF™	Baden Aniline and Soda Factory
BM-3	*Bacillus megaterium* Cytochrome P450
BxBx	Benzoxazineless
CAA	Case Amino Acids
CAM	Camphor
cDNA	Complementary DNA
CdRP	1-(*O*-carboxyphenylamino)-1-deoxyribulose-5-phosphate
CYP	Cytochrome P450
DIMBOA	2,4-*di*hydroxy-7-*m*ethoxy-1,4-*b*enzox*a*zin-3-one
DNA	Deoxyribonucleic Acid
DPM	Disintegrations per min
EtOH	Ethanol
g	Gram
HPLC	High Performance Liquid Chromatography
IAA	Indole-3-acetic acid
IAOx	Indole-3-acetaldoxime
I3GP	Indole-3-Glycerolphosphate
IGL	Indole-3glycerol Phosphate Lyase
IND	Indole oxidase from *C. violaceum*
INO	Indole oxygenase
IS	Indole synthase
LB	Luria Bertoni
M	Molar
MetOH	Methanol
mg	Milligram

mL	Millilitre
mM	Millimolar
MMLV	Moloney Murine Leukemia Virus
M9 MM	M9 Minimal media
NADPH	Nicotinamide adenine dinucleotide phosphate
ng	Nanogram
nm	Nanomolar
NPR	NADPH P450 reductase
PCR	Polymerase Chain Reaction
PRPP	Phosphoribosylpyrophosphate
rpm	Revolutions per min
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
µg	Microgram
µL	Microlitre

List of Figures

List of Tables

1 Chapter One: Review of literature

1.1 History of indigo dyeing

Indigo is a distinctive blue substance renowned for its attractiveness and vibrant colour. Although it may be considered a 'cheap' blue pigment it has been the centre of uprisings and conflict through the ages. It has been prevalent throughout history as a dye for use in printing and textile staining (Clark *et al*, 1993). Indigo dyeing practices are believed to have originated in India and other vicinities of Asia, but were specific to regions inhabited by plants harbouring the blue compound. Evidence of indigo dyed garments has also been discovered in ancient Egyptian and Incan tombs and it is generally accepted that indigo was of great interest to other ancient civilisations found in Greece, Rome, Africa and the Middle East with the dye not only used for staining materials but for used as pen ink as well as being utilised for body art (Balfour-Paul, 1998).

From the medieval times until the late 15th century the use of indigo in the majority of Europe was scarce, it was primarily imported from Asia to only a few nations through Middle Eastern countries that imposed excessive trade laws and provided treacherous and unreliable land routes (Sequin-Frey, 1981; Clark *et al*, 1993). When Dutch and Portuguese explorers discovered an easy and direct sea route to India, primitive trade companies were formed and the utilisation of indigo in Europe increased significantly. Since the beginning of the 16th century, powerhouse European nations have made extensive profits by establishing indigo plantations in regions with tropical climates (Balfour-Paul, 1998), benefitting from the popularity of dyeing fabrics such as denim and other everyday materials. Indigo has remained a fashionable and popular dye in modern times, with pioneering interest in employing the dye not only for fabric staining but for applications in the field of science, as a potential direction for the 'blue' colouration of flowers and plant fibres being pursued (Gillam and Guengerich, 2001; Warzecha *et al,* 2007).

1.1.1 Cultural significance of indigo

Many culturally significant events have revolved around indigo. In India, the indigo revolt of Bengal stemmed from the uprising of indigo farmers against indigo planters in 1859, an especially important event in Bengali history (Bhattacharya, 1977). Indigo harvesting also had links to slavery in the Deep South of America prior to the emancipation proclamation implemented by then President Abraham Lincoln (Sharrer, 1971). Indigo developed into a culturally significant pigment in many other countries due to its association with dyeing of cotton garments. Throughout the 'Edo Period' in Japan, established in 1603 and ending in 1868, cotton was an alternative to prohibited fabrics such as silk and satin and became desirable due to its ability to accommodate the blue dye with relative ease compared to silk and satin (Sandberg, 1989).

Indigo still remains the predominant colour of clothing during the summer months in Japan, signifying the 'deep blue sea' revered so much by the Japanese people (Hayashi, 1978). In addition to the dye representing certain cultures connection with the environment, indigo was seen as a way to indicate status amongst primitive societies. Strength and prosperity was attributed to hunters, nomads and warriors in West African countries when their headwear was dyed the characteristic blue shade (Balfour-Paul, 1998). Most of these customs are still practiced today and indicate the importance indigo holds when linked to cultural traditions.

1.1.2 Advances in dyeing practices

Indigo is classified as a vat dye due to the archaic method of processing indigo by fermenting vast amounts of raw plant material in large vats (Sequin-Frey, 1981). This method, believed to have originated in India was quite crude, utilising the only readily available bases such as lye or in the majority of cases stale urine to treat the plant matter. These practices were also observed in other particularly unsophisticated cultures (Balfour-Paul, 1998; Oberthur *et al*, 2004). The fermentation process reduced vegetation to a colourless soluble substance which

was then applied to material and textiles, after which a 'blue' colour would develop (Ferreira *et al*, 2004). To obtain dyed fabric of the uppermost quality, the entire fermentation process occurred over several months with air tight vats maintained preventing premature oxidation and ensuring an efficient reduction process, guaranteeing the best possible extraction of indigo (Liles, 1990).

Modern indigo vat dyeing techniques still make use of air tight containers, however primitive bases have been replaced by more accurately measured substitutes such as bran, madder and lime. Modern vats generally have a capacity of over 1,000 litres permitting enormous amounts of indigo to be developed. By employing superior bases and fermenting under more stringent conditions, a significant decrease in dyeing time is achievable (Clark *et al*, 1993).

1.1.3 The era of synthetic indigo

Rudimentary extraction of indigo was superseded when the chemical pathway to indigo synthesis was discovered in the 1870s by the Nobel Prize winning German chemist Johann Friedrich Wilhelm Adolf Von Baeyer, whose experiments with indigo led to chemical substitutes being preferred over natural indigo production (Li, 2003). The method developed by Adolf Von Baeyer to synthesise indigo was via an aldol condensation reaction currently accepted as the Baeyer-Drewson model (Figure 1.1). When acetone was added to O-nitrobenzaldehyde and vigorously agitated indigo appeared as a dark blue precipitate (Russell and Kaupp, 1969; Steingruber, 2004).

This method of indigo synthesis was not economically viable and was impractical; however rather than repressing additional research this ground breaking initiative encouraged large chemical companies to further investigate chemically induced indigo

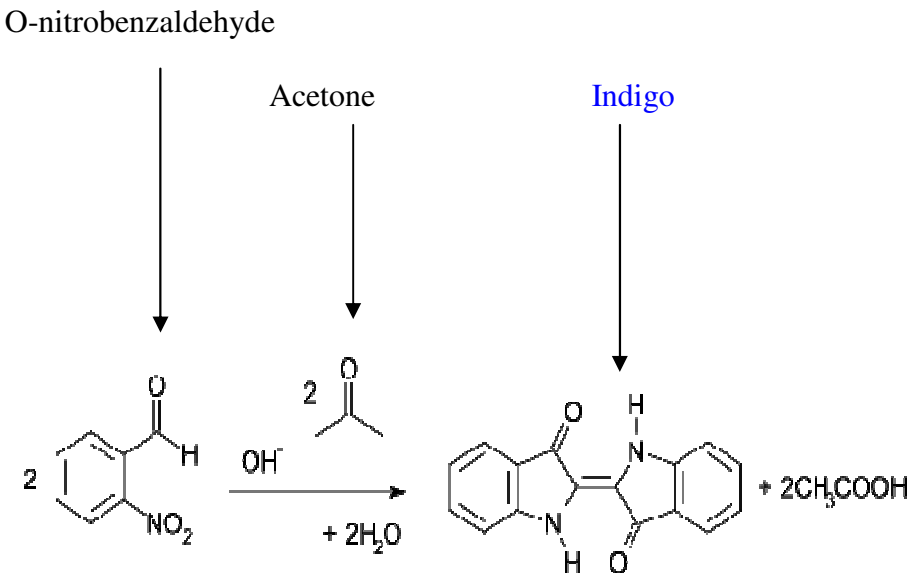


Figure 1.1 The Baeyer-Drewson model for synthetic indigo production

synthesis (Reed, 1992). In 1890, Swiss scientist Karl Heumann, contracted to the German based Baden Aniline and Soda Factory (BASF™), then the worlds leading chemical company, proposed a similar approach that resulted in a more commercially feasible process (Nagendrappa, 2003). Heumann utilised chemical alternatives that were cheaper and more readily available than those employed by Von Baeyer. When anthranilic acid was used as the preliminary chemical and combined with chloracetic acid, phenylglycine-O-carboxylic acid was produced. Subsequent heating in conjunction with the addition of a strong base led to the synthesis of indoxyl which was converted to indigo through exposure to air and natural oxidation (McKee and Zanger, 1991).

In 1897 BASF™, having invested nearly 18 million gold marks into indigo research, implemented synthetic production. This caused a dramatic decrease in the amounts of indigo sourced naturally, dropping from 19,000 tons per annum to 1,000 tons at the start of World War II. Currently, the German based company still has a monopoly over the indigo industry, manufacturing over a third of the 20,000 tons of synthetic indigo produced annually by chemical factories world wide (Balfour-Paul, 1998). The faith shown by BASF™ in the late 19th century has been rewarded with the company currently grossing in excess of €50 billion annually and employing over 100,000 workers to continue the pioneering work initiated by Von Baeyer and Heumann (Moore-Braun, 2010).

1.1.3.1 Indigo white

Due to the insoluble nature of indigo, an alternative, water soluble form of the dye was developed and patented for use in textile dyeing. Much like traditional dyeing practices, this water soluble form of indigo arises from a reduction-oxidation and is affixed to garments in airtight sealable vats (Clark *et al*, 1993). When reduced with sodium dithionite, indigo converts to the compound indigo white. Alternatively named leucoindigo, this particular derivative is able to dissolve in water and allows for easier dyeing of fabrics. Simulating

early dyeing techniques, indigo white is applied to material and exposed to air; subsequent oxidation leads to the reconversion back to the insoluble indigo compound ultimately producing indigo coloured garments and textiles (Clark *et al*, 1993; Hardwood *et al*, 1999). These advances in organic chemistry now provide indigo not only for textile and fabric staining but also for the synthesis of food dyes and a variety of other colourants such as laboratory indicator dyes (Steingruber, 2004).

1.2 Natural indigo production

Prior to the introduction of synthetic indigo, plants were the sole naturally occurring resource harvested for the blue dye (Warzecha *et al*, 2007). Whilst indigo is recognised as an aesthetically pleasing compound, there is no clear evidence of it having any valid function in living plants (Warzecha *et al*, 2007), although the indigo pathway may have a speculative role in defence mechanisms opposing plant pathogens. This theory has arisen due to an association with specific secondary metabolite precursors involved in protection from a vast range of predatory insects, bacteria and fungi (Melanson *et al*, 1997). However, it is more probable that indigo acts as a repugnatory defence mechanism in damaged leaves to protect the remainder of an otherwise healthy plant.

1.2.1 Indigo producing plants

Historically, the most common agricultural crops used for the generation of indigo dye are from the genus *Indigofera* (Balfour-Paul, 1998). Containing over 700 species of flowering plants and renowned for its many pain relief properties, *Indigofera* belongs to the family Fabaceae, the third largest family of angiosperms behind Orchidaceae and Asteraceae and is commonly referred to as the legume or pulse family (Parotta, 2001). Like all indigo bearing plants, *Indigofera* was first harvested for indigo dye in India before early trade was implemented (Gilbert and Cooke, 2001). While indigo synthesis from plants has spread world wide, the use of *Indigofera* as a source for the dye, has for the majority, remained localised to India, Pakistan and regions of South East Asia due to its reliance on the tropical and subtropical climate and its inability to flourish in harsher weather (Chanayath *et al*, 2002). The maceration of leaf tissue from a diverse range of these indigo producing plants is directly responsible for the production of indigo (Chanayath *et al*, 2002). Rising concerns over pollution output stemming from the use of the potential carcinogenic sodium dithionite and insoluble indigo by synthetic indigo manufacturing factories has led to a renewed interest in natural indigo production from these indigo bearing plants (Garcia-Macias and John, 2004).

1.2.1.1 Isatis tinctoria

The angiosperm, *Isatis tinctoria* (Figure 1.2), was another major crop used for the generation of indigo (Hurry, 1973). Unlike *Indigofera*, *I. tinctoria* belongs to the *Brassiceae* family and is commonly referred to as Dyer's Woad. It is a perennial herb indigenous to the southern hills of Russia, but over time has spread across Europe and into central Asia. Dyer's Woad, or simply Woad, is cultivated in temperate climates and is able to withstand colder and harsher environments than *Indigofera* and was the preferred source for indigo extraction from the initial stages of natural indigo dyeing until the implementation of synthetic indigo production (Gilbert and Cooke, 2001).

Until the introduction of other indigoferous plants through trade and exchange, Woad was the prevailing resource for the dye in Northern Europe during the 'golden age' of indigo production from the 12th century onwards (Gilbert and Cooke, 2001). Such was the dominance of the Woad industry in these parts that the prohibition of other indigo bearing plants in some European countries was implemented, with severe reprimands and punishment inflicted on those who manufactured the dye from the banned plants (Ferreira *et al*, 2004). Great economic importance was placed on *I. tinctoria*, especially in Germany and England, where Woad manufacturing factories were central to employment and financial stability in many cities.

Whilst Woad was generally recognised as one of the main plants used for the generation of indigo dye, it was also a common medicinal plant used to treat tumours, ulcers, haemorrhoids and snake bites, as well as acting as a powerful anti-inflammatory (Oberthur *et al*, 2004; Zou and Koh, 2007). An extremely versatile and robust plant, Woad existed in England as a crude resource for indigo until the introduction of the synthetic indigo industry in 1930. Since then, the use of Woad vats has become obsolete with the implements and devices used in vat dyeing now resting in museums and commemorative houses (Gilbert and Cooke, 2001).

Figure 1.2 Mature *I. tinctoria* commonly known as Dyer's Woad

1.2.1.2 Polygonum tinctorium

The genus of plants denoted *Polygonum* are from the family *Polygonaceae*. Collectively known as Knotweeds or Smartweeds, many species of plants from the *Polygonaceae* family are invasive weeds. However, a vast selection are considered valuable due to their anti-oxidant and anti-microbial properties, as well as being a pleasant additive to herbal teas and in some Asian cultures being used as a snack food as well as their capacity to produce indigo (Burns *et al*, 2002).

One species within the *Polygonaceae* family, *Polygonum tinctorium*, otherwise referred to as 'Japanese Indigo', was a frequently used plant in the generation of indigo (Kim *et al*, 1996; Mitich, 1998). The scientific name is derived from the Greek description of the characteristic swollen jointed nodes displayed on their stems, *Poly* denoting many and *gonu* meaning knee, referencing the enlarged knot like protuberances (Mitich, 1998). Recently reclassified under the genus *Persicaria*, *P. tinctorium* is a popular ornamental plant still used by artists from Far East Asia for natural indigo dyeing (Figure 1.3). Although synthetic indigo production is preferred over natural indigo production by chemical companies, many of those who make use of indigo believe the natural blue colour

Figure 1.3 Mature *P. tinctorium* commonly known as Japanese Indigo

harvested from *P. tinctorium* is far superior to the synthetically produced hue (Shim *et al*, 1998). This belief may be attributed to the extremely high concentration of indigo precursors within the plant, up to 2-5% dry weight of the vegetative material (Xia and Zenk, 1992; Gilbert *et al*, 2004).

Japanese Indigo was most commonly utilised in China and Japan but its use gradually extended to more temperate parts of Europe when natural indigo production was at its most popular (Angelini *et al,* 2004). Due to the heavy reliance on the provision of water required for optimal growth, *P. tinctorium* is usually grown and cultivated in tropical environments. This is opposed to other indigo bearing plants that can flourish without the need for constant uptake of water, plants that can be grown in less tropical environments due to their ability to withstand more stressful conditions (Campeol *et al*, 2006).

Although Japanese Indigo contains substantial reservoirs of natural indigo, harvest yields were unreliable and it provided a perennial source of frustration due to its inability to flourish in the European hub of the indigo trade in comparison with other indigoferous plants. This was a major reason why it was less popular than Woad in the European countries with harsher climates (Balfour-Paul, 1998). Improved knowledge of growing conditions and advances in harvesting techniques have allowed the re-introduction of *P. tinctorium* into Europe as a source of natural indigo (Angelini *et al*, 2004).

1.2.2 Indigo synthesis in plants

Renewed interest in natural indigo production has provided a stimulus for further examination into the mechanism by which indigo is created and other potential biological applications of this blue compound (Warzecha *et al*, 2007). Reports of synthetic indigo being linked to allergies and illness (Bunting, 2000), as well as heightened awareness of the detrimental

effects on the environment caused by indigo factories has led to increased interest in the plant indigo pathway and the possible manipulation of indigo biosynthesis in plants.

Indigo production in plants relies on two key chemical components, indole and indoxyl. Indoles and indolic compounds have become the subject of increasing interest since the discovery of their association with plant growth hormones, specifically auxins (Kende and Zeevaart, 1997). Whilst there are several documented indolic compounds within plants many are of unknown function, further intensifying the study of those indoles that do play a role in the biosynthesis of tryptophan and of those involved in the stabilising the auxin regulatory pathway (Powell, 1963; Roychowdhury and Basak, 1975).

Indole is an aromatic heterocyclic organic compound that is a derivative of the tryptophan biosynthetic pathway (Radwanski and Last, 1995). The formation of indole and the subsequent conversion to the isomeric indoxyl has long been the focus of research due to the fact that indole is the common precursor to indigo in all indigo producing plants (Xia and Zenk, 1992). These compounds display toxicity towards the host cell, however harmless glucoside derivatives are formed spontaneously preventing cellular damage ensuing from the toxic nature of indole and indoxyl (Kim *et al*, 2009). It is evident in the associated literature that a limited amount of study has been solely devoted to indoxyl. This may be due to the fact that aside from being a major precurser in the indigo pathway, like indigo, indoxyl appears to have no known function in plants (Gilbert and Cooke, 2000).

1.2.2.1 Indican and isatin B

In indigo producing plants the glucoside moieties resulting from indoxyl conversion accumulate in the vacuole as colourless derivatives in contrast to the vibrant 'blue' indigo compound produced further along the indigo pathway (Kokubun *et al*, 1997). Since the synthetic indigo boom of the mid to late 1800's, extensive research has provided much

information regarding these colourless glucoside residues known as indican and isatin (Kim *et al*, 1996; Kokubun *et al*, 1998; Gilbert *et al*, 2004). In *P. tinctorium* indican is the sole indoxy-glucoside derivative of indoxyl that is generated and stored. The concentration of this amassed indican is significantly higher compared to the amount of indican stored in *I. tinctoria* (Gilbert *et al*, 2004).

While *P.tinctorium* harbours indican as the solitary indoxyl derivative, *I. tinctoria* accumulates indican as well as isatin, specifically isatin B, another colourless indoxy-glucoside (Oberthur *et al*, 2004). Within the plant indigo pathway, the conversion of indoxyl to either derivative is controlled by unspecified glucosyl-transferase enzymes. This stabilises the indoxyl component permitting accumulation in the undisrupted vacuole of the usually toxic indoxyl (Figure 1.4) (Warzecha *et al*, 2007).

A discovery in the mid 1990's described the mechanism by which indoxyl reforms from indican permitting indigo production in *P. tinctorium* (Figure 1.5) (Minami *et al*, 1996). Upon maceration of the plant tissue, the vacuole in which indican is stored is disrupted, and the indoxy-glucoside is released and exposed to a β-glucosidase enzyme with high substrate specificity for indican. This indican specific enzyme, occasionally referred to as indican synthase (Davies, 2004), located in the chloroplast of plant mesophyll cells are able to produce glucose and indoxyl through a hydrolysis reaction (Minami *et al*, 1996; Minami *et* al, 1997). Subsequent dimerisation of two indoxyl molecules mediated by atmospheric oxygen results in the formation of indigo (Maugard *et al*, 2002).

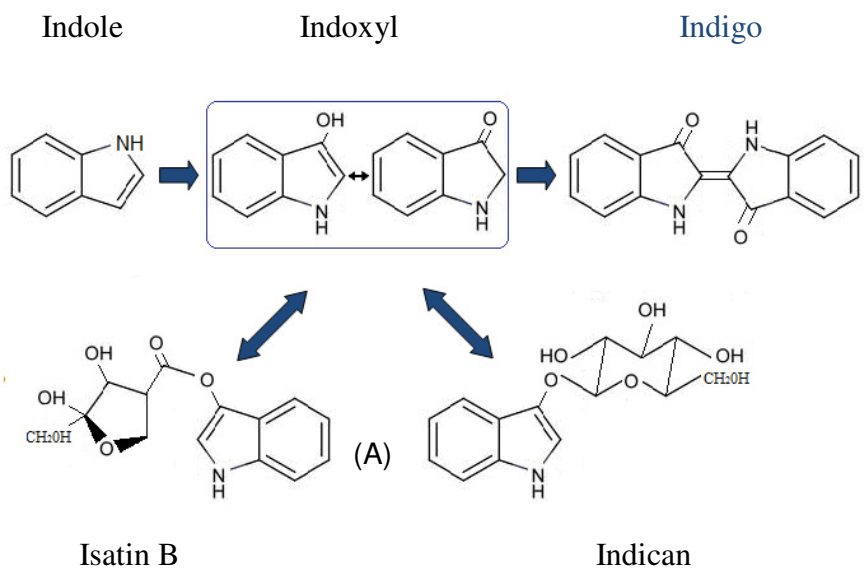


Figure 1.4 Indigo conversion in *I. tinctoria* and *P. tinctorium*. (A) Specific glucosyl transferase enzymes mediate the conversion of indoxyl to non-toxic indoxy-glucosides for storage in the vacuole. Conversion of indoxyl and subsequent indigo formation occurs when the vacuole is disrupted

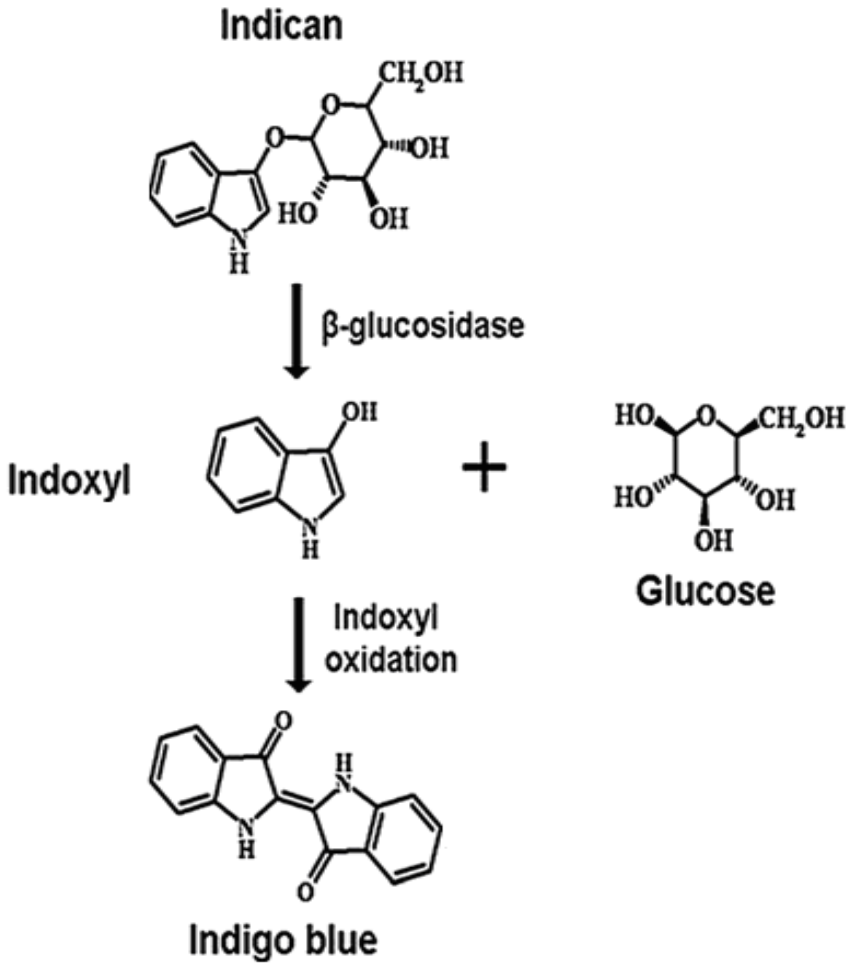


Figure 1.5 β-glucosidase mediated hydrolysis of indican and subsequent indigo production from indoxyl in *P. tinctorium* after disruption of the vacuole by leaf tissue maceration

Recent studies have concluded that although both derivatives are stored in *I. tinctoria*, the concentration of both indican and isatin B is three fold less than the concentration of stored indican in leaves of *P. tinctorium* (Kim *et* al, 1996; Gilbert *et al*, 2004). This may explain why *P.tinctorium* is able to produce more vibrant and more abundant quantities of indigo compared to *I. tinctoria* and as a result is the preferred resource for the dye.

Isatin B in Woad has been the focus of intense research since the introduction of indigo dye factories in the late 19th century. Originally named 'Woad indican', it was long believed to be identical in structure and behaviour to indigo precursers in *Polygonum* and *Indigofera* plants, a belief refuted by studies in the early 1900's when a definitive distinction was made between the indigo precursors through the discovery of differing chromatographic behaviour of the indoxyl derivatives and the different chemical properties of their carbohydrate moiety (Karapetyan *et al*, 2011). During this same time period the structure of indican was discovered and was alternatively termed 'indoxyl-β-D- glucoside'. The structure of 'Woad indican' was proposed in the 1930's with conclusions offering the scientific term of 'indoxyl-5-ketogluconate' and the common name of Isatin B in 1967 (Oberthur *et al*, 2004).

1.2.3 Indole conversion

1.2.3.1 Hydroxylase mediated indole conversion

In mammals and higher plants cytochrome P450's are heme-thiolate proteins often involved in xenobiotic metabolism (Gillam *et al*, 1999; Gillam *et al*, 2000), and play a pivotal role in the biosynthesis of fatty acids, flavonoids, glucosides and alkaloids as well as being vital in defense mechanisms, successful growth and development (Morant *et al*, 2003; Yamada *et al*, 2000).

Cytochrome P450 enzymes that control hydroxylation of indolic compounds are regularly referred to as CYP's. These CYP's are included in a superfamily of related enzymes that are

categorised into families and sub-families on the basis of their amino acid and nucleotide homology (Hanioka *et al*, 2000; Gillam and Guengerich, 2001). The hydroxylation of indole to indigo mediated by CYP's is a process found throughout the phylogenetic spectrum from mammalian tissue to microbial systems (Gillam *et al*, 2000; Kelly *et al*, 2007).

Within the vitally important tryptophan biosynthetic pathway, indole is a key intermediate having a significant role in the degradation of tryptophan and subsequent metabolism of its derivatives. Due to the association of such indolic compounds and tryptophan metabolism, research has been dedicated to further understanding the role indolic compounds play within the tryptophan biosynthetic pathway and other pathways of corresponding importance (Normanly *et al*, 1993). The relationship between a variety of indolic compounds and CYP's is of major interest.

Recent research has described the identification a specific CYP variant from *Arabidopsis thaliana* (*cyp83B1)* that catalyses the oxidative degradation of indole glucosinolates and indole-3-acetaldoxime (IAOx) (Smolen and Bender, 2002). Both are secondary metabolites associated with defense against plant pathogens (Ludwig-Muller *et* al, 1999) while IAOx is also an intermediate in the synthesis of indole-3-acetic acid (IAA), an extremely important plant growth hormone (Gillam and Guengerich, 2001; Pollmann *et al*, 2009). Similarly, Hull *et al* (2000), describe the isolation of CYPs from *A.thaliana*, namely CYP79B2 and CYP79B3 that can convert tryptophan to IAOx. Considering the proposed link the indigo pathway has with plant defense, the production of indolic compounds by these CYP variants and the conversion of indole in plants is an area for further investigation.

The formation of indigo by recombinant mammalian cytochrome P450's has also been reported (Gillam *et al*, 2000). When recombinant human CYP's are co-expressed with Nicotinamide adenine dinucleotide phosphate (NADPH)-P450 reductase in *E.coli*, the

formation of a blue pigment in bacterial cultures is observed. This product was confirmed to be indigo by high performance liquid chromatography (HPLC) and mass spectroscopy (Gillam *et al*, 2000).

1.2.4 Microbial biosynthesis of indigo

In addition to indigo producing plants, an abundance of microorganisms harbour the necessary chemical precursors for indigo synthesis and are capable of producing indigo as a by-product of various metabolic processes. Evidence of indigo accumulation in various environments, from soil, mammalian intestinal tracts and even urine due to bacterial indole oxidation is confirmation of microbial indigo synthesis (Warzecha *et al*, 2007). Bacterial oxygenases and hydroxylases that catalyse indigo synthesis are significant in that they are the basis for an alternative to chemical indigo production (Sun *et al*, 2003).

1.2.4.1 Naphthalene dioxygenase

A particularly well studied model for indigo production from indole in microbes is the breakdown of the polycyclic aromatic hydrocarbon naphthalene by bacteria such as *Pseudomonas* sp (Figure 1.6). The complex pathway by which various *Pseudomonas* sp. and other naphthalene degrading bacteria synthesise indigo is initiated by naphthalene dioxygenase (NDO) (Ensley *et al*, 1983; Han *et al*, 2008). This particular enzyme is part of a multicomponent system that mediates the transfer of electrons from nicotinamide adenine dinucleotide phosphate (NADPH) to a terminal oxygenase (Parales *et al*, 2000). When indole is supplied to this complex the dioxygenation occurs and transforms indole into *cis*-2,3-dihydroxy-2,3-dihydroindole. This *cis*-diol intermediate is rapidly dehydrated

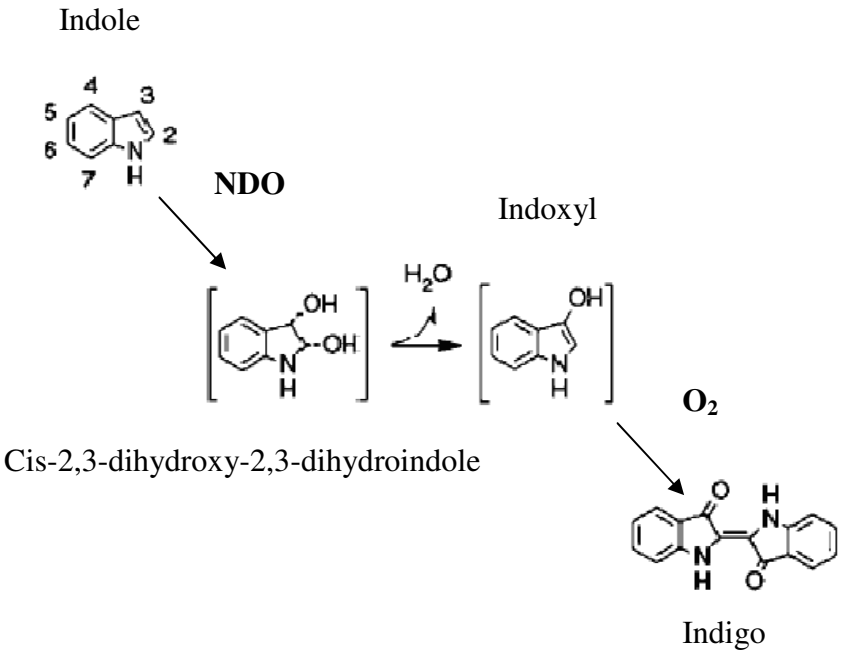


Figure 1.6 NDO mediated formation of indigo from indole in bacteria

to indoxyl, subsequent dimerisation of indoxyl molecules then forms indigo (Pathak and Madamwar, 2009).

An NDO from *Pseudomonas putida* when expressed in tryptophan producing strains of *E. coli* has been reported to catalyse the synthesis of indigo (Berry *et al*, 2002). This synthesis uses glucose as an initial precursor, in contrast to the use of the more expensive and environmentally unfriendly indole (O'Connor and Hartmans, 1998). However, this process is not yet commercially viable due to low productivity stemming from the instability of the dioxygenase enzyme (Kim *et al*, 2003).

1.2.4.2 Bacterial CYP's

Bacterial mediated indigo synthesis is a complex mechanism that requires natural redox partners and involves the biosythesis of numerous intermediates and by-products. In the human intestinal system, indole formed from tryptophan by various gut bacteria is absorbed into the liver and metabolised to the indican derivative indol-3-yl sulphate (metabolic indicant) and excreted in the urine. Bacteria present in urine then mediate the formation of indoxyl and subsequent oxidation results in the production of indigo (Gillam *et al*, 2000).

Microbial CYP's are diverse and have origins dating from Archaebacteria, with extensive research discovering the importance these enzymes have in the production of metabolic compounds and detoxification of xenobiotics as well as the hydroxylation of indolic compounds (Kelly *et al*, 2005). The most widely studied bacterial CYP is the monoxygenase $P450_{CAM}$ from *P. putida*, an enzyme involved in camphor hydroxylation and the primary model for the study of other P450 monooxygenases (Celik *et al*, 2005). Recently there has been the report of the identification of $P450_{CAM}$ mutants that have the ability to catalyse the hydroxylation of indole for subsequent indigo synthesis within a bacterial culture (Celik and

Speight, 2005). Cytochrome P450 BM-3 from *Bacillus megaterium* also catalyses NADPH mediated indole hydroxylation in an alkaline environment (Li *et al*, 2005).

The research performed on bacterial CYP's has indicated that they are extremely effective at producing indigo in liquid culture. Like the CYP described by Gillam *et al*, (1999), indigo production is reliant upon the co-expression with a natural redox partner. Recombinant human P450 co-expressed with NADPH p450 reductase (NPR) in *E. coli* catalysed the formation of indole derivatives, due to oxidative coupling within the active site of P450's (Gillam *et al*, 2000). Theoretically, bacterial CYP's have the ability to aid in indigo synthesis, perhaps through a mechanism comparable to the means by which various P450's from other species are able to produce the blue pigment.

1.2.4.3 Indole oxygenase

Indole oxygenase (INO) is an enzyme that is closely associated with the formation of blue pigments in bacterial cultures (Hart *et al*, 1990; Doukyu *et al*, 2003). Studies have reported a 2.1kb DNA fragment isolated from a *Rhodococcus* strain that in *E.coli* gave rise to the formation of indigo (Hart *et al*, 1990) and indirubin, a less concentrated pink coloured dye also formed from the oxidation of indole in plants and microbes (Kokubun *et al*, 1998). Although capable of producing both pigments, INO has mainly been studied for its potential in indigo associated colouration of plant tissues (Cheah *et al*, 1998). Further investigation has revealed that a 1,161bp gene encodes an enzyme responsible for the conversion of indole to indigo. Although no indigo production is observed in *Rhodococcus* sp, accumulation of indigo in *E. coli* cultures utilising the *Rhodococcus* INO has been described (Hart *et al*, 1992).

Chromobacterium violaceum is a Gram-negative bacteria is also linked with INO activity and is commonly found in tropical environments. Its name is derived from the purple pigment it produces, vioacein that serves as an antibiotic and discourages protozoan predation (Cheah *et*

al, 1998). Within a specific gene cluster of *C. violaceum*, a particular gene, *IND*, encoding Indole oxygenase activity has been characterised and shown to display the same tendancy to catalyse indigo synthesis in *E. coli* as the INO from *Rhodococcus*. Both of these catalytic processes from *Rhodococcus* and *Chromobacterium* display an efficient mechanism for indigo synthesis.

Little is known about the mechanism by which indole is converted to indoxyl as a precursor to indigo synthesis in plants. An insight into this process may be obtained by exploring comparable enzyme systems from other organisms. Perhaps an analogous single polypeptide catalysed reaction exists in plants. What is acknowledged is that the conversion of indole to indoxyl in plants is secondary compared to the primary role of indole: that is as a fundamental intermediate critical to tryptophan biosynthesis.

1.3 Tryptophan Biosynthesis

Tryptophan is an amino acid essential to basic protein assembly in both eukaryotes and prokaryotes (Crawford, 1989). In microorganisms and higher plants tryptophan is a direct derivative of the shikimate pathway (Hermann, 1995). The shikimate pathway exclusive to higher plants and microorganisms is an essential metabolic process (Hermann, 1995) and was discovered through the concerted efforts of Bernhard Davis and David Sprinson and many other collaborators. It is not only essential in the formation of tryptophan but other vital aromatic amino acids such as tyrosine and phenylalanine (Figure 1.7) (Gibson and Pittard, 1968) and a variety of other metabolites. These include the important structural biopolymer lignin (Bentley, 1990) and critical precursors for a number of secondary metabolites such as plant pigments and defence compounds acting to protect against insects and other herbivores (Hermann, 1995).

1.3.1 The tryptophan pathway

The pathway for tryptophan biosynthesis has been extensively studied and has been found to show remarkable conservation between prokaryotes and eukaryotes (Crawford, 1989). That the tryptophan biosynthetic pathway in plants is localised to the chloroplast is confirmed by evidence from *in vitro* assays and immunoblot experiments verifying the presence of all pathway enzymes within the chloroplast fraction (Zhao and Last, 1995). Products of the tryptophan biosynthetic pathway are not only incorporated into polypeptides but are also metabolised into different indolic compounds utilised in diverse roles by plants (Radwanski and Last, 1995).

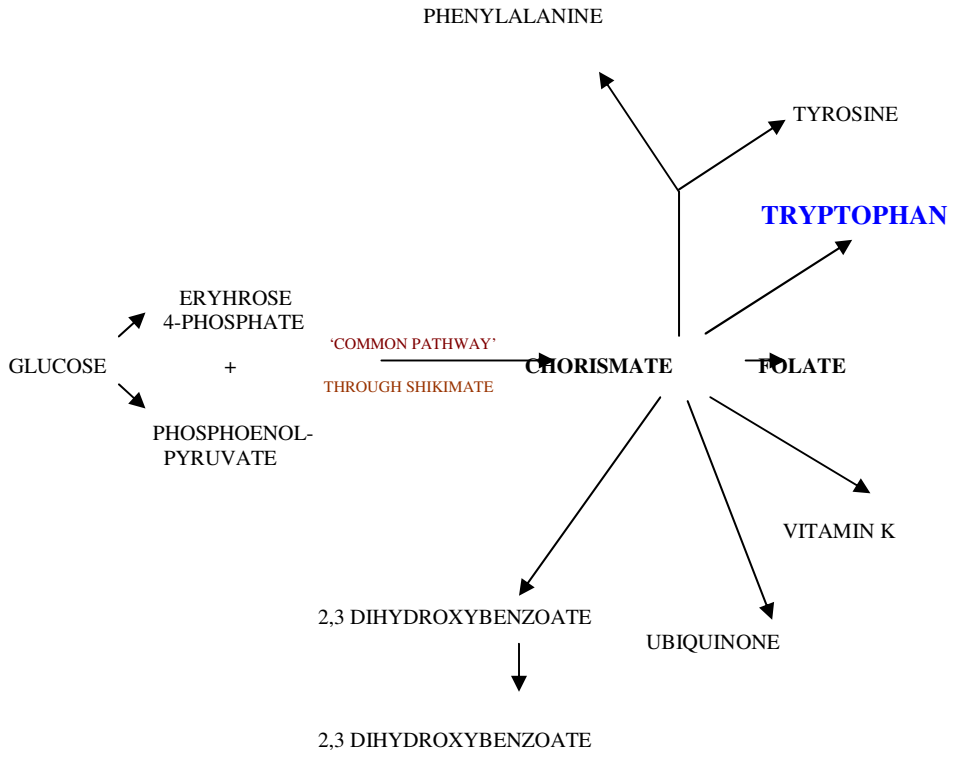


Figure 1.7 Branching properties of chorismate for the formation of essential aromatic amino acids in *E. coli*

Tryptophan synthesis utilises glucose as the starting substrate through a 'common pathway' resulting in the formation of chorismate. Chorismate acts as a branch point from which several individual pathways deviate (Gibson and Pittard, 1968). Anthranilate is readily synthesised from chorismate through catalysis by anthranilate synthetase. This enzymatic mechanism is very complex, so much so that intermediates between the two compounds have been suggested (Ratledge, 1964; Srinivasan, 1965; Edwards and Jackman, 1965).

Enzymatic evidence suggested that conversion of chorismate to anthranilate is completed in two steps and that it is likely a protein complex is involved (Figure 1.8: A) (Nagano *et al*, 1970). An αβ enzyme complex is required for successful conversion, more specifically the larger α subunit binds chorismate and mediates its aromatization followed by glutamine amidotransferase activity via the β subunit acting as an auxillary protein (Radwanski and Last, 1995; Gibson and Pittard, 1968). The conversion of anthranilate to indole-3-glycerolphosphate is performed in three steps crucial to metabolism of secondary indolic compounds. Firstly, PR anthranilate transferase catalyses the transfer of a phosphoribosyl moiety to anthranilate from phosphoribosylpyrophosphate (PRPP) (Figure 1.8: B), followed by the irreversible action of anthranilate isomerise catalysing the rearrangement of N-(5'-phosphoribosyl)-anthranilate to 1-(*O*-carboxyphenylamino)-1-deoxyribulose-5-phosphate (CdRP) (Figure 1.8: C). Finally, Indole-3-Glycerolphosphate (I3GP) is formed from CdRP mediated by I3GP synthase (Figure 1.8: D).

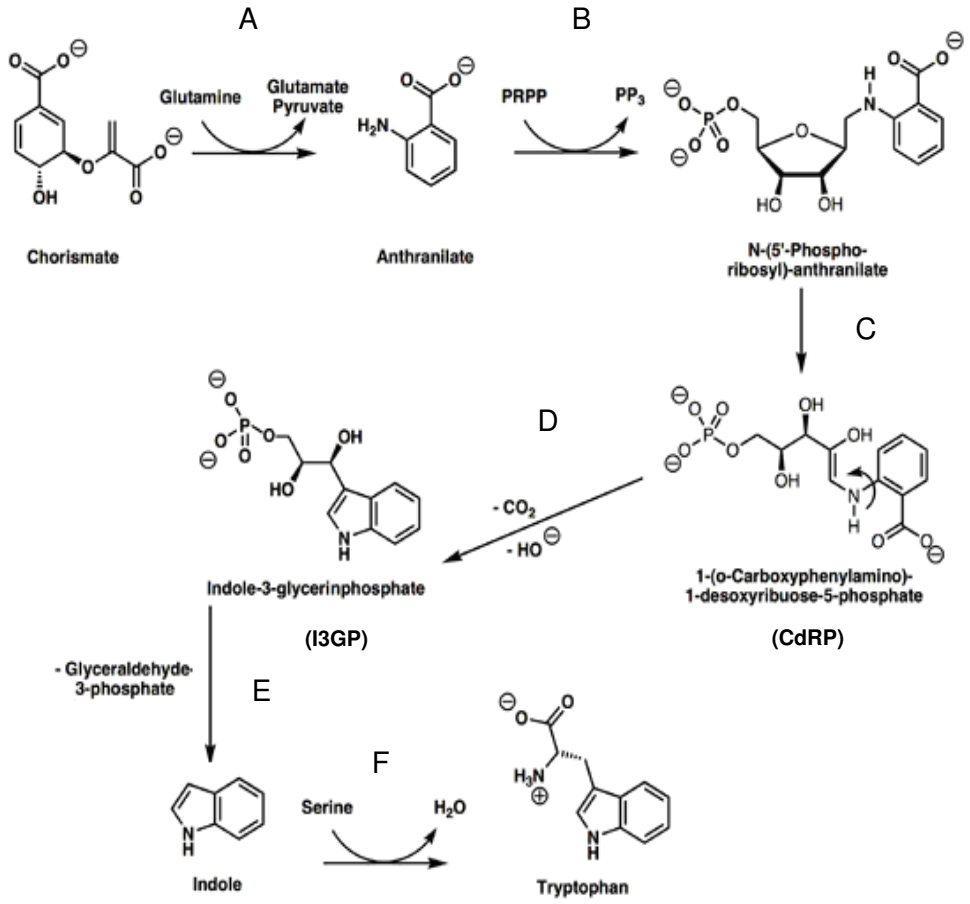


Figure 1.8 The Tryptophan Biosynthetic Pathway: (A) Anthranilate α/β (B) PR Anthranilate transferase (C) Anthranilate isomerise (D) I3GP synthase (E) TSα (F) TSβ

1.3.2 Tryptophan synthase

The final step in the tryptophan pathway is the synthesis of tryptophan from I3GP via the activity of tryptophan synthase. Since the elucidation of its three dimensional structure (Hyde *et* al, 1988) the multifaceted properties of this enzyme are now more clearly understood, with the identification of two active sites leading to the conclusion that tryptophan synthase is also an enzyme complex (Crawford, 1989). In most organisms tryptophan synthase is an $\alpha_2\beta_2$ tetramer structure arranged in an α-β-β-α linear format, initially discovered by X-ray crystallography studies of the multi-enzyme complex in *Salmonella typhirium* (Hyde *et al*, 1988). Activity involves the active site of the smaller α sub-unit cleaving I3GP to produce indole and glycerol (Figure 1.8: E). Subsequent to this cleavage the active site of the much larger β sub-unit mediates the synthesis of tryptophan from indole and serine (Figure 1.8: F). This mechanism involves a condensation reaction using pyridoxyl phosphate as an enzyme cofactor (Dunn *et al*, 1990; Lane and Kirschner, 1983).

Some unusual features of the TS $\alpha_2\beta_2$ include the observation that indole does not leave the enzyme during the conversion process of I3GP to tryptophan. Indole is confined within an intramolecular 'tunnel' connecting the first (α) and second (β) active sites (Figure 1.9) (Banik *et al*, 1995). This tunnel, consisting mainly of β-chain residues and 25 Angstrors (Å) in length, channels the hydrophobic indole between active sites (Hyde *et al*, 1988; Crawford, 1989). The enzyme complex is also reliant on the functionality of both sub-units, meaning a functional α sub-unit is required for the β subunit to function and vice-versa. These properties lead the complex to be termed a channelling enzyme (Radwanski and Last, 1995).

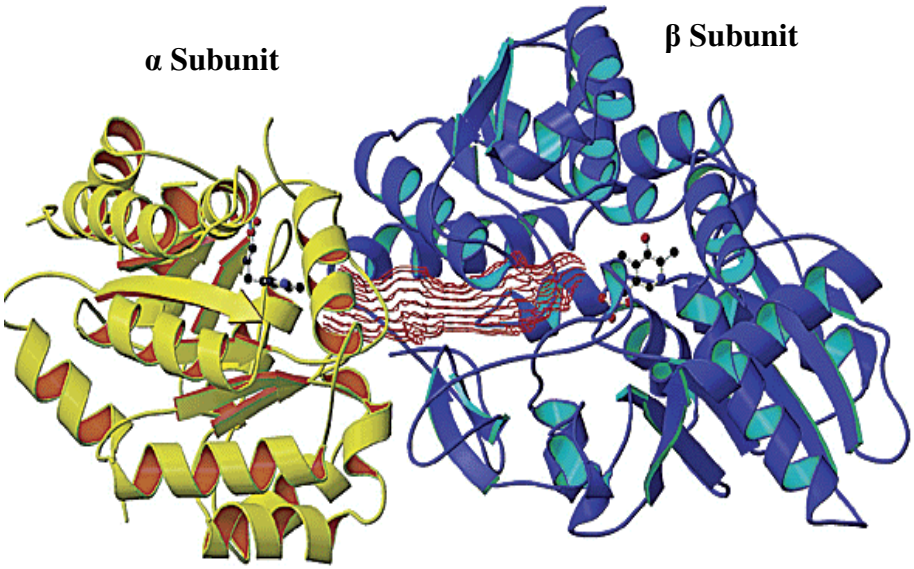


Figure 1.9 TSαβ complex with intramolecular 'tunnel' shown in red (Banik *et al*, 1995).

1.3.3 Enzyme mutants within the tryptophan biosynthetic pathway

Bacterial tryptophan biosynthesis can provide a model pathway for the study of gene structure and enzyme function due to the vast compilation of clones, mutants, mapped loci and antibodies unique to tryptophan synthesis. This is in contrast to the gathering of information regarding tryptophan biosynthesis in plants, which is in its infancy (Radwanski and Last, 1995; Hutter and Niederberger, 1986; Crawford, 1989). However, the prominent conservation between organisms has been exploited with the discovery of key structural genes and enzyme systems for each step of the pathway (Rose and Last, 1994). Characterisation of the pathway in *Arabidopsis thaliana* has been possible due to the availability of genetic mutants or antisense plants for nearly all enzyme steps in the pathway, except for Indole-3-glycerolphophate synthase (Table 1.1) (Radwanski and Last, 1995).

The step that is of most relevance to this thesis is the last step of tryptophan conversion by the TSαβ enzyme complex. Tryptophan synthase α is the best characterised of the pathway subunits within prokaryotes, however, limited knowledge exists in regards to plant TSα's. Although the information available on TSα from *A. thaliana* is scarce in comparison to microbial TSα's there has been successful identification and characterisation of an *A.thaliana* TSα variant renamed TSα1. When this newly identified cDNA clone (TSα1) was subject to experiments with *A. thaliana* cDNA expression libraries, the failure to function in the absence of the *A. thaliana* β subunit was observed indicating the inability to produce indole without a functional *A. thaliana* TSβ subunit (Radwanski *et al*, 1995). This *A. thaliana* TSα1 appeared to contain all of the amino acids required for TSα catalytic activity, conserved in TSα's from other organisms (Radwanski and Last, 1995).

Table 1.1 Genetic mutations for key enzymes within the tryptophan biosynthetic pathway in plants

Enzyme	Structural Gene[a]	Mutant Designation[b]
Anthranilate synthase		
• α Subunit	*ASA1* *ASA2* *AMT-1*	 *amt-1*
• β Subunit	*ASB1* *ASB2* *ASB3*	*trp4*
PR-anthranilate transferase	*PAT1*	*trp1*
PR-anthranilate isomerase	*PAI1* *PAI2* *PAI3*	antisense
Indole-3-glycerolphosphate synthase	*IGS1*	-
Tryptophan synthase		
• α Subunit	*TSA1*	*trp3*
• β Subunit	*TSB1* *TSB2* *TSB1* (Maize) *TSB2* (Maize)	*trp2* *orp1* *orp2*

[a] All genes are from *A. thaliana* unless otherwise stated

[b] Those genes for which mutant alleles have been identified are indicated. Antisense indicates that a mutant has been created by antisense suppression

More information is available on the TSβ subunit from plants due to the high phylogenetic conservation it possesses in comparison with other enzymes in the pathway (Miles, 1991). This conservation enabled the isolation of a 'TSβ like' gene that encoded a protein with > 95% amino acid identity to the established *A. thaliana* TSβ by screening *A. thaliana* cDNA libraries with a TSβ from yeast (Berlyn *et al*, 1989; Last *et al*, 1991). Subsequently renamed TSβ2, it was discovered that this gene was a duplicate structural gene of the already characterised TSβ and that it may play a role in the existence of a possible alternative tryptophan biosynthetic pathway (Last *et al*, 1991; Normanly *et al*, 1993).

The discovery of this duplicate *A. thaliana* TSβ sub-unit led to led to the pursuit of similar enzymes within other plants. Two cDNA clones from maize were subsequently identified on the basis of homology to the *A. thaliana* TSβ1, displaying greater than 80% similarity at the amino acid level (Wright *et al*, 1992; Last *et al*, 1991). This enabled the isolation of tryptophan synthase β deficient mutants in *A. thaliana* and maize (Radwanski and Last, 1995). The *A. thaliana trp*2-1 mutation, a point mutation (Jing *et* al, 2009) decreases the activity of TSβ to approximately 15% of wild type levels leading to diminished growth under normal growth conditions requiring tryptophan supplementation to restore normal growth. A strong indole odour emanating from seedlings is also an indication of the decreased activity of the mutated TSβ resulting from the inability to catalyse indole condensation (Last *et al*, 1991).

1.3.4 Sub-unit association

The plant TSαβ enzyme complex is believed to strongly resemble that of prokaryotes, an $\alpha_2\beta_2$ tetramer arranged in the linear form α-β-β-α (Anderson *et al*, 1991; Radwanski and Last, 1995). This plant αβ enzyme complex has also been described as a channelling enzyme, with both the α and β sub-units dependant on each other for full catalytic activity (Anderson *et* al, 1991; Berlyn *et* al, 1989; Radwanski and Last, 1995).

Catalytic activity of the α and β sub-units of the plant TSαβ are regulated by the products of two separate genes, *trp*A and *trp*B (Crawford, 1989; Berlyn *et al*, 1989). While similar catalytic functions exist in the eukaryotic fungus *Neurospora* and some yeast, each catalytic domain is controlled by a single gene (Crawford, 1989). Fusion of the *Trp*A and *Trp*B genes within these organisms occurred as evolution of the tryptophan pathway progressed, with the NH_2 terminal regulating the TSα sub-unit and the COOH terminal regulating activity of the TSβ sub-unit (Crawford, 1979; Berlyn *et al*, 1993). The discovery of this single gene controlling both α and β sub-units led to the identification of the duplicate TSβ from *A. thaliana* described previously. This was achieved by probing *A. thaliana* cDNA libraries with the COOH terminal of the *TRP5* gene from yeast, found to be homologous to the duplicate structural enzyme TSβ2, providing a possible duplicate route for tryptophan synthesis (Last *et al*, 1991).

There are several lines of evidence suggesting individual sub-units within enzyme complexes associated with tryptophan biosynthesis do have the capability to engage in multiple protein interactions (Radwanski and Last, 1995). Research has already been pursued to identify such sub-units (Hommel *et al*, 1989). The stabilisation of active protein conformations or the channelling of unstable substrates to the next active site may result from various protein interactions with the dual active sites of the tryptophan synthase complex (Radwanski and Last, 1995). Perhaps the presence of free indole, a compound known to be toxic and detrimental to plant health is one such substrate that permits atypical protein sub-unit activity.

1.4 The indigo-tryptophan paradigm

The synthesis of indigo by plants and bacteria is dependant upon the presence of indole, the key chemical intermediate crucial to a functional indigo pathway. As previously discussed, indole is also a fundamental intermediate in the tryptophan biosynthetic pathway. The last two steps of tryptophan biosynthesis are mediated by tryptophan synthase, the multifaceted αβ enzyme complex confining precursor indole to an intramolecular tunnel connecting both sub-units (Anderson *et al*, 1991). There exists in plants and bacteria enzymes that are able to catalyse the production of indole that is not dedicated to tryptophan biosynthesis. Being the key precursor to indigo synthesis, the generation of vast amounts of indole not exclusively linked to the tryptophan biosynthetic pathway is essential to further understanding the mechanisms by which indigo producing plants are able to generate indigo. This makes the identification of enzymes capable of synthesising vast amounts of indole of great interest.

1.4.1 Indole synthase

In 1995, Kramer and Koziel cloned a 'TSα-like' gene from *Zea mays*. When introduced into a *trp*A defective *E. coli* strain this cDNA sequence supported growth on minimal media in the absence of tryptophan supplementation. This indicates an ability of the IS gene to perform the *trp*A function in *E. coli* (Kramer and Koziel, 1995), thus capable of generating indole in the absence of a plant TSβ sub-unit. This cloned gene encoded an enzyme that showed significant homology with TSα sub-units from a variety of indigo producing organisms particularly in the residues dedicated to catalytic activity (Crawford, 1989), a significant fact considering that this maize gene, encoding a 'TSα like' protein was the first *trp*A gene isolated from a plant (Kramer and Koziel, 1995).

As previously described, certain intermediates within the tryptophan pathway are vital not only for tryptophan biosynthesis but they also represent a branch point for the pathways that produce secondary metabolites, some associated with plant defense mechanisms. The

hydroxamic acid pathway is one such pathway in which indole is the last intermediate common to both it and the tryptophan biosynthetic pathway (Figure 1.10) (Melanson *et al*, 1997). The hydroxamic acid pathway is instrumental in the synthesis of cyclic hydroxamic acids present as glucosides in several grasses, including maize. DIMBOA (2,4-*di*hydroxy-7-*m*ethoxy-1,4-*b*enzox*a*zin-3-one) is one cyclic hydroxamic glucoside generated by this pathway and is synthesised in response to attack from pathogenic bacteria, fungi and insects (Frey *et al*, 1997; Melanson *et al*, 1997). The most studied mutation within the DIMBOA biosynthetic pathway of maize is the *BxBx* (benzoxazineless) line, a mutation that renders the plant deficient in the synthesis of DIMBOA (Niemeyer, 1988; Melanson *et al*, 1997).

It was discovered that the maize 'TSα like' gene isolated by Kramer and Koziel, (1995) is the site of a genetic lesion responsible for the *BxBx* mutation (Melanson *et al*, 1997) evidence supported by the finding that the indole synthase maps to the same location as the *Bx*1 on the short arm of chromosome 4 (Figure 1.11). Melanson *et al*, (1997) propose that this 'TSα like' enzyme is not part of the tryptophan biosynthetic pathway but is a specific DIMBOA synthesis enzyme and that a separate TSα is expressed at a later stage of development for tryptophan biosynthesis. That this gene encoded an enzyme able to catalyse the formation of indole and was not an enzyme dedicated to tryptophan biosynthesis, this 'TSα like' enzyme was subsequently renamed Indole synthase (IS) (Melanson *et al*, 1997).

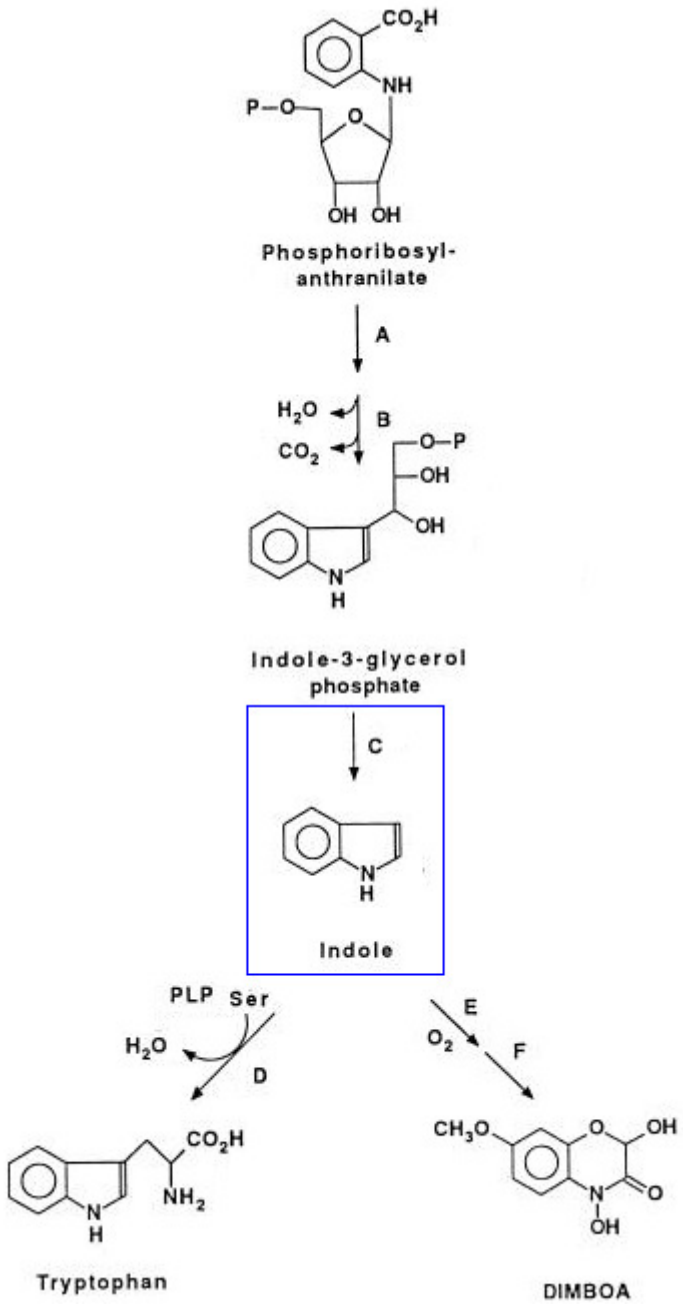


Figure 1.10 Schematic of the biosynthetic pathway of DIMBOA and tryptophan in maize. (*A*) Phosphoribosyl-anthranilate isomerase. (*B*) Indole-3-glycerol phosphate synthase. (*C*) Indole synthase (tryptophan synthase α). (*D*) Tryptophan synthase β. (*E*) A series of subsequent cytochrome P450-mediated oxygenation steps. (*F*) Methyl transferase (Melanson *et al*, 1997). Indole is the last common intermediate shared by each pathway.

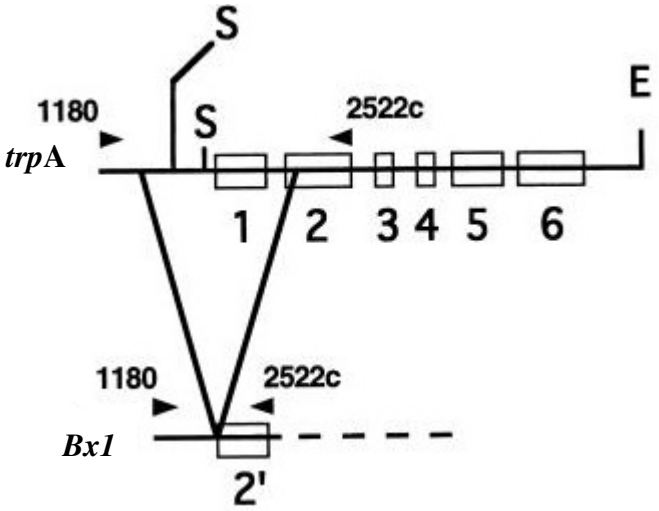


Figure 1.11 Schematic structure of the *trp*A gene and the location of the deletion on the short arm of chromosome 4 in the *BxBx* maize line. The open boxes on the line represent the gene exons (Melanson *et al*, 1997).

1.4.2 Tryptophanase

Tryptophanase (tna) is a catabolic enzyme present in gram negative bacteria active in the synthesis of indole. In *E. coli*, the presence of excess tryptophan induces the tryptophanase operon (*tna*) activating catabolite repression and anti-termination of tryptophanase transcription (Gong *et al*, 2001). The *tna* operon of *E. coli* encodes two structural genes, *tna*A, encoding tryptophanase and *tna*B that codes for tryptophan permease. Tryptophanase catalyses the formation of indole, pyruvate and ammonia from tryptophan (Evans *et al*, 1941; Watanabe and Snell, 1972). This process permits tryptophan to be utilised as a carbon and nitrogen source, however, the tryptophan-indole reaction can be catalysed in both directions by tryptophanase, favoured in the direction of tryptophan biosynthesis (Gong *et al*, 2001). As yet there has been no discovery of a plant *tna*A or *tna*A like gene encoding an enzyme that can catalyse the formation of indole from tryptophan. Due to the reliance upon the tryptophan biosynthetic pathway for more than just protein synthesis, plants may contain other unknown strategies for regulating tryptophan biosynthesis and tryptophan levels not found in microbes.

1.4.3 Rationale

Based on the information regarding the identification of an indole synthase in maize, perhaps there is a specialised indole synthase existing in other plants dedicated to the generation of large amounts of indole necessary for the synthesis of indigo. Figure 1.12 details the proposed relationship between the indigo and tryptophan pathways in plants, in contrast to the relationship observed in bacteria (Figure 1.13). Both the indigo and tryptophan biosynthetic pathways within plants and bacteria utilise indole as a key intermediate, implying competition for indole between these pathways.

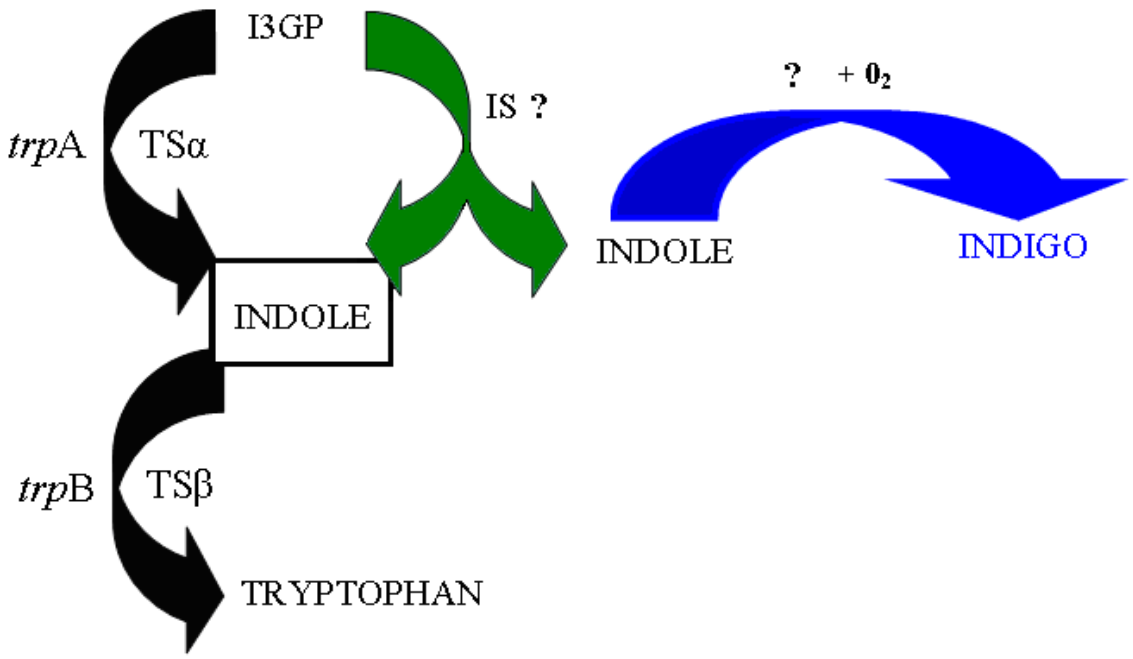


Figure 1.12 Proposed Indigo pathway in plants, branching off the tryptophan biosynthetic pathway.

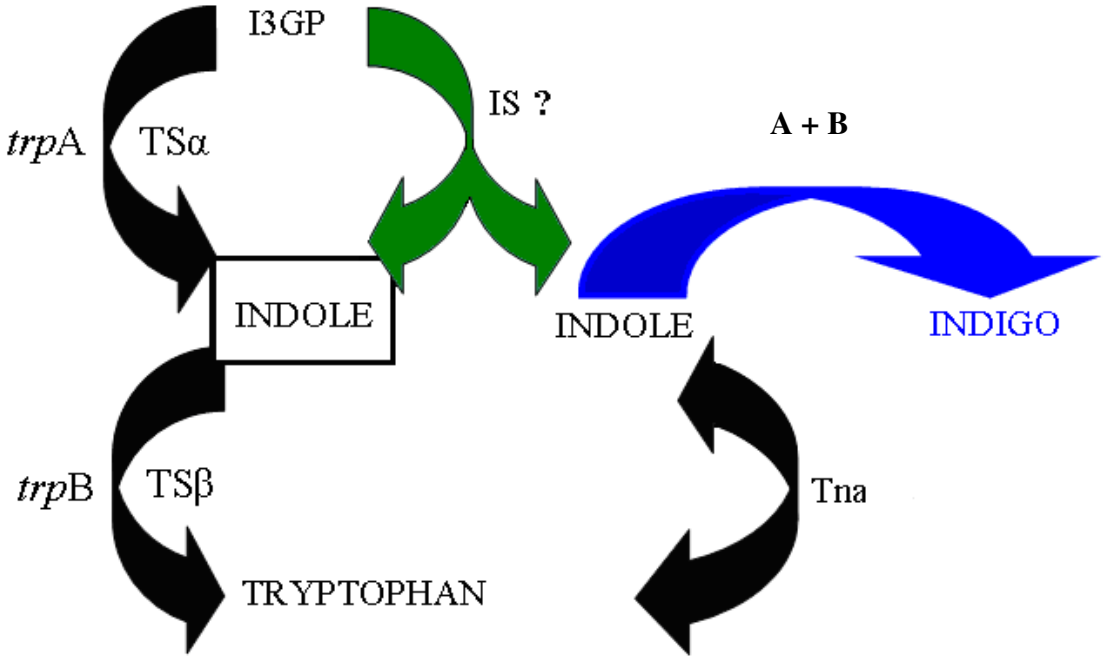


Figure 1.13 The proposed indigo pathway in *E. coli*, branching off the tryptophan biosynthetic pathway. (A) The conversion of indole to indoxyl, (B) Natural oxidation required to produce indigo. Tryptophanse carries out the trp-indole reaction in both directions however the reaction is favoured in the tryptophan to indole direction via tryptophan degradation.

There is evidence that perhaps an alternative route for tryptophan biosynthesis exists in plants (Last *et al*, 1991), yet the complete characterisation of the tryptophan synthase αβ complex in plants remains undefined and remains an area for further investigation. The research within this thesis will aim to build knowledge regarding the mechanisms by which tryptophan is produced from the sub-unit activity of the tryptophan synthase enzyme complex as well as the potential role of indole synthase in tryptophan biosynthesis. The identification of a gene in indigo producing plants that encodes an indole synthase will also be performed. With renewed interest in natural indigo dyeing, the pursuit of an indole synthase like gene in indigo producing plants is fundamental to further understanding indigo production.

1.5 Aims of the present study

Little is known about the plant tryptophan synthase enzyme complex in comparison to the microbial TSαβ enzyme complex. With the recent discovery of plant tryptophan synthase variants, unknown aspects of the tryptophan biosynthesis have been revealed increasing interest in the manipulation of the tryptophan biosynthetic pathway and the perpetually linked indigo pathway. The maize indole synthase (IS), having the ability to complement an *E. coli* trpA mutant is one such gene that has significance in possible pathway manipulation. Perhaps examination of this enzyme may aid in the elucidation of how certain plants are able to generate vast quantities of indigo. Examination of the mode of complementation of a specific *E. coli trp*A auxotroph with the maize indole synthase will be explored. Kramer and Koziel (1995), describe this enzyme as enabling growth due to co-expression with an *E.coli* TSβ sub-unit. Does this plant enzyme form a functional enzyme complex with an *E. coli* β sub-unit or does the IS generate indole for trpyptophanase based tryptophan biosynthesis? The indigo producing capabilities of indole synthase will also be examined as a precursor to a potential novel indole synthase assay system.

The specific aims of chapter three include:

- Examination of the mode of complementation by the indole synthase in an *E. coli trp*A mutant. Specifically how does activity of the maize IS ultimately generate tryptophan?
- The construction of an *E. coli* double mutant deficient in both TSα and tryptophanase activity for the purpose of analysing indigo accumulation. Can the IS generate indole for both indigo and tryptophan biosynthesis.
- Construction of a novel vector harbouring the indole oxygenase gene from *Rhodococcus* vital to indigo synthesis in bacteria.
- The generation of an indole synthase/indole oxygenase sequential enzyme complex for potential indigo accumulation in bacteria.

The identification of genes in indigo producing plants *I. tinctoria* and *P. tinctorium* that encode enzymes with the same enzymatic capabilities as the maize IS is the basis for chapter four. The identification of a gene or genes encoding an enzyme capable of catalysing 'free' indole production not exclusive to the tryptophan biosynthetic pathway in *P. tinctorium* will also be pursued. This will be performed by screening a *P. tinctorium* cDNA library for a gene homologous to the maize indole synthase as well as performing a functional assay to determine if there are any genes capable of emulating the maize IS within the cDNA library.

The specific aims of chapter four include:

- The isolation of *trp*A genes from *I. tinctoria* and *P. tinctorium* encoding TSα sub-units
- The potential complementation of an *E. coli trp*A auxotroph with the cloned *trp*A genes to ascertain the enzymatic capabilities of the various plant TSα sub-units
- The construction of a *P. tinctorium* cDNA library
- The pursuit of a gene encoding an indole synthase enzyme by screening the library via hybridisation with the maize IS gene and via functional analysis

This thesis aims to dissect the indigo pathway with aspirations to identify the mechanisms by which indigo producing plants are able to generate such vast amounts of indigo. Analysis of the tryptophan and indigo pathways will provide insight in to the renewed interest in natural indigo dyeing and possible pathway manipulations with potential for indigo accumulation in plant fibres via tissue culture experiments and various other transgenic techniques.

2 Chapter Two: General materials and methods

2.1 General procedures

All solutions were prepared using Millipore MilliQ grade deionised H_2O unless otherwise stated. All media and glassware was autoclaved at 121°C for a period of 20 mins unless otherwise stated. Solutions ranging from 0.2 µL to 5 mL were dispensed using Eppendorf Research micropipettes, for all volumes above 5mL measuring cylinders were utilised. All used glassware was washed with Pyroneg detergent and rinsed with tap water and Millipore MilliQ grade deionised H_2O.

2.2 Media and antibiotics

All media and antibiotic solutions were weighed out and prepared using an AFP-36L ADAM Balance machine from Crown Scientific.

2.2.1 Antibiotic stock solutions

Ampicillin: Amp was dissolved in sterile water to a concentration of 100 mg/mL stock solution. One millilitre aliquots were dispensed and stored at -20°C. The working concentration of Amp was 100 µg/mL.

Chloramphenicol: Cm was dissolved in MetOH to a concentration of 30 mg/mL stock solution. One millilitre aliquots were dispensed and stored at -20°C. The working concentration of Cm was 30 µg/mL.

Kanamycin: Kan was dissolved in deionised H_2O to a concentration of 50 mg/mL stock solution. One millilitre aliquots were dispensed and stored at -20°C. The working concentration of Kan was 50 µg/mL.

Tetracycline: Tet was dissolved in a 1:1 volume of deionised H_2O:EtOH to a concentration of 12.5 mg/mL stock solution. One millilitre aliquots were dispensed and stored at -20°C. The working concentration of Tet was 12.5 µg/mL

2.2.2 LB agar

Tryptone (1% w/v), Yeast Extract (0.5% w/v), NaCl (0.5% w/v) and Bacteriological Agar (1% w/v) were dissolved in deionised H_2O and autoclaved at standard conditions outlined in General Procedures (Section 2.1).

2.2.3 LB broth

Tryptone (1% w/v), Yeast Extract (0.5% w/v), and NaCl (0.5% w/v) was dissolved in deionised H_2O and autoclaved at standard conditions outlined in General Procedures (Section 2.1).

2.2.4 M9 minimal media (+ CAA) agar

Bacteriological Agar (1.5% w/v) was dissolved in deionised H_2O and autoclaved at standard conditions outlined in General Procedures (Section 2.1). Once sufficiently cooled, 20% Glucose (2% v/v), 10 x M9 Salts (10% v/v), 1M $MgSO_4$ (0.2% v/v), 1M $CaCl_2$ (0.01% v/v) and 5% CAA (10% v/v) were aseptically supplemented. (NB: The 10 x M9 Salts and 5% CAA solutions were previously autoclaved, while the 20% Glucose, 1M $MgSO_4$ and 1M $CaCl_2$ were previously filter sterilised).

2.2.5 M9 minimal media (+ CAA) broth

20% Glucose (2% v/v), 10 x M9 Salts (10% v/v), 1M $MgSO_4$ (0.2% v/v), 1M $CaCl_2$ (0.01% v/v) and 5% CAA (10% v/v) were added to deionised H_2O aseptically.

2.3 Bacterial strains and plasmids

Table 2.1 Bacterial strains used in this study

Bacterial Strain	Genotype
Eschericia coli	
BW25141[1]	Δ(*araD-araB*)567,Δ*lacZ*4787(::rrnB-3),*lacIp*-4000(lacI^{Q}),Δ(*phoB-phoR*)580,λ^{-},*gal*U95,Δ*uid*A3::pir+
BW25113[1]	Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), lambda^{-}, rph-1, Δ(rhaD-rhaB)568, hsdR514
DH5α[2]	*supE ΔlacU169 (ϕ80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*
W3110trpA33[3]	F-, trpA33, rpH1, IN(rrnD-rrnE)1
XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacI*qZΔ*M15* Tn*10* (Tetr)].
BM25.8	F′ [*traD36 lacI*^{q} *lacZ*ΔM15 *proAB*] *supE thi* Δ(*lac-proAB*) [P1] Cm^{r} r^{-} m^{-} λimm^{434} Kan^{r}
Salmonella typhimurium	
LT2-9121[4]	*leu hsdL trpD2 rpsL120 ilv452 metE551 metA22 hsdA hsdB*

[1] Reference/Source: (Datsenko and Wanner, 2000)

[2] Reference/Source: RMIT Plant Biotechnology Dept.

[3] Reference/Source: CGSC, Yale University, Connecticut, U.S.A.

[4] Reference/Source: RMIT Biotechnology Dept

Table 2.2 Plasmids used in this study

Vector	Description
pUC18[1]	2686bp, pMB1 Ori, LacZ(α), Amp^R
pUC120[1]	3168bp, derivative of *pUC118* that possesses a unique *Nco*1 restriction site at the LacZ(α) start codon
pISP1[2]	3458bp, *pUC18* backbone with *Z. mays* IS insert, Amp^R
pISP2[2]	3518bp, *pUC18* backbone with *I. tinctoria* TSα insert, Amp^R
pISP3[2]	3518bp, *pUC18* backbone with *P. tinctorium* TSα insert, Amp^R
pISP4[2]	3518bp, *pUC18* backbone with *A. thaliana* TSα1 insert, Amp^R
pISP5[2]	3443bp, *pUC18* backbone with *A. thaliana* TSα2 insert, Amp^R
pNAV79[1]	4294bp, *pUC120* backbone with *Rhodococcus* INO from *p5129*, Amp^R
pACYC184[1]	4245bp, p15A Ori, Tet^R, Cm^R
pNAV110[2]	3196bp, *pACYC184* with Cm^R excluded via digestion with *Xmn*1/*Sca*1 and re-ligation, Tet^R

pNAV111[2]	3841bp, *pNAV110* backbone with LacZ(α) insert from *pUC120*, Tet^R
pNAV112[2]	4966bp, *pNAV111* backbone with INO gene insert interrupting the LacZ(α), Tet^R
pKD13[3]	3.4kb plasmid, oriR6Kγ, $tL3_{\lambda}$(Ter), bla(ApR), kan, rgn(Ter)
pK46[3]	6.3kb plasmid, repA101(ts), araBp-gam-bet-exo, bla(ApR), oriR101

[1] Source: RMIT Plant Biotechnology Dept.

[2] Source: Plasmid constructed for this study

[3] Source: RMIT Biotechnology Dept.

2.3.1 Storage of bacterial strains

Long Term All bacterial strains were stored at -80°C in 10% sterile glycerol (v/v)

Short Term All bacterial strains were stored on LB agar plates at 4°C. The bacteria were re-streaked every 4-6 weeks from single colonies.

2.3.2 Bacterial culture conditions

All *E. coli* and *Salmonella* strains were grown in liquid broth at 37° C for 16 h on an orbital shaker at 200 rpm.

2.4 General DNA and RNA techniques

2.4.1 Extraction of DNA

2.4.1.1 Small Scale plasmid DNA extraction using PROMEGA SV miniprep kit

Small scale plasmid DNA extraction was achieved as per the manufacturer's instructions (Promega, U.S.A). The purification procedure was performed according to the manufacturer's instructions. A 5 mL culture of LB media and the appropriate antibiotic was collected by centrifugation at 6, 000 rpm for 10 mins at 4°C. The pellet was resuspended in 250 µL of resuspension buffer to which 250 µL of lysis buffer was added, the microcentrifuge tube was then inverted several times to mix solutions and incubated for 5 min sat room temperature to lyse the cells. Alkaline protease was added (10 µL) to eliminate unwanted precipitated protein followed by addition of 350 µL neutralisation buffer and a subsequent 10 min incubation step once again at room temperature. The supernatant was collected after centrifugation at 14,000 rpm for 10min, decanted into a Promega spin column and spun for 30 seconds in the centrifuge after which the flow through was discarded. The bound DNA was washed with 250 µL of wash buffer and spun for 1 min in the centrifuge after which flow through was once again discarded. Residual buffer was eliminated by a secondary 30 second spin in the centrifuge. The Promega spin column was placed in a new microcentrifuge tube and the DNA was eluted with 50 µL of 10 mM Tris, pH 8.0.

2.4.1.2 DNA quantification

DNA was quantified by determining the concentration of eluted DNA by reading absorbance levels at 260 nm. Alternatively the concentration of DNA was determined by viewing aliquots on a 1% TBE or TAE (w/v) agarose gel after electrophoresis.

2.4.2 Extraction of RNA

2.4.2.1 Small scale RNA extraction

RNA was extracted from leaf tissue using the RNeasy® Plant Mini Kit as per the manufacturer's instructions (Qiagen®, Australia). No more than 100 mg was ground in liquid nitrogen and decanted into a 1.5 mL microcentrifuge tube. Buffer RLT (450µL) was added to the ground plant tissue and vortexed to homogenise the lysate. The lysate was then transferred to a Qiashredder spin column and centrifuged at 14,000 rpm for 2 min. The collected supernatant was combined with 0.5 volumes of 95% EtOH (v/v) and mixed thoroughly. The sample was then transferred to an RNeasy spin column and centrifuged for 15 seconds at 10,000 rpm with any flow through discarded. Buffer RW1 (700 µL) was added to the column which was once again spun for 15 seconds with the flow through discarded, 500 µL of buffer RPE was then added to the column to wash the RNA, any residual buffer was eliminated by an additional 1 min centrifugation step. The RNeasy column was then placed in a new 1.5mL microcentrifuge tube and RNA was eluted with 50 µL of 10 mM Tris, pH 8.0. Resultant RNA was cleaned up via DNase digestion following the protocol provided in the RNeasy® Plant Mini Kit (Qiagen®, Australia).

2.4.2.2 RNA quantification

RNA was quantified by determining the concentration of eluted RNA by reading absorbance levels at 260 nm. Alternatively the concentration of RNA was determined by viewing aliquots on a 1% TAE (w/v) agarose gel or a 1% Formaldehyde gel (v/v) after electrophoresis.

2.4.3 cDNA synthesis

Oligo DT (0.5 µg) and 2 µg of RNA was added to deionised H_2O for a resultant 15 µL volume solution. The solution was incubated at 70°C for 5 min sand immediately chilled on ice. After a short centrifuge spin the RNA/Oligo DT solution was supplemented with 5 µL

of 5 x MMLV Buffer, 1.25 µL 10mM dATP, 1.25 µL 10mM dCTP, 1.25 µL 10mM dGTP, 1.25 µL 10mM dTTP, 25 units of RNasin® RNase Inhibitor (30 U/µL), 200 units of MMLV Reverse Transcriptase (200 U/µL) and made up to a total volume of 25 µL with deionised H_2O. The solution was then mixed and incubated for 1 h at 37°C and stored at -20°C in preparation for amplification via PCR. This was the method of cDNA synthesis for all PCR gene templates unless otherwise specified. (NB: All reagents were provided by Promega, Australia).

2.4.4 Ethanol precipitation

This method was used to concentrate DNA samples by decreasing the total volume. One-fifth of the sample volume of 10 M ammonium acetate and 2.5 volumes of ice-cold absolute ethanol were added to the DNA preparation. The sample was stored at -70°C for 30 min. The DNA was collected by centrifugation in an Eppendorf microcentrifuge at 10,000rpm for 15 min. The pellet was washed twice with 70% (v/v) EtOH and was then air-dried. The pellet was resuspended in 10 mM Tris, pH 8.0.

2.4.5 Gel electrophoresis

Gel electrophoresis was performed using 1% gels (w/v) to separate and analyse DNA. Electrophoresis was performed at 100 volts generated by a BIORAD Power Pack™.

2.4.5.1 DNA gels

1% TAE or TBE (w/v) gels were supplemented with ethidium bromide (0.5 µg/mL) and used to analyse DNA by gel electrophoresis.

2.4.6 DNA purification

DNA from agarose was purified via the QIAGEN® Gel Extraction Kit using the manufacturer's instructions. DNA was viewed under ultraviolet light and specific bands were excised and resuspended in buffer QG then heated to 50°C for 10 mins to dissolve the agarose. The solution was then applied to a QIAGEN® spin column to and subject to centrifugation at 14,000 rpm for 1 min to bind the DNA to the column. The column was washed with buffer PE via another centrifuge step after which the wash buffer was discarded and a secondary centrifuge step was pursued to eliminate any residual buffer PE. The QIAGEN® spin column was then placed in a new 1.5mL microcentrifuge tube and the purified DNA was eluted with 50 µL of 10 mM Tris, pH 8.0. All PCR products and pDNA extracts were purified.

2.4.7 Polymerase chain reaction (PCR)

2.4.7.1 Primer design

Primers were designed with a GC content of between 40-60% to ensure stable binding to the template DNA with a Tm°C in the range of 50-80°C. Primers were designed from sequences obtained from the NCBI website and certain restriction enzyme sites were integrated into primers to aid in cloning procedures. The primers used in this investigation are listed in Table 2.3.

2.4.7.2 General PCR procedure

The PCR amplifications were carried out in 25 µL reactions containing: 12.5 µL 2 X GoTaq® Green Master Mix, 2 µL template DNA (50-100 ng cDNA or 0.1-1 ng plasmid DNA), 2 µL 10 µM forward primer, 2 µL 10 µM reverse primer and 6.5 µL of nuclease free H_20. The PCR cycling conditions for all reactions are detailed in Table 2.4. Annealing temperature was dependant on the primers used in each specific PCR reaction.

Table 2.3 Primers used in this study

Primer	Sequence 5' → 3'	Target	Tm°C
M13(-20)F	gtaaaaccacggccagt	Universal Primer	47
M13R	caggaaacagctatgac	Universal primer	44
INO(F)	acaccaagcttaaccatggacatcacccgc	INO gene	64
INO(R)	gatttccatggatcctcagaggagcgggtc	INO gene	66
AtTSA1(F)	acatggatccttctctctccacctcttcta	*A.thaliana* TSα1	66
AtTSA1(R)	cggcaagcttgtcaaagaagagcagattta	*A. thaliana* TSα1	66
AtTSA(2)F	acatggatccttcacaaggcaaagtggctc	*A. thaliana* TSα2	66
AtTSA(2)R	cggcaagcttgtcaagagacaagagcagac	*A. thaliana* TSα2	66
ItTSA(F)	acatggatcctgccctctccacctcttctc	*I. tinctoria* TSα	67
ItTSA(R)	atcgaagctttatcaaagaagagcagattt	*I. tinctoria* TSα	59
ZmIS(F)	acatggatcctgcgctcatggccaagggca	*Z. mays* IS	68
ZmIS(R)	tgacaagctttttcatggcagcgttctt	*Z. mays* IS	59
SmartIV	aagcagtggtatcaacgcagagtggccattacggccggg	Serves as template at 5' end of mRNA	70
CDSIII/3' PCR Primer	attctagaggccgaggcggccgacatg-d(T)$_3$ N$_{-1}$N-3'	Primes first strand synthesis reaction	70
β-glucF	acatggatccataagcattccgcttaagcggac	*P. tinctorium* β-glucosidase	64

β-glucR	cggcaagcttgctgc gattatacccttttgat c	*P. tinctorium* β-glucosidase	64
5' Seq	ctccgagctctggac gagc	*pTriplEx2* vector	57
3' Seq	taatacgactcactat aggg	*pTriplEx2* vector	53
LacZFBlunt	ccaatacgcaaacc gcctctccccg	LacZ(α) from *pUC120*	64
LacZRBlunt	aacgtggcgagaaa ggaagggaaga	LacZ(α) from pUC120	59
P1(*pKD13*)	atggaaaactttaaac atctccctgaaccgtt ccgcattcgtgttatt gagtgtaggctgga gctgcttc	*pKD13* (template for arabinose induced recombination)	74
P2(*pKD13*)	ttaaacttctttcagttt tgcggtgaagtgacg caatacttttggttcgt attccggggatccgt cgacc	*pKD13* (template for arabinose induced recombination)	74
Kt	cggccacagtcgat gaatcc	Kan resistance gene	56
K2	cggtgccctgaatga actgc	Kan resistance gene	56

Table 2.4 PCR cycling conditions

Number of Cycles	Temp °C	Duration
1 (Denaturation)	94	3 min
30 (Denaturation)	94	30 sec
30 (Annealing)	50-70	1 min
30 (Extension)	72	30 sec/kb
1 (Extension)	72	5 min

2.4.8 DNA sequence analysis

DNA sequence analysis was performed at AgGenomics (Victorian AgriBiosciences Centre). All reactions were carried out at the sequencing facility. Sequencing reaction constituents are listed in Table 2.5.

Table 2.5 Sequencing reaction set-up

Component	Volume (µL)
Template DNA (300-600ng)	x
Primer (3.2µM)	1.0
ABI BigDye terminator v3.1	1.0
5 X Sequencing Buffer	3.5
MilliQ Water	y
Total Volume	**20**

Sequencing cycling conditions are listed below:

- 96°C 1 min
- 96°C 10 sec
- 50°C 5 sec
- 60°C 4 min

} 25 cycles

2.5 Recombinant DNA techniques

2.5.1 Cloning of PCR products

The PCR amplified genes were first digested with the restriction enzymes integrated into the PCR product via amplification with primers inclusive of specific restriction enzyme sites. The concentration of PCR products was determined via DNA quantification (Section 2.4.1.2). Ligation of PCR products into similarly digested cloning vectors was performed using T4 DNA ligase. Subsequent ligation solutions were used to transform electrocompetent *E. coli* cells.

2.5.2 Restriction enzyme digestion

Table 2.6 describes all restriction enzymes used in this study and details their recognition sequences. All digestions were performed at 37°C over a duration ranging from 2 h to overnight. Constituents of the reactions were as follows:

- 50-100 ng DNA
- 5 µL 10 X Restriction Enzyme Buffer
- 1 U of Restriction Enzyme

} Up to 50 µL with nuclease free water

Table 2.6 Restriction enzymes used in this study and their recognition sequences

Restriction Enzyme	Recognition Sequence
*Bam*H1	5' G/GATCC 3' 3' CCTAG/G 5'
*Hind*III	5' A/AGCTT 3' 3' TTCGA/A 5'
*Sfi*1	5' GGCCNNNN/NGGCC 3' 3' CCGGN/NNNNCCGG 5'
*Dpn*1	5' GA/TC 3' 3' CT/AG 5'
*Xmn*1	5' GAATT/AATTC 3' 3' CTTAA/TTAAG 5'
*Sca*1	5' AGT/ACT 3' 3' TCA/TGA 5'
*Nco*1	5' C/CATGG 3' 3' GGTAC/C 5'
*Pst*1	5' CTGCA/G 3' 3' G/ACGTC 5'

2.5.3 Ligation of DNA

Ligation reactions were performed after DNA was purified via gel extraction using the QIAGEN® Gel Extraction Kit. Ligation reactions were made up to 20 µL and carried out at room temperature for a minimum of 4 h or overnight at 4°C. Listed below are the constituents of all ligation reactions:

- Plasmid DNA – 100 ng
- Insert DNA – x µL
- 10 X Ligation Buffer – 2 µL
- T4 DNA Ligase – 1 U

} Up to 20 µL with nuclease free water

The total amount of insert DNA was calculated using the following formula:

$$\frac{\text{Vector (ng) X Size of Insert (kb)}}{\text{Size of Vector (kb)}} \text{ X } \frac{3}{1} = \text{Insert (ng)}$$

2.5.4 Transformation

All transformations were performed using the BIORAD™ Electroporator.

2.5.4.1 Generation of electrocompetent cells

Bacterial cells were grown in 10 mL of LB broth supplemented with appropriate antibiotic on an orbital shaker at 37°C for 16 h. Five millilitres of this culture was used to inoculate 200 mL of pre-warmed LB broth in a flask and shaken vigorously at 37°C until approximately mid-log phase, when an absorbance at 600 nm of 0.4-0.5, was reached. The cells were chilled on ice for 30 min and collected by centrifugation at 10,000 rpm for 15 min at 4°C. The supernatant was decanted and discarded before the cells were washed in 200 mL of ice-cold H_2O and collected by centrifugation at 10,000rpm for 15 min at 4°C. This step was repeated with 100 mL of ice-cold H_2O. The cell pellet was resuspended in 10 mL of ice-cold 10% glycerol (v/v) and centrifuged as before. The cells were resuspended in a final volume of 1 mL of ice-cold 10% glycerol and divided into 50 µL aliquots and stored at -80°C in preparation for transformation via electroporation.

2.5.4.2 Electroporation

Frozen electro-competent cells were thawed on ice for 15 mins, mixed with DNA (0.5 µL pDNA, 2.0 µL ligation reaction) and transferred to an ice-cold electrocuvette. For electrocompetent *E. coli* cells the pulse settings used to deliver DNA into the cells were 2.5kV, 25mF and 200Ω. After the electroporation, the transformed bacteria was added to 1 mL of liquid LB media and cultures were either plated onto LB media immediately or

incubated at 37°C for 1 h depending on the specific selection antibiotic. One hundred microliters of transformed culture was plated out onto LB agar containing the appropriate antibiotic, while the remaining transformation mixture was concentrated by centrifugation at 10,000 rpm for 5 min and resuspended in 100 µL of LB media. This was also plated onto LB containing the appropriate antibiotic.

2.5.4.3 Selection

To select for successful transformants, the media was supplemented with the corresponding antibiotic to that of the cloning vector. Blue/white screening was performed where specified via addition of IPTG and and X-Gal to a final concentration of 1 mM and 20 µg/mL to the autoclaved media to identify colonies with inserts.

3 Chapter Three: Examination of the role of a maize indole synthase in an *E. coli trp*A mutant

3.1 Introduction

In indigo producing plants and microbes, indole is a fundamental intermediate in both the tryptophan and indigo biosynthetic pathways. As discussed previously, an IS gene (*Bx1*) isolated from maize has the ability to catalyse the synthesis of indole that is not utilised for tryptophan biosynthesis. This was discovered via the complementation of a strain of *E. coli* deficient in TSα activity due to a mutation in the *trp*A gene with the maize IS gene. This indicated that the maize IS could perform the *trp*A function in *E. coli* (Kramer and Koziel, 1995). Therefore the question is raised, does this enzyme produce indole that is channelled directly through an *E. coli* TSβ sub-unit for tryptophan biosynthesis, or does it produce indole for tryptophan biosynthesis via another mechanism? Although Kramer and Koziel, (1995), describe the ability of the maize IS to complement a *trp*A *E. coli* mutant, they do not report the specific mechanism by which tryptophan is ultimately produced.

In gram negative bacteria, specifically *E. coli*, there are enzymatic reactions that utilise indole as a substrate. The degradation of tryptophan in *E. coli* is regulated by transcription of the tryptophanase operon (tna) (Gish and Yanofsky, 1993) and subsequent synthesis of tryptophanase (Stewart *et al*, 1986), encoded by *tna*A, a gene encoding the catabolic enzyme tryptophanase that utilises tryptophan as a substrate for production of indole, pyruvate and ammonia (Bilezikian *et al*, 1967; Watanabe and Snell, 1972; Gong *et al*, 2001). Tryptophanase is able to carry out the tryptophan-indole reaction in both directions (Watanabe and Snell, 1972; Deeley and Yanofsky, 1981). This represents a possible mode of tryptophan biosynthesis that Kramer and Koziel failed to consider.

In this chapter, the construction of a bacterial assay system for both tryptophan and indigo biosynthesis, mediated by action of the maize indole synthase and a subsequent indole oxygenase activity will be presented. This sequential enzyme mechanism should provide a pathway to indigo production that is independent of a pathway involving tryptophan degradation catalysed by tryptophanase. However, while the tryptophanase gene remains functional in *E. coli* indigo accumulation may be attributed to *tna*A encoded tryptophanase activity and not the indole synthase/indole oxygenase sequential enzyme system under examination. It is therefore necessary to 'knockout' the *tna*A gene and consequently eliminate any tryptophanase activity.

In this study, inactivation of the chromosomal *tna*A gene will be performed via the use of a Red recombinase plasmid system that permits mutation through gene targeting (Datsenko and Wanner, 2000). Red recombination relies on homologous recombination and adapted technology from systems found in bacteriophages encoding their own homologous recombination systems (Smith, 1988; Datsenko and Wanner, 2000). In this instance the system will aim to replace a section of the *tna*A gene with a kanamycin resistance cassette thereby allowing the selection of recombinants via the ability to grow in the presence of kanamycin. The expression plasmid *pKD46* (Datsenko and Wanner, 2000), a Red recombinase plasmid used to 'knockout' *E. coli* chromosomal genes, is a low copy helper plasmid containing three genes (γ, β and *exo*) whose respective products, Gam, Bet and Exo drive enhanced recombination of linear DNA to disrupt target genes (Winans *et al*, 1985; Datsenko and Wanner, 2000). Targeted mutagenesis is performed utilising specific PCR products generated from a template plasmid *pKD13*. Amplification of a specific region of *pKD13* results in the production of a linear DNA fragment containing a kanamycin resistance cassette flanked by Flippase (FLP) recombination enzyme recognition target (FLT) sites. Subsequent integration of this resistance cassette into a target gene is mediated by the FLT sites that promote a recombination event induced by arabinose (Datsenko and Wanner, 2000;

Baba *et al*, 2006). Previously, most bacterial species were not easily transformable with linear DNA fragments due to the presence of internal exonucleases that degrade linear DNA (Lorenz and Wackernagel, 1994). However, the Red recombinase system enables easy transformation with PCR products due to the strains origins having been derived from *E. coli* mutants lacking exonuclease activity but remaining proficient in recombinase activity. This makes the Red recombinase system an ideal tool for targeted mutagenesis studies.

The construction of a double mutant will be performed to assay the function of indole synthase and indole oxygenase in *E. coli* cultures. The mutant bacterial strain *E. coli* W3110*trp*A33 is deficient in TSα activity and as a result lacks the ability to synthesise tryptophan. This *trp*A auxotroph is an ideal candidate for targeted mutagenesis of the *tna*A gene using the Red recombinase system thereby creating a double mutant deficient in both TSα and tryptophanse activity. The specific aim of the work reported in this chapter was to generate indigo using glucose as the sole carbon source within the *E. coli* W3110*trp*A33*tna*A double mutant strain. This will be achieved via the introduction of an indole synthase gene that will be co-expressed with an exogenous indole oxygenase gene. Whether this strain can be used to assess the indigo producing capabilities of indole synthase will be evaluated. The premise that precursor indole can be utilised for both tryptophan and indigo synthesis will also be explored. This will then provide the basis of a functional screen for putative indole synthase genes from other plant species known to produce large quantities of indigo.

3.2 Materials and Methods

3.2.1 Construction of a plasmid harbouring the indole oxygenase gene

All constructs were verified for the successful integration of inserts via PCR (Section 2.4.7) or restriction enzyme digestion (Section 2.5.2). All products were viewed on DNA gels after electrophoresis (Section 2.4.5) and ethanol precipitation was performed where necessary as described in Section 2.4.4.

3.2.1.1 pNAV110 construct

Construct *pNAV110* was produced via excision of the chloramphenicol resistance gene from cloning vector *pACYC184* through restriction enzyme digestion with *Xmn* 1 and *Sca*1 (Section 2.5.2). Religation of the digested plasmid was performed as described in section 2.5.3. Construct *pNAV110* was then used to transform electrocompetent *E. coli* DH5α cells (Section 2.5.4) and resultant cultures were plated onto LB media supplemented with tetracycline. Plasmid DNA was then extracted from resulting colonies via small scale DNA extraction described in Section 2.4.1.1.

3.2.1.2 pNAV111 construct

Construct *pNAV111* was produced via insertion of the LacZ(α) fragment from the cloning vector *pUC120*. The LacZ(α) fragment was amplified using *pUC120* as a PCR template (Section 2.4.7) and primers detailed in Table 2.3. Blunt ended purified PCR products were inserted into *pNAV110* digested with *Xmn*1 and *Sca*1 (Section 2.5.2), via blunt end ligation (Section 2.5.3). Construct *pNAV111* was then used to transform electrocompetent *E. coli* DH5α cells (Section 2.5.4) and resultant cultures were plated onto LB media supplemented with tetracycline, IPTG and X-gal. Blue colonies were selected and plasmid DNA was then extracted via small scale DNA extraction described in Section 2.4.1.1.

3.2.1.3 pNAV112 construct

Construct *pNAV112* was produced via insertion of the *Rhodococcus* indole oxygenase gene (GenBank: accession number M30195) from vector *pUC79* (provided by Prof. David Stalker of the RMIT Plant Biotechnology Dept.) into *pNAV111*, interrupting the LacZ(α) and making blue/white screening possible. The indole oxygenase was excised from *pUC79* via restriction enzyme digestion with *Nco*1 and *Pst*1 (Section 2.5.2) and ligated into the similaraly digested *pNAV111* as described in Section 2.5.3. Construct *pNAV112* was then used to transform electrocompetent *E. coli* DH5α cells (Section 2.5.4) and resultant cultures were plated onto LB media supplemented with tetracycline, IPTG and X-gal. White colonies were selected and plasmid DNA was then extracted via small scale DNA extraction described in Section 2.4.1.1.

3.2.2 The indole synthase/indole oxygenase sequential enzyme system

Co-existance of *pNAV112* (harbouring the INO gene) and *pISP1* (harbouring the IS gene) in *E. coli* was pursued to generate an indole synthase/indole oxygenase sequential enzyme system to permit indigo accumulation in bacteria.

3.2.3 Red disruption system

Red recombinase mediated mutagenesis was required to knockout the *tna*A gene from *E. coli* W3110*trp*A33.

3.2.3.1 Extraction of pKD46 from E. coli BW25113

Red helper plasmid *pKD46* was extracted from *E. coli* BW25113 and purified using the small scale plasmid DNA extraction method described in Section 2.4.1.1. Cultures were grown at 30°C overnight due to the temperature sensitive replicon of *pKD46* which allows replication of the plasmid.

3.2.3.2 Extraction of pKD13 from E. coli BW25141

Plasmid *pKD13* was extracted from *E. coli* BW25141 and purified using the small scale plasmid DNA extraction method described in Section 2.4.1.1. Cultures were grown at 30°C overnight due to the temperature sensitive replicon of *pKD46* which allows replication of the plasmid.

3.2.3.3 Transformation of E. coli W3110trpA33 with pKD46

The transformation protocol of *E. coli* W3110*trp*A33 is described in Section 2.5.4.1 and 2.5.4.2. Selection was done on LB media supplemented with ampicillin. Transformants were examined for the presence of *pKD46* via restriction enzyme digestion and subsequent viewing on a DNA gel after electrophoresis.

3.2.4 Amplification of kanamycin^{R} cassette from *pKD13* template

3.2.4.1 Primer design

Primers were designed as described in Section 2.4.7.1. Primers of 70 nucleotides in length were generated for both forward and reverse orientation. Each forward and reverse primer incorporated 50 nucleotides homologous to the *tna*A gene in *E. coli* included to the 5' ends of each primer and downstream of 20 nucleotides complementary to the *pKD13* template (Datsenko and Wanner, 2000).

3.2.4.2 Cycling conditions

PCR reactions were performed in a total volume of 50 µL for this specific task. Quantities of reagents described in Section 2.4.7.2 were doubled.

Table 3.1 Specific PCR cycling conditions for amplification of kanamycin^R cassette

Number of Cycles	Temp °C	Duration
1 (Denaturation)	94	5 min
35 (Denaturation)	94	1 min
35 (Annealing)	70	2 min
35 (Extension)	72	2 min
1 (Extension)	72	10 min

3.2.5 Kanamycin^R cassette amplification

After digestion with restriction endonuclease *Dpn*1 (Section 2.5.2), verification of PCR amplification was performed via gel electrophoresis described in Section 2.4.5.1.

3.2.6 Integration of kanamycin^R cassette into the *E. coli* W3110*trp*A33 chromosome

3.2.6.1 Transformation

Transformation of W3110*trp*A33 harbouring *pKD46* is described in Section 2.5.4. The generation of electrocompetent *E. coli* W3110*trp*A33 cells harbouring *pKD46* is described in Section 2.5.4.1, while the electroporation protocol is detailed in Section 2.5.4.2. Resultant cultures were incubated at 30°C for 3 h before being plated on selection media.

3.2.6.2 Selection media

Cultures were plated onto LB media containing kanamycin (50 µg/mL) to select for transformants. Plates were incubated at 30°C overnight due to the temperature sensitive replicon, which allows the replication of plasmid at 30°C.

3.2.6.3 Verification of integration

Transformants were screened via colony PCR. A single colony was picked and added to 10 µL of sterile water which was then heated to 95°C for 10 mins and chilled on ice. That bacterial suspension was then used as a template for a PCR reaction described in Section 4.2.5.2. Successful transformants were used as the basis for the generation of an *E. coli* mutant deficient in both TSα and tryptophanase activity (W3110*trp*A33*tna*A) (Section 2.5.4.1).

3.2.7 Double mutant complementation

3.2.7.1 Transformation

Transformation of *E. coli* W3110*trp*A33*tna*A with INO (*pNAV112*) was performed as described in Section 2.5.4. Electrocompetent cells were generated which were subsequently transformed with maize IS (*pISP1*).

3.2.7.2 Selection media

Resultant cultures were plated onto M9 minimal media with or without added IPTG, 1mM indole and 1 mM tryptophan. Ampicillin was supplemented to select for transformants containing *pISP1*.

3.2.7.3 Verification

Verification of the presence of both indole synthase and indole oygenase genes in transformants was achieved via colony PCR. A single colony was picked and added to 10 µL of sterile water which was then heated to 95°C for 10 mins and chilled on ice. That bacterial suspension was then used as a template for a PCR reaction described in Section 2.4.7.2. The verification of a functional indole synthase/indole oxygenase enzyme system was qualified by the accumulation of indigo in resultant colonies.

3.3 Results

3.3.1 Construction of a plasmid harbouring *INO*

Plasmid *pACYC184* (Figure 3.1) was selected as the backbone for construction of a novel plasmid harbouring *INO* due to the presence of a tetracycline resistance marker and its origin of replication derived from p15A. This origin of replication is from a different incompatibility group than the origin of replication from *pUC* vectors and therefore these vectors are able to co-exist in bacteria.

3.3.2 Synthesis of *pNAV110*

A strategy was pursued to digest the *pACYC184* with restriction endonucleases *Xmn*1 and *Sca*1 leaving a blunt ended 3196bp backbone and a linear fragment of 1049bp in length inclusive of the chloramphenicol resistance gene (Figure 3.2). Re-ligation of those blunt ends and selection on tetracycline media ultimately identified a 3196bp plasmid with a tetracycline marker as the sole resistance gene (Figure 3.3 A). This newly synthesised plasmid was renamed *pNAV110* (Figure 3.3 B).

3.3.3 Integration of LacZ(α) into *pNAV110*

Cloning vector *pUC120* was used as a template for PCR amplification of the LacZ(α) gene downstream of the promoter, the region that facilitates transcription. Primers were designed to ultimately produce a blunt ended PCR product suitable for ligation into an *Xmn*1/*Sca*1 digested *pNAV110* plasmid. LacZ(α) fragments of 645bp in length were generated, purified via gel excision purification (QIAGEN) and run on an electrophoresis gel to determine purity and quantify the DNA (Figure 3.4 A). Ligation of the purified LacZ(α) fragment into the similarly blunt ended *pNAV110* was achieved and subsequent transformation of DH5α electrocompetent cells was performed via electroporation. Resultant cultures were incubated at 37°C for 30 mins and plated on LB^{tet} plates and left to incubate overnight once again at 37°C.

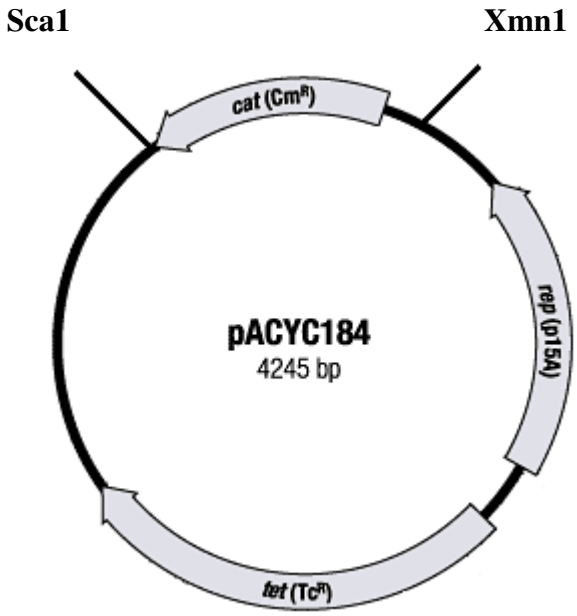


Figure 3.1 Cloning vector *pACYC184* displaying *Xmn*1 and *Sca*1 restriction enzyme sites, two resistance genes (Cm and Tet) and the rep p15A origin of replication.

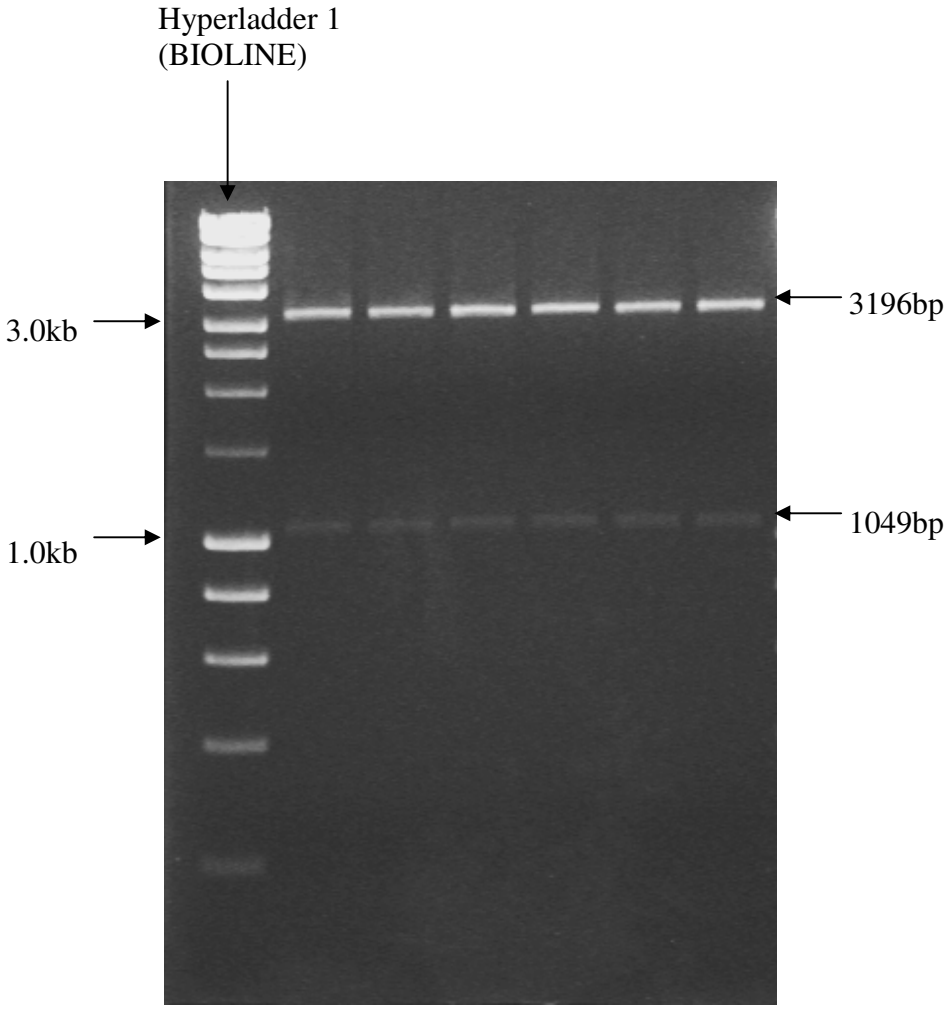
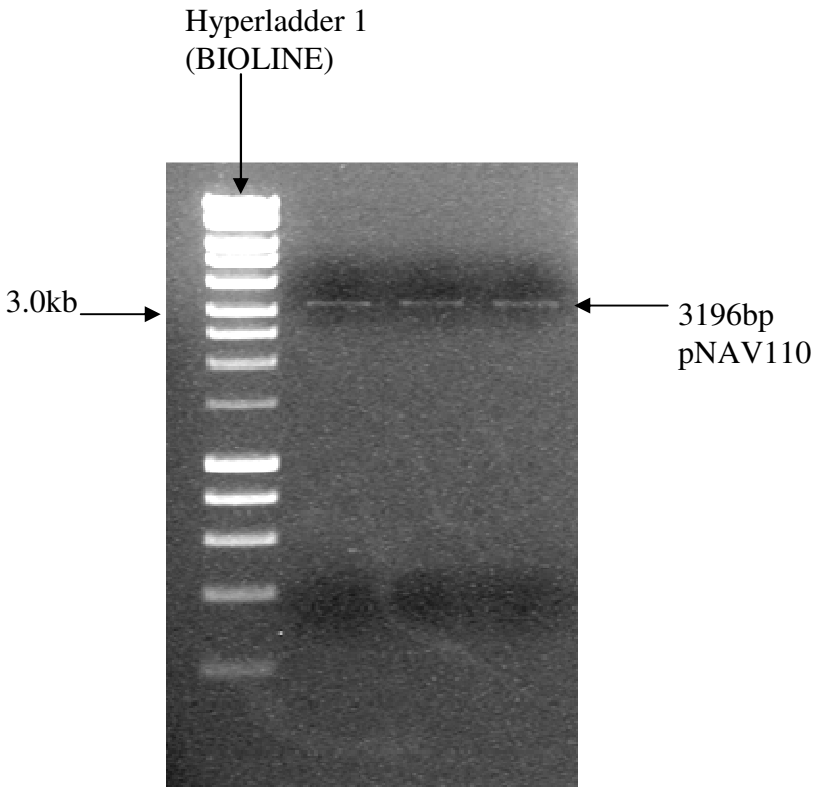


Figure 3.2 Digestion of *pACYC184* with *Xmn*1 and Sca1 produced a 3196bp backbone and a 1049bp fragment inclusive of the chloramphenicol resistance marker.

(A)



(B)

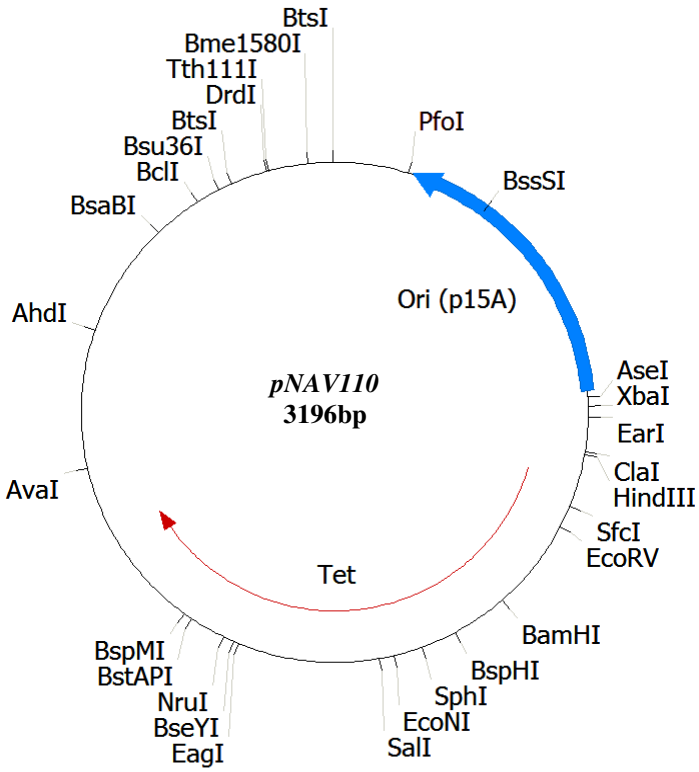


Figure 3.3 (A) Linearised *pNAV110* (3196bp) (re-ligated *pACYC184* omitting the Cm resistance marker) (in triplicate). (B) *pNAV110* construct

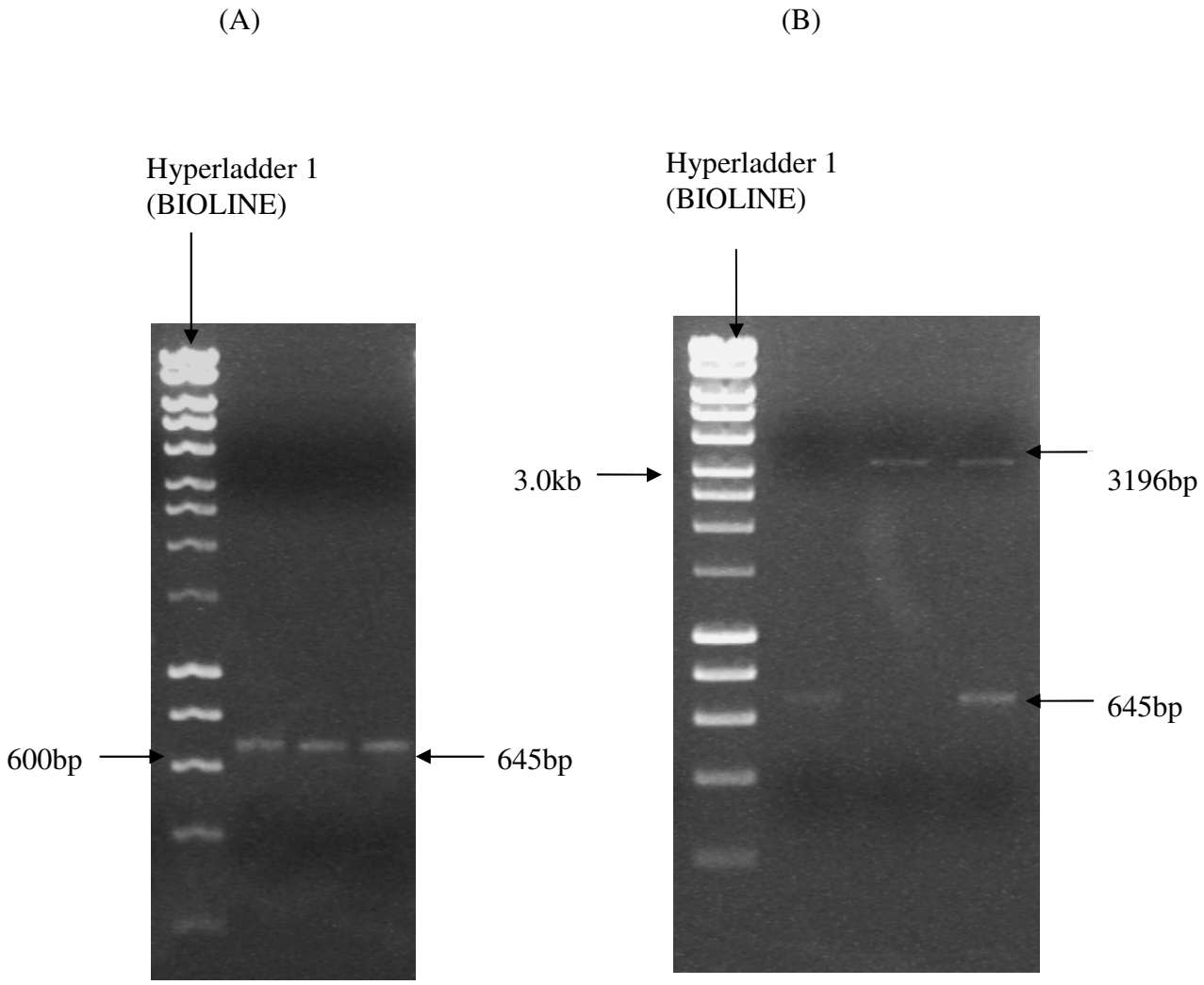


Figure 3.4 (A) Purified LacZ(α) PCR fragments (In triplicate). (B) Ligation of LacZ(α) into *pNAV110*, Lane 2: Purified LacZ(α) DNA, Lane 3: Digested *pNAV110* (*Xmn*1/*Sca*1), Lane 44: Digested *pNAV111* (*Xmn*1/*Sca*1) showing 645bp LacZ(α) and *pNAV110* backbone (3196bp)

IPTG and X-gal was added to the media enabling blue white colony screening. Blue colonies were screened for the presence of the LacZ(α) insert via restriction enzyme digestion (Figure 3.4 B) of the newly synthesised plasmid DNA classified as *pNAV111* (Figure 3.5).

3.3.4 Integration of *INO* into *pNAV111*

The construct *pNAV79* (created and supplied by Professor David Stalker) consists of a *pUC120* backbone with an 1163bp INO insert originally cloned from *Rhodococcus* sp. The cloning plasmid *pUC120* contains a unique *Nco*1 restriction site at the LacZ(α) start codon to facilitate IPTG inducible expression of cloned inserts. The construct *pNAV79* makes use of this feature by harbouring the INO (Figure 3.6). Twenty two base pairs downstream of the INO stop codon is the restriction site Pst1. Restriction enzyme digestion with *Nco*1 and *Pst*1 'drops out' the 1185bp INO fragment ready for ligation into *pNAV111* also digested with *Nco*1 and *Pst*1.

Figure 3.7 shows the 1185bp INO fragment excised from *pNAV79* while Figure 3.8 shows the ligation strategy of inserting the digested *INO* fragment into *pNAV111* directly interrupting the LacZ(α) gene and providing another selection process (Blue/White screening). This newly synthesised plasmid is denoted *pNAV112* (Figure 3.9). After ligation, transformation of *E. coli* DH5α electrocompetent cells was performed via electroporation and resultant cultures were incubated at 37°C for 30 mins and plated on LB^{tet} plates and left to incubate overnight once again at 37°C. Extraction of plasmid DNA was performed as per the instructions provided in Section 2.4.1.1.

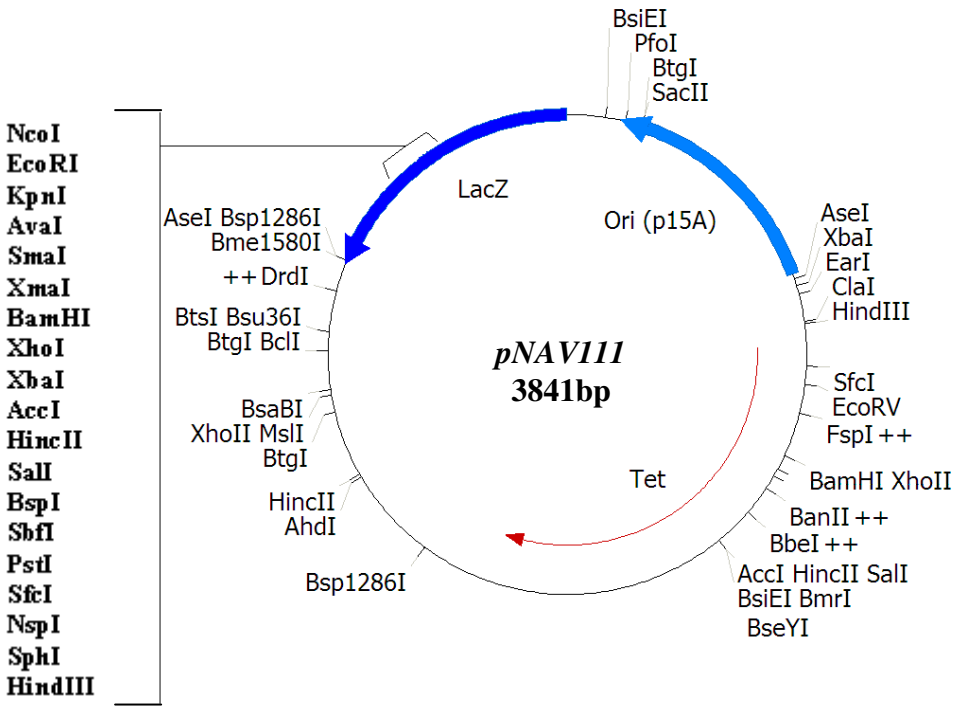


Figure 3.5 ***pNAV111*** **construct with LacZ(α) insert**

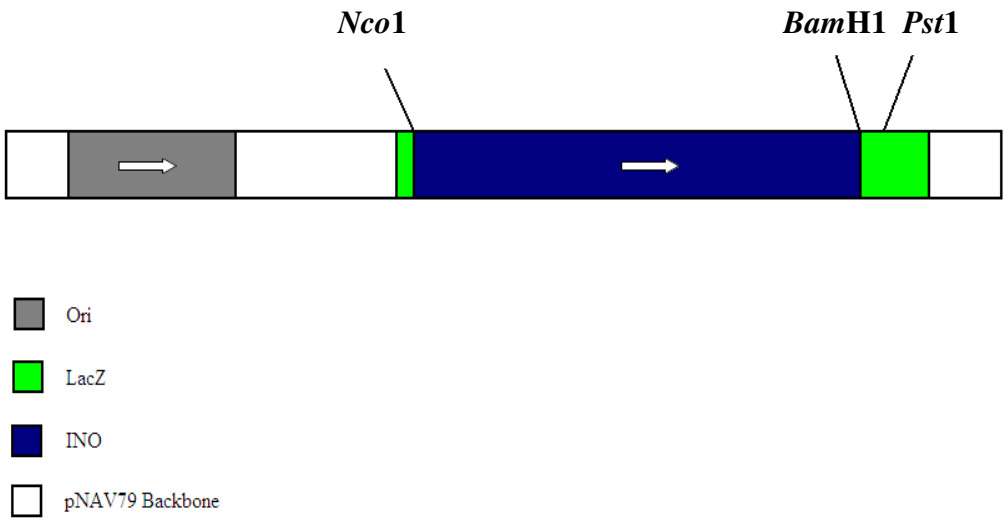


Figure 3.6 Linear representation of *pNAV79* with *INO* interrupting the LacZ(α) gene

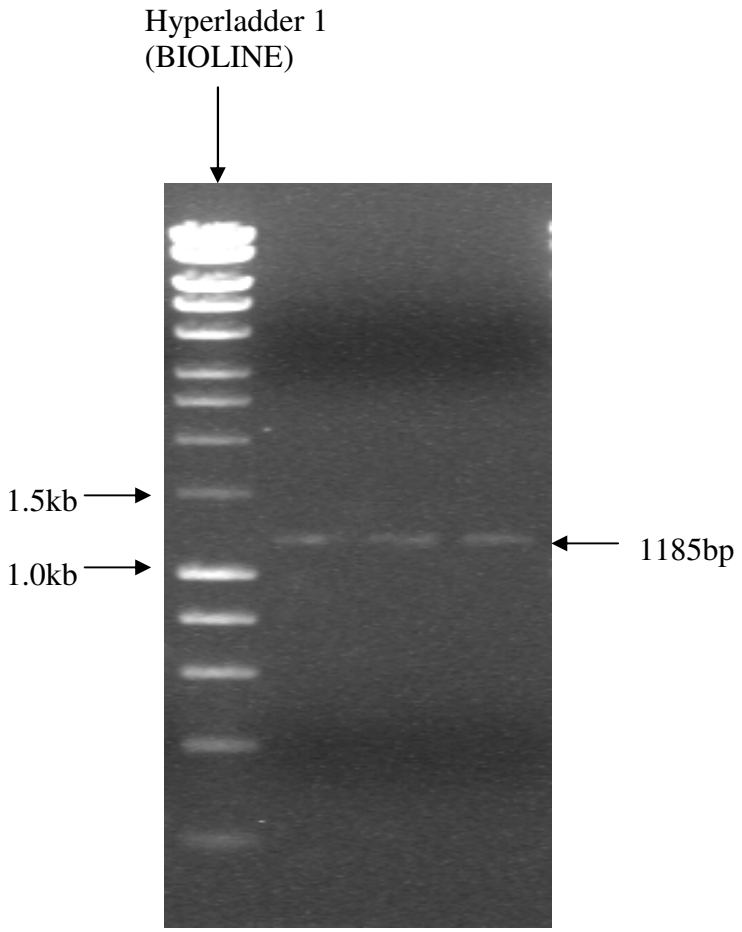


Figure 3.7 Excised *INO* fragment (Excised from *pNAV79* via digestion with *Nco*1 and *Bam*H1) (In triplicate)

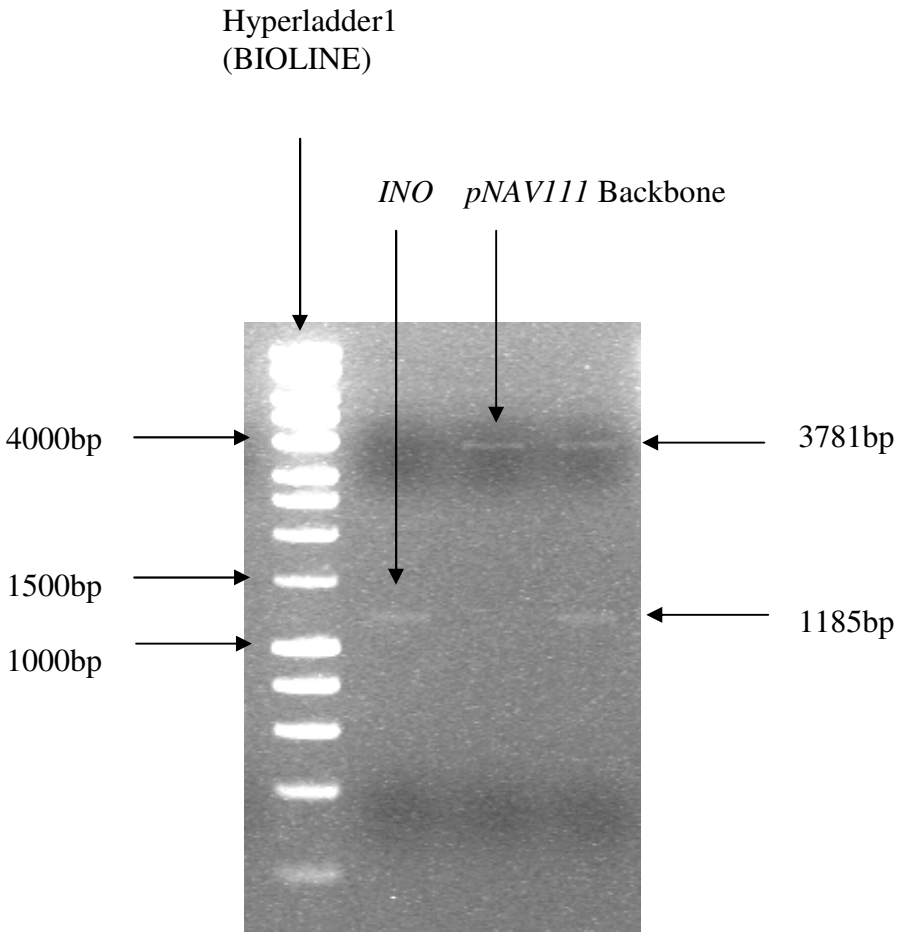


Figure 3.8 Ligation of INO into *pNAV111*, Lane 2: Purified INO digested with *Nco*1 and *Pst*1, Lane 3: Digested *pNAV111* (*Nco*1/*Pst*1), Lane 4: Digested *pNAV112* (*Nco*1/*Pst*1) showing linearised *pNAV111* backbone (3781bp) and *INO* insert (1185bp).

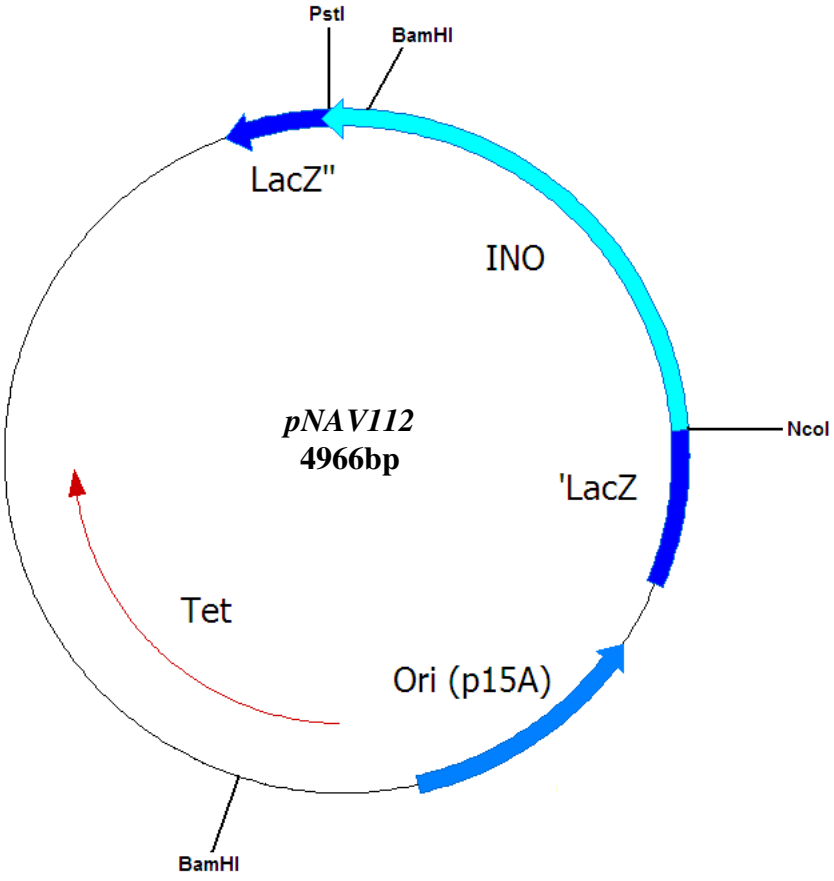


Figure 3.9 *pNAV112* construct with *Bam*HI, *Nco*I and *Pst*I restriction enzyme sites

3.3.5 Complementation controls

The parent strain (*E. coli* W3110*trp*A33) was plated onto M9 minimal media with and without tryptophan and with and without indole supplementation. The strain did not grow in the absence of either supplement, however the strain was rescued by 0.5 mM tryptophan and 1.0 mM indole (Figure 3.10). Both tryptophan and indole supplementation permitted growth as detailed in the corresponding pathway schematics (Figure 3.11). Replicating the experiment performed by Kramer and Koziel, (1995), introduction of the maize indole synthase complemented W3110*trp*A33 (Figure 3.12). The supplementation of tryptophan and indole, in W3110*trp*A33/IS, up to six fold that of the required amount to rescue W3110*trp*A33, in M9 minimal media and subsequent plating of the parent was performed to ascertain whether the presence of additional indole, either by tryptophan degradation or direct supplementation, enabled indigo synthesis. Although growth was observed, indigo accumulation did not occur (Figure 3.13).

The pathways corresponding to bacterial growth are also detailed (Figure 3.14). The introduction of INO into the parent strain and subsequent plating onto M9 minimal media supplemented with and without tryptophan and indole led to indigo accumulation in W3110*trp*A33/INO (Figure 3.15). The corresponding pathways are shown in Figure 3.16. When the *Rhodococcus INO* is expressed in W3110*trp*A33, the oxidation of 'unbound' indole leads to indigo synthesis. Considering the analytical nature of this potential bacterial assay system, tryptophanse based indole synthesis and subsequent indigo accumulation must be eliminated to accurately assess the indigo producing capabilities of any putative indole synthases.

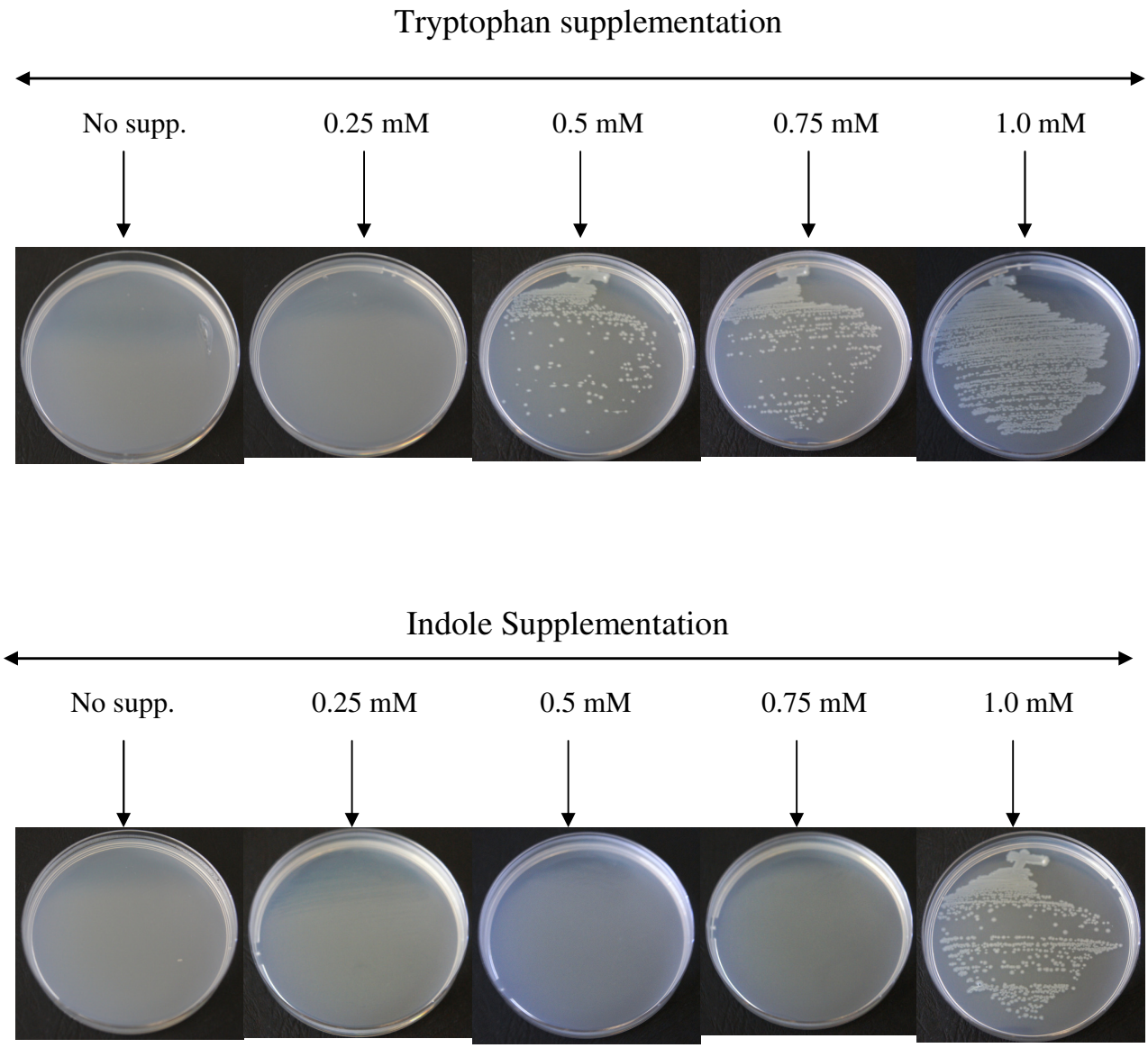


Figure 3.10 Tryptophan and indole supplementation in M9 minimal media, *E. coli* W3110*trp*A33.

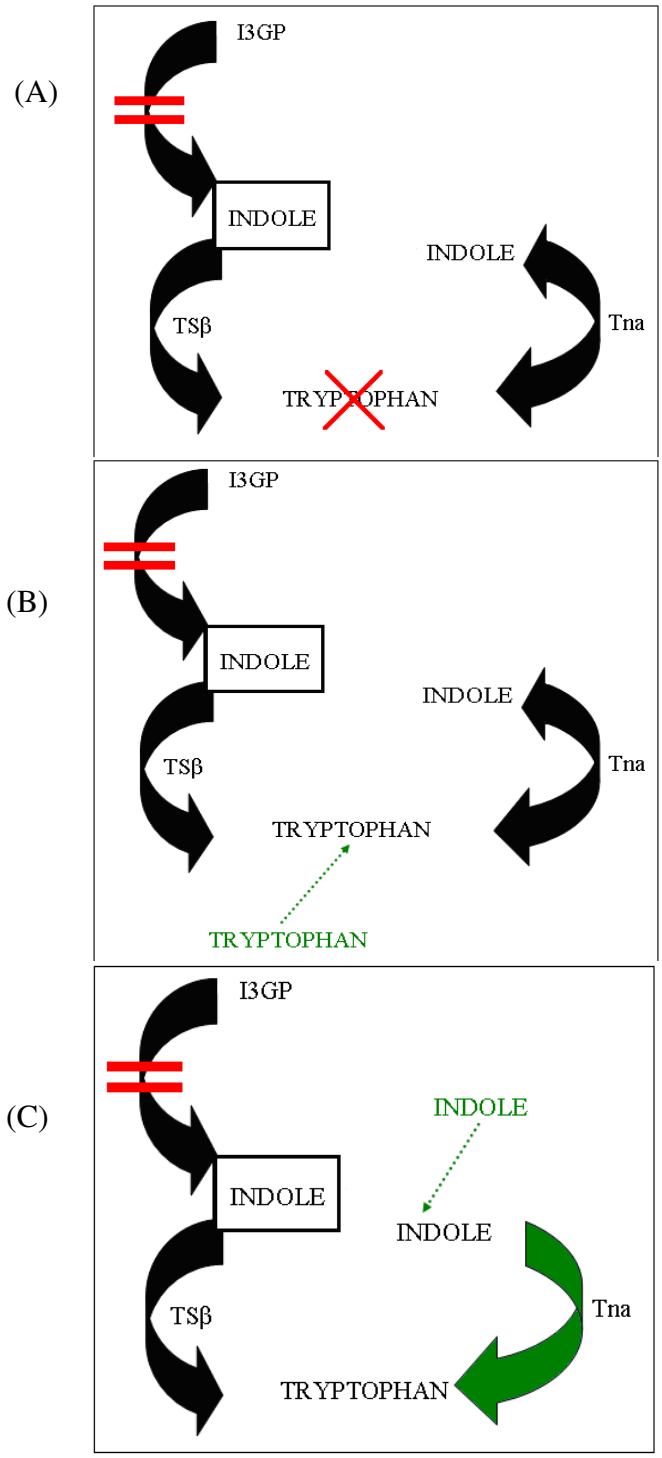
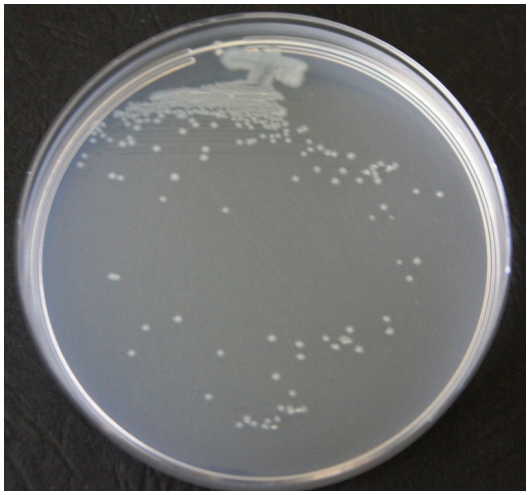


Figure 3.11 (A) Scheme representing tryptophan biosynthetic pathways in *E. coli* W3110*trp*A33. (B) The strain is rescued by 0.5 mM tryptophan supplementation and 1.0 mM indole supplementation (C). The double red lines indicate deficient TSα activity due to the *trp*A mutation. The bold green arrows represent the mechanism activated when the strain is rescued by indole. The dotted green arrows indicate which chemical is supplemented.

(A)



(B)

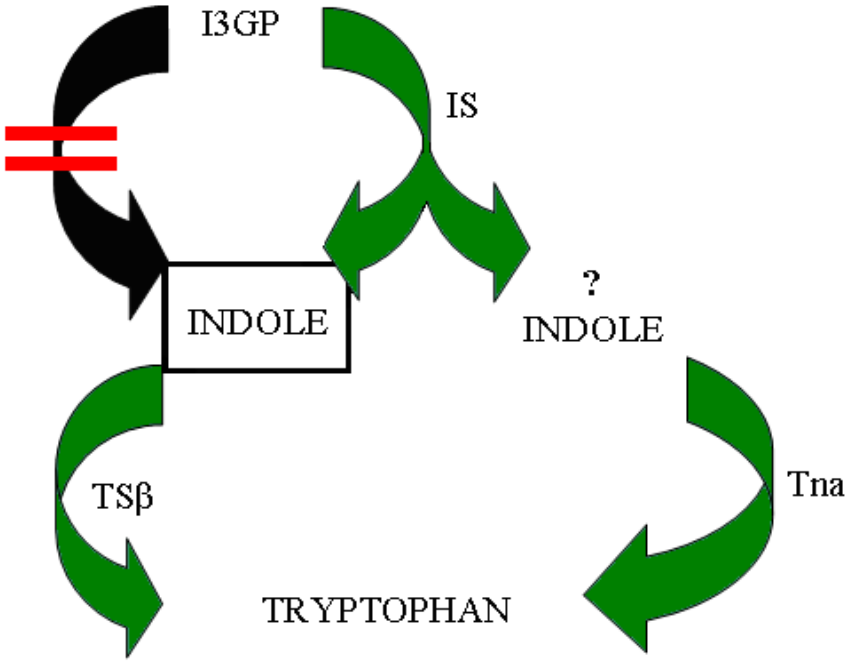


Figure 3.12 (A) Complementation of *E. coli* W3110*trp*A33 with the maize indole synthase (W3110*trp*A33/IS). Growth of colonies on M9 minimal media with no tryptophan or indole supplementation indicates successful complementation (B) Pathway detailing mode of complementation. The double red lines indicate deficient TSα activity due to deletion of the *trp*A gene. The bold green arrows represent the potential activated mechanisms when the strain is complemented by IS.

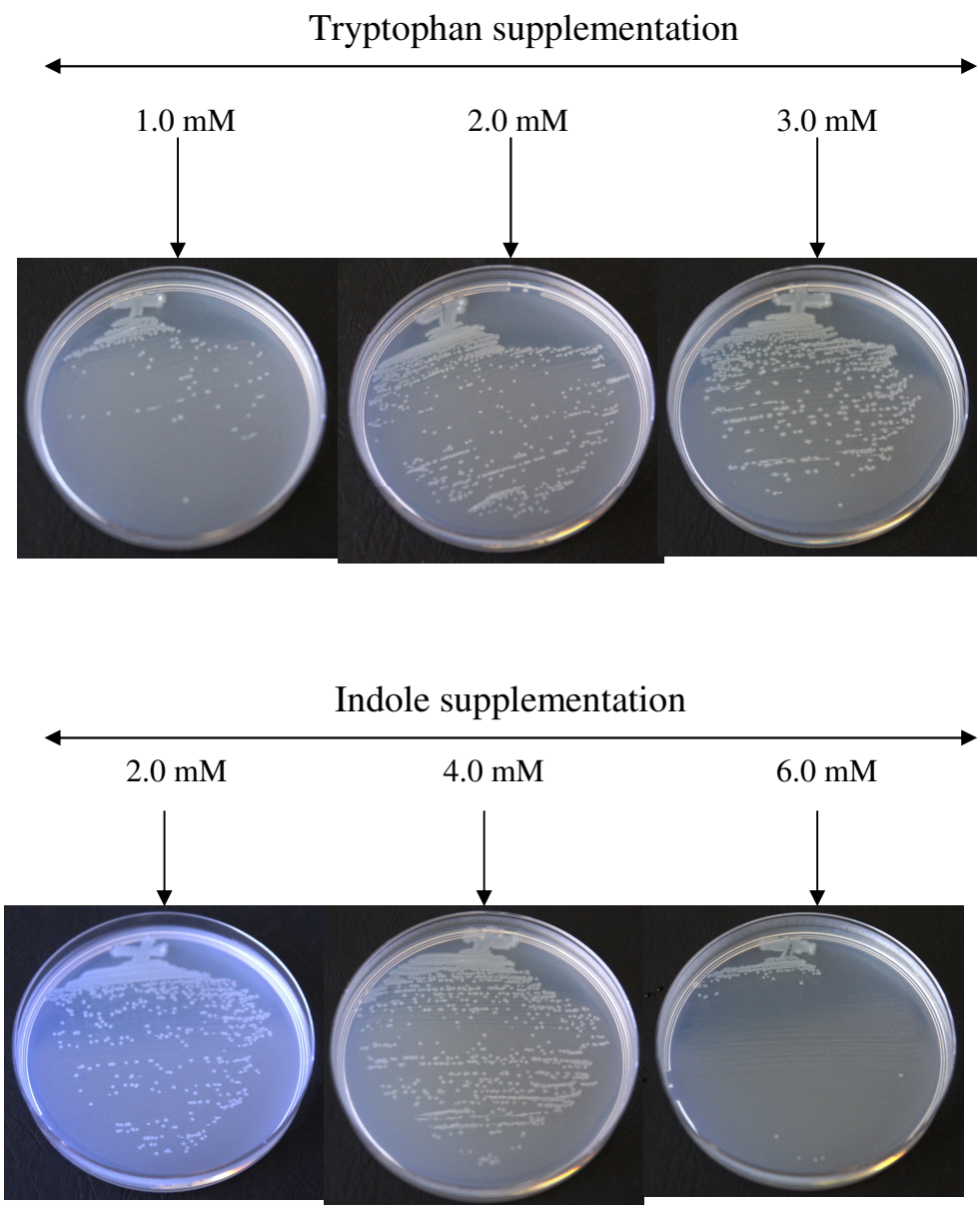


Figure 3.13 Supplementation with varying levels of tryptophan and indole in M9 minimal media to determine if the presence of excess indole enables indigo accumulation in *E. coli* W3110*trp*A33/IS.

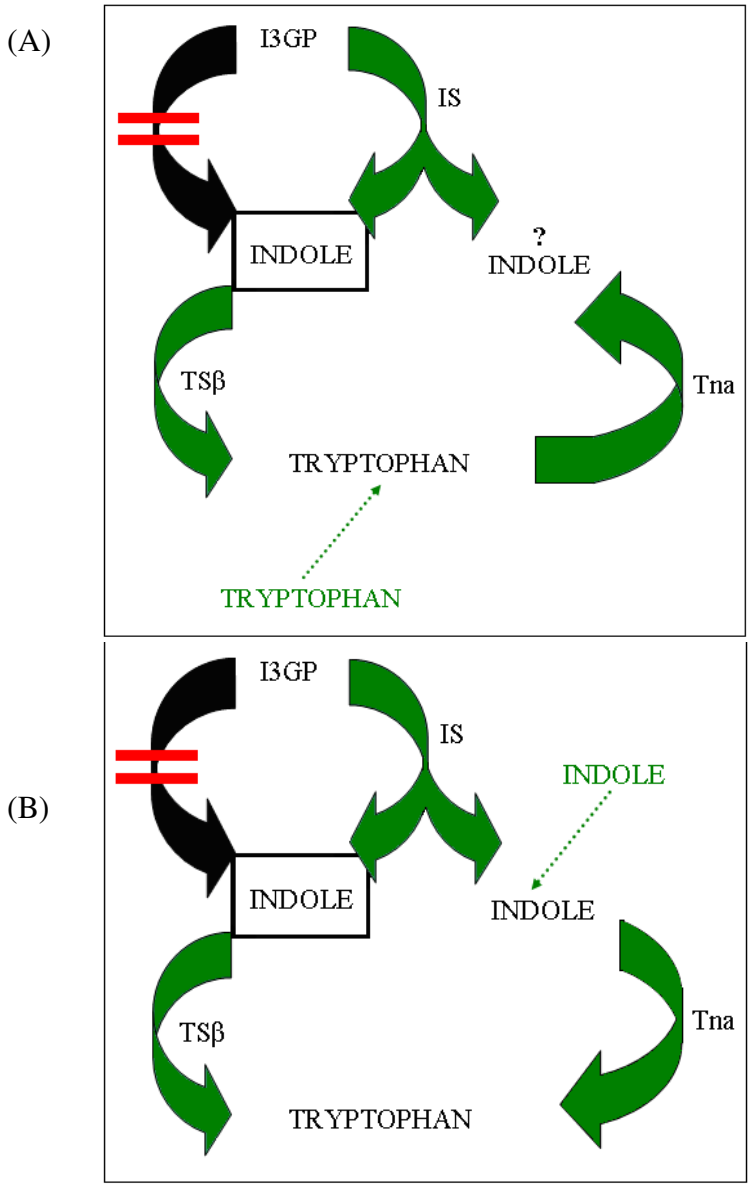


Figure 3.14 Comparison of tryptophan biosynthetic pathways when *E. coli* W3110*trp*A33/IS is grown in the presence of tryptophan supplementation (A) and excess indole supplementation (B). The double red lines indicate deficient TSα activity due to deletion of the *trp*A gene.

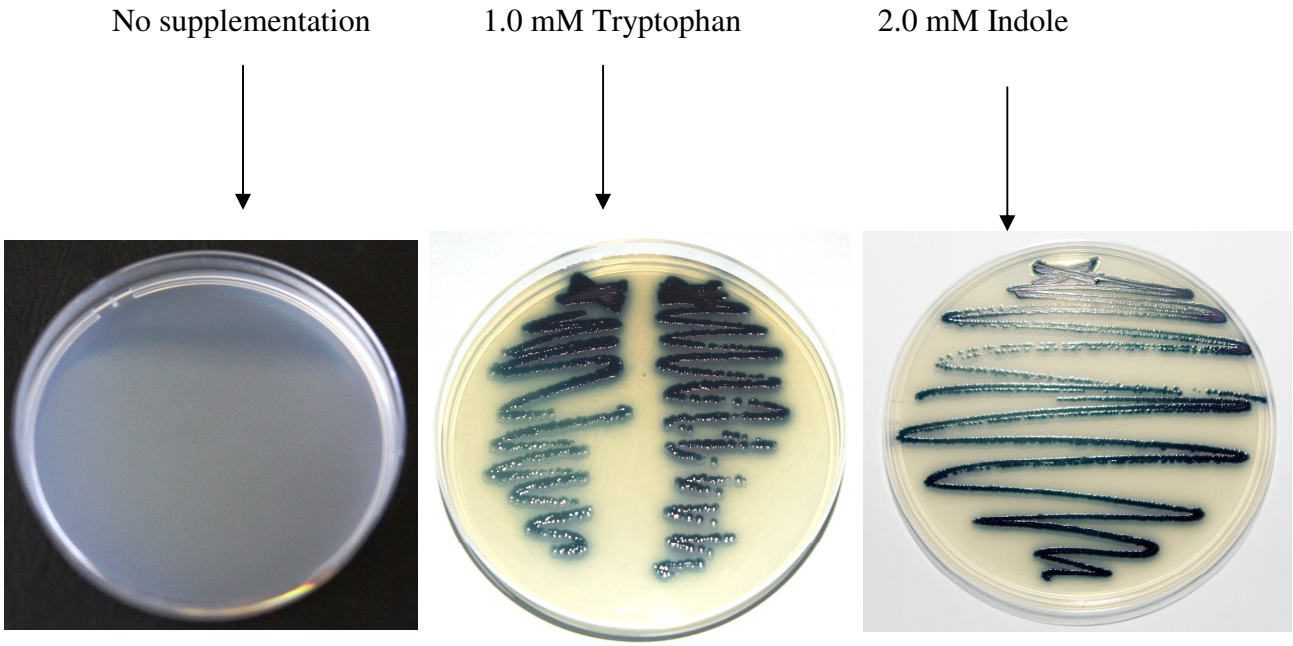
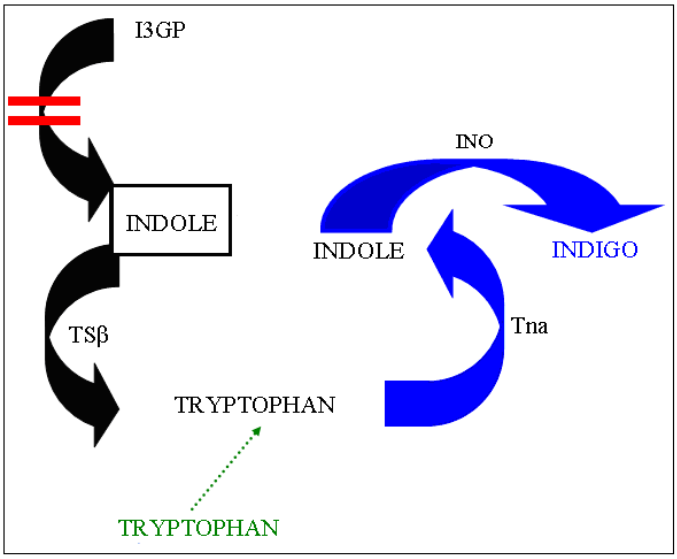


Figure 3.15 *E. coli*W3110*trp*A33 transformed with INO and plated on M9 minimal media with indole or tryptophan supplementation.

(A)



(B)

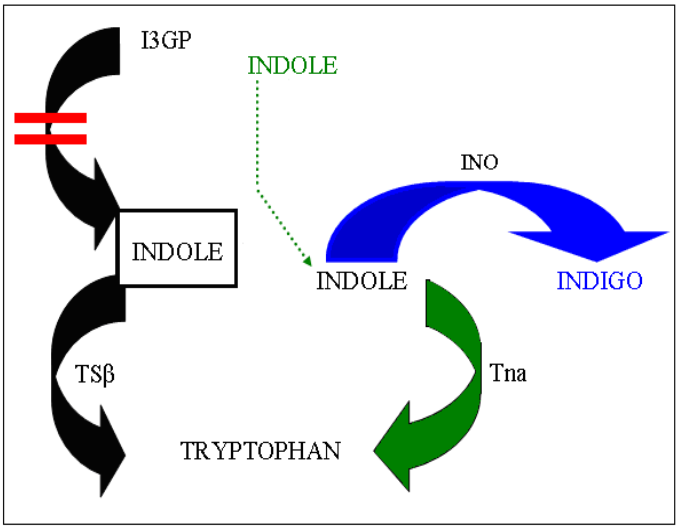


Figure 3.16 Comparison of tryptophan biosynthetic pathways when *E. coli* W3110*trp*A33 transformed with INO is rescued by 1.0 mM tryptophan supplementation (A) and 2.0 mM indole supplementation (B). The double red lines indicate deficient TSα activity due to deletion of the *trp*A gene. The bold black arrows represent the inactive mechanisms within the pathway. The bold blue arrows represent the mechanisms responsible for indigo accumulation when the strain is rescued by either tryptophan or indole. The bold green arrow represents mode of tryptophan biosynthesis. The dotted green arrows indicate which chemical is supplemented.

3.3.6 Red disruption system

As described in the introduction, *pKD46* harbours three genes, γ, β and *exo* that are responsible for enhanced recombination efficiency (Datsenko and Wanner, 2000). The products of these genes are Gam, Bet and Exo respectively. Gam inhibits exonuclease activity permitting Bet and Exo to access the linear DNA and initiate recombination. W3110trpA33 was transformed with *pKD46* after the plasmid was passaged through the *S. typhimurium* strain (Table 2.1). The *pKD46* plasmid carries a temperature sensitive conditional replicon that allows replication of plasmid at 30°C only. The lambda red genes which drive the recombination of short linear PCR fragments are expressed under the control of the arabinose promoter (araC) (Figure 3.17). The same process was performed for *pKD13* maintained in *E. coli* BW25141 to extract the plasmid to be used as a template for PCR amplification of the kanamycin resistance gene. The plasmid replicates under the control of the R6K origin of replication and the kanamycin resistance gene is flanked by FRT sites (Figure 3.18 A).

3.3.7 Amplification of the kanamycin^{R} cassette from *pKD13*

Gene knockout primers were designed from the *pKD13* template (Datsenko and Wanner, 2000). Both primers have 20 nucleotide 3' ends for priming upstream (P1) and downstream (P2) of the FRT sites flanking the kanamycin resistance gene and 50 nucleotide 5' ends homologous to upstream (H1) and downstream (H2) chromosomal sequences for the *tna*A gene in W3110trpA33 (Figure 3.18 B). H1 includes the Met start codon for *tna*A while H2 includes the stop codon and 29 nucleotides downstream. The 1.4kb PCR products were purified and digested with *Dpn*1 (Figure 3.19) to eliminate unamplified template DNA and to ensure an efficient integration of the resistance cassette into the *tna*A gene of W3110trpA33.

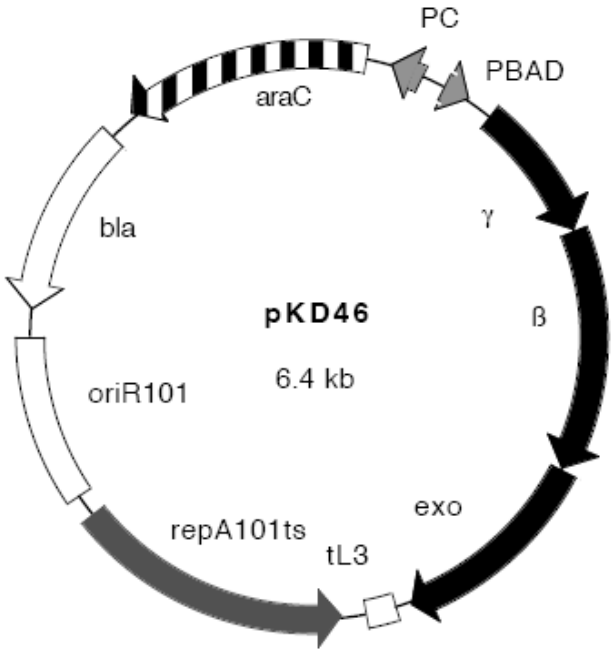
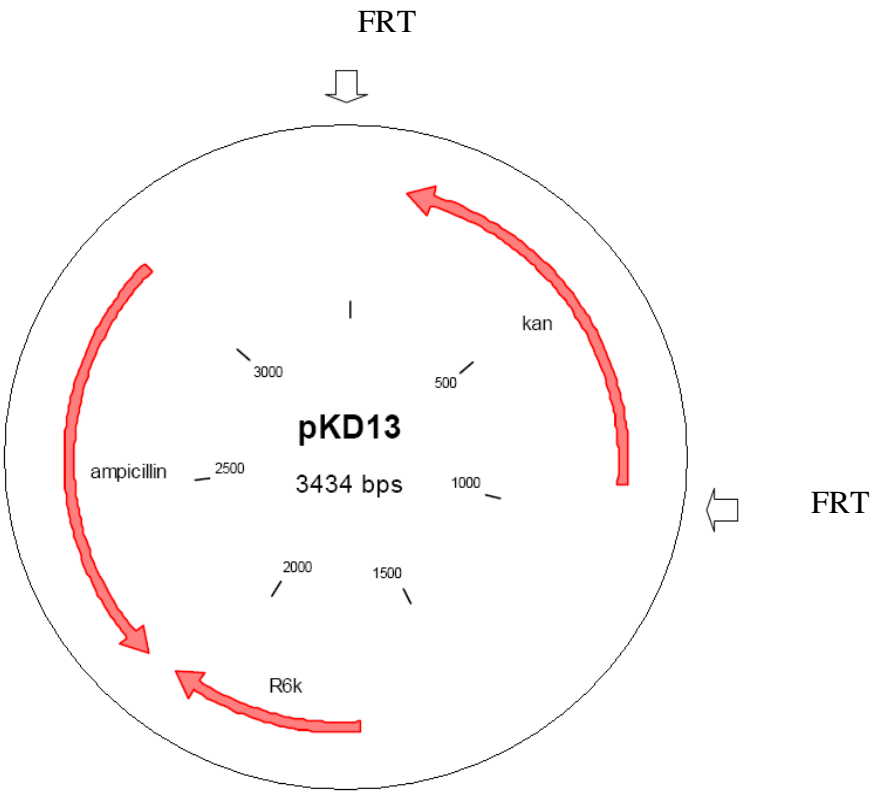


Figure 3.17 Red recombinase expression plasmid *pKD46*

(A)



(B)

Figure 3.18 (A) The *pKD13* plasmid with the R6K origin of replication. The kanamycin antibiotic cassette is flanked by FRT sites. (B) Primer design for 'knockout' gene experiment (*pKD13* used as primer template)

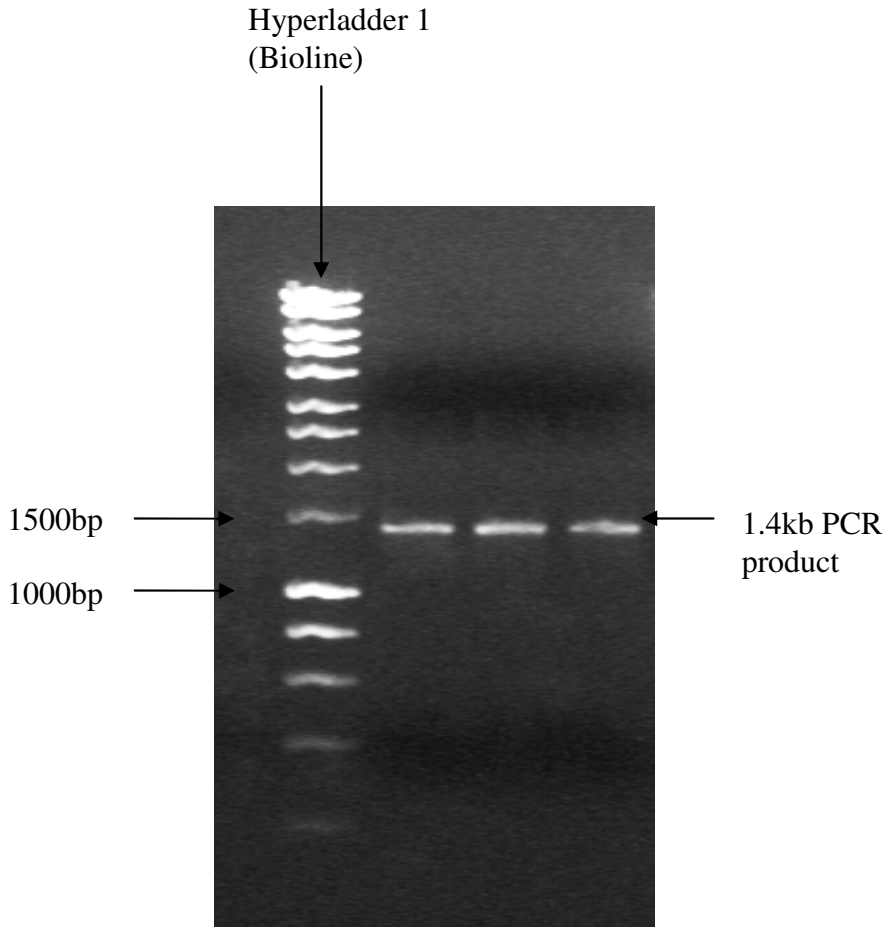


Figure 3.19 Purified *pKD13* PCR generated DNA fragment (In triplicate: lanes 2-4) amplified with primers flanking the kanamycin resistance gene. PCR products digested with endonuclease *Dpn*1

3.3.8 Integration of the PCR product into the *tna*A gene of *E. coli* W3110*trp*A33 and PCR verification

Freshly made electrocompetent W3110*trp*A33 cells harbouring the *pKD46* plasmid were transformed with the linear DNA fragments generated by PCR via electroporation. Resultant 1mL cultures were incubated at 37°C for 2 h and then spread onto LB media supplemented with kanamycin to a final concentration of 20mg/mL and grown overnight. Positive transformants were screened via colony PCR.

Two separate PCR's with different primer combinations were used to identify successful integration. Primers P1 and P2 were used to show the successful integration of the entire 1.4kb linear PCR fragment containg the kanamycin cassette with flanking FRT sites, while control primers (Kt and K2) were designed from the *pKD13* (Figure 3.20) template to confirm successful recombination by identifying a 471bp kanamycin fragment.

Amplification with both sets of primers (Figure 3.21) indicates that the mutants have integrated the cassette. For both sets of primers the *pKD13* plasmid backbone was used as a positive control, while PCR performed using no template DNA was used as a negative control. A representation of successful recombination is also shown (Figure 3.22). The double mutant *E. coli*, strain W3110*trp*A33 deficient in both TSα activity and tryptophanase activity and was renamed W3110*trp*A33*tna*A.

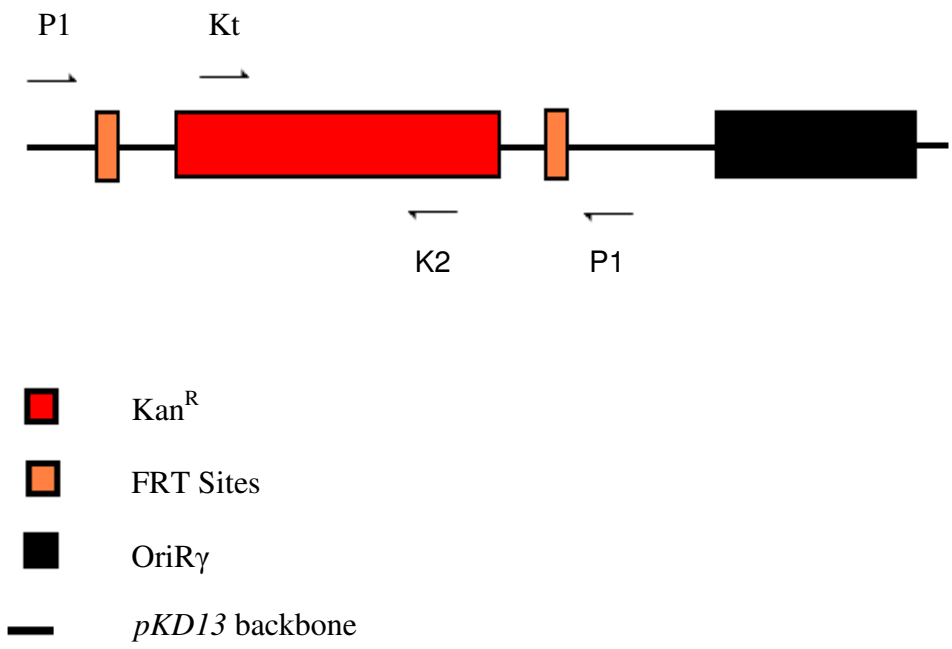


Figure 3.20 Linear representation of *pKD13* plasmid template with priming sites

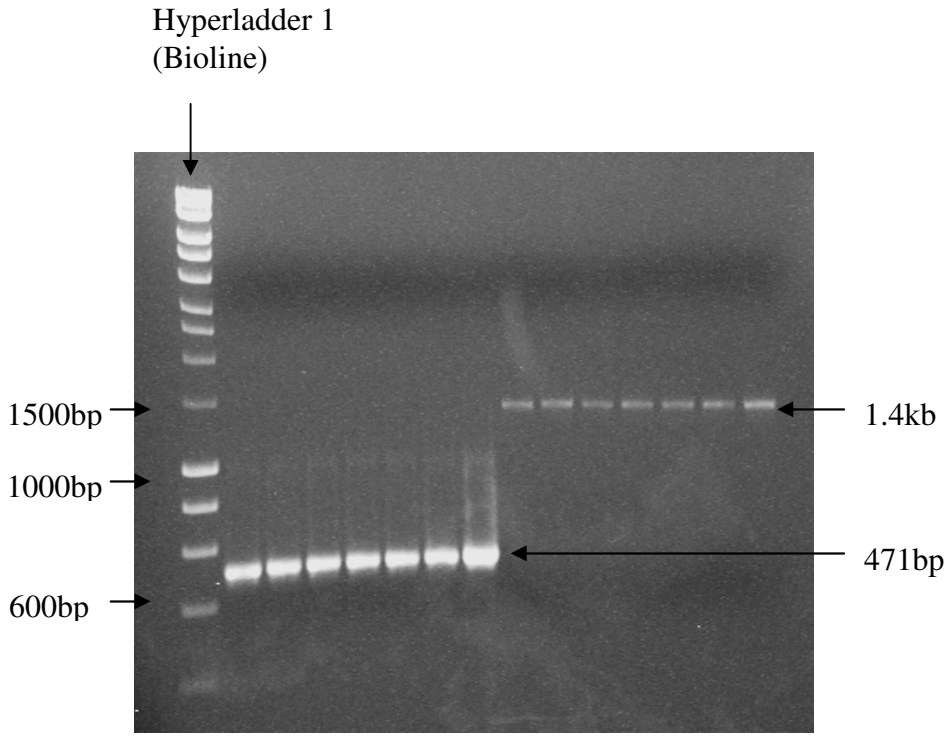


Figure 3.21 Colony PCR verification of kanamycin^{R} cassette integration into *tna*A gene of *E. coli* W3110*trp*A33 (Lane 2-7: Kan fragments, lane 8: Kan positive control, lane 9-14: fragments amplified with P1 and P2 primers, lane 15: P1 and P2 positive control, (negative controls not shown)

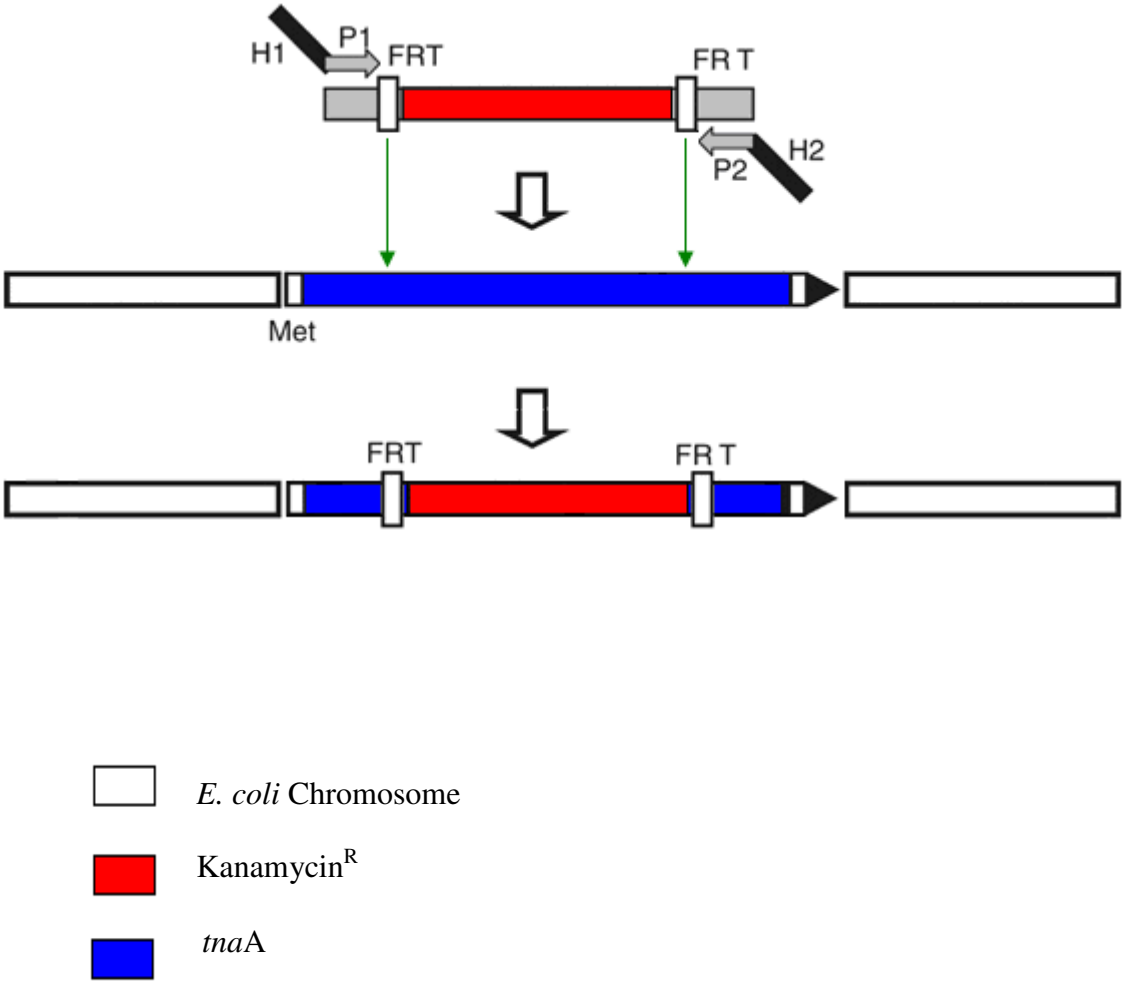


Figure 3.22 Construction of single gene 'knockout' mutant. Sequential gene disruption strategy. H1 and H2 refer to *tna*A homology regions. P1 and P2 refer to priming sites. Arabinose supplementation induces recombination event (green arrows) and subsequent disruption of the *tna*A gene rendering it non functional and 'knocking it out' of the *E. coli* chromosome.

3.3.9 Indole synthase/indole oxygenase sequential enzyme system

3.3.9.1 Double mutant cell line controls

Electrocompetent cells were prepared from *E. coli* strain W3110*trp*A33*tna*A and transformed with INO (*pNAV112*). Electrocompetent cells of W3110*trp*A33*tna*A harbouring *pNAV112* were made in preparation for cross complementation experiments. The newly synthesised cells were renamed W3110trpA33*tna*A/INO. The construction of a double mutant deficient in both TSα and tryptophanase is necessary to ensure indigo accumulation does not result from tryptophanase activity (Figure 3.23).

3.3.9.2 Double mutant control experiments

To examine the effects of the knockout of the *tna*A gene in the W3110*trp*A33 chromosome, the strain, W3110*trp*A33*tna*A was plated onto M9 minimal media and with indole and tryptophan supplementation. Once again, 0.5 mM of tryptophan was able to rescue the strain, however indole supplementation failed to rescue the double mutant strain, indicating failure of tryptophanase catalysing tryptophan biosynthesis from indole confirming the 'knockout' of the tryptophanase gene (Figure 3.24).

3.3.9.3 Complementation of W3110trpA33tnaA with maize IS

To assess whether the maize indole synthase makes a functional enzyme complex with an *E. coli* TSβ sub-unit, the double mutant strain was transformed with *pISP1* (IS) and plated onto M9 minimal media. Growth of colonies indicated successful complementation, allowing the conclusion that the maize indole synthase is able to complement the double mutant strain when co-expressed with an *E. coli* TSβ sub-unit (Figure 3.25).

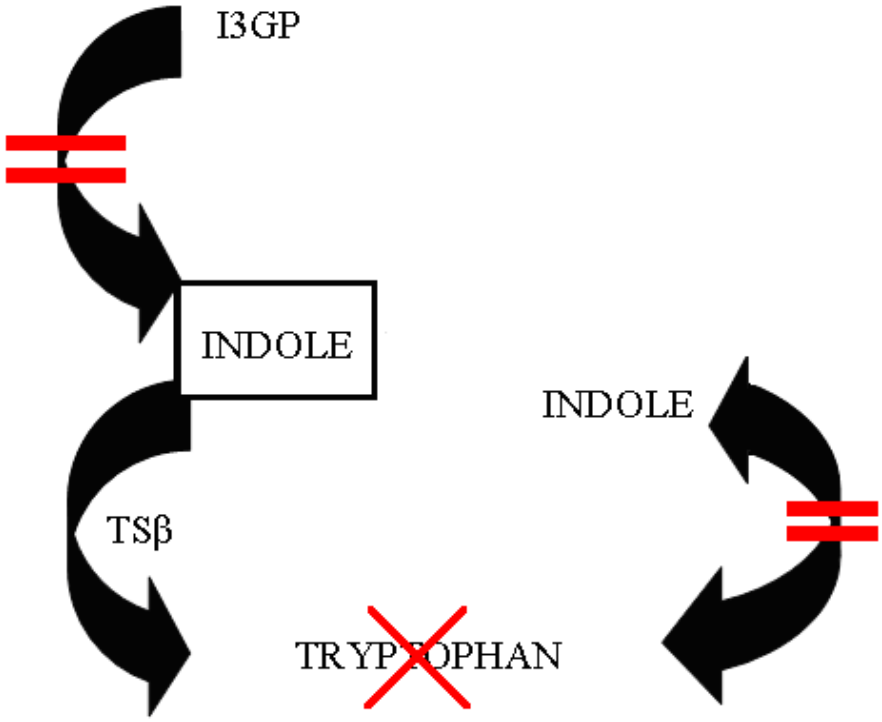


Figure 3.23 Schematic of the last two steps of tryptophan biosynthesis in *E. coli* W3110*trp*A33*tna*A. The double red lines indicate deficient TSα and Tna activity due to mutation of the *trp*A and *tna*A genes.

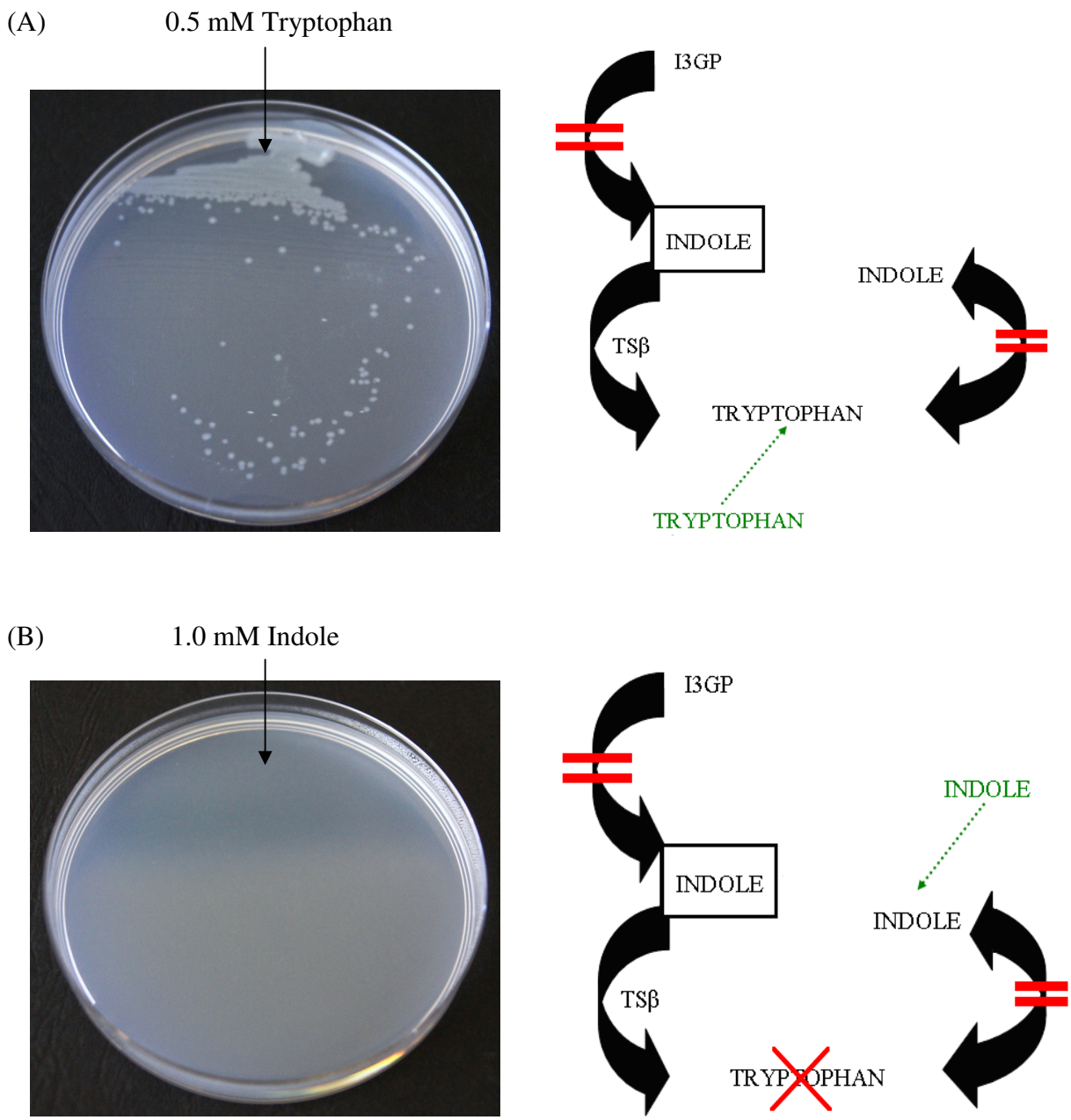


Figure 3.24 (A) *E. coli* W3110*trp*A33*tna*A rescued by 0.5 mM tryptophan. (B) 1.0 mM fails to rescue W3110*trp*A33*tna*A confirming elimination of the *tna*A gene. Corresponding pathways are detailed beside images of each plate. The double red lines indicate deficient TSα and Tna activity due to mutation of the *trp*A and *tna*A genes. The dotted green arrows indicate which chemical is supplemented.

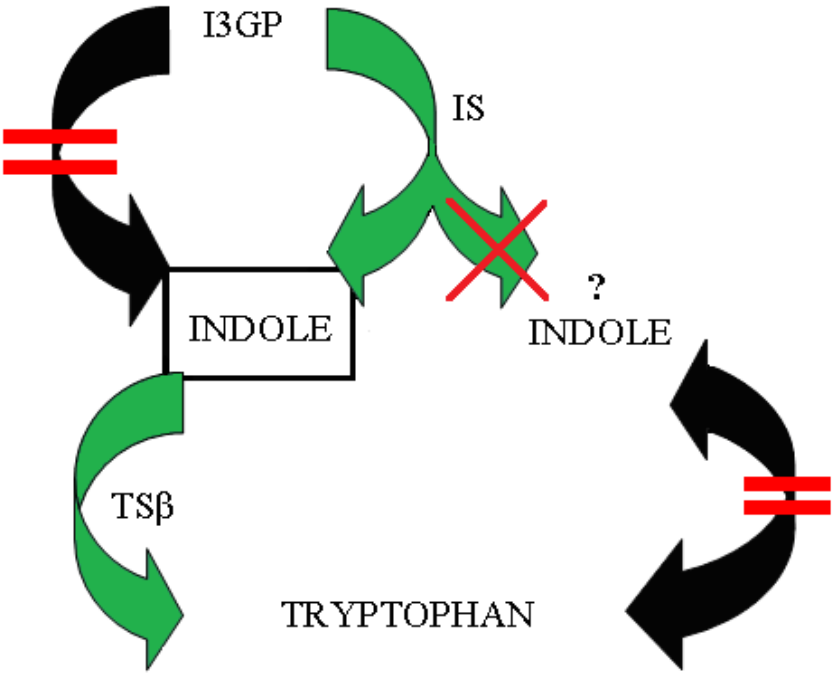


Figure 3.25 Complementation of *E. coli* W3110*trp*A33*tna*A with the maize IS. The double red lines indicate deficient TSα and Tna activity due to mutation of the *trp*A and *tna*A genes. The bold green arrows represent the mechanism by which tryptophan is produced.

3.3.9.4 Indigo accumulation in W3110trpA33tnaA

To assess the functionality of the IS/INO complex within *E. coli*, W3110*trp*A33*tna*A/INO was transformed with the maize IS to assess the amount of indigo that accumulated in the double mutant *E. coli* strain. Plating of W3110*trp*A33*tna*A/INO transformed with the IS on M9 minimal media alone and supplemented with 1.0 mM tryptophan produced white colonies, however 2.0 mM of indole stimulated the synthesis of indigo coloured bacteria (Figure 3.26). The corresponding pathways detail the mode of tryptophan biosynthesis and indigo production (Figure 3.27).

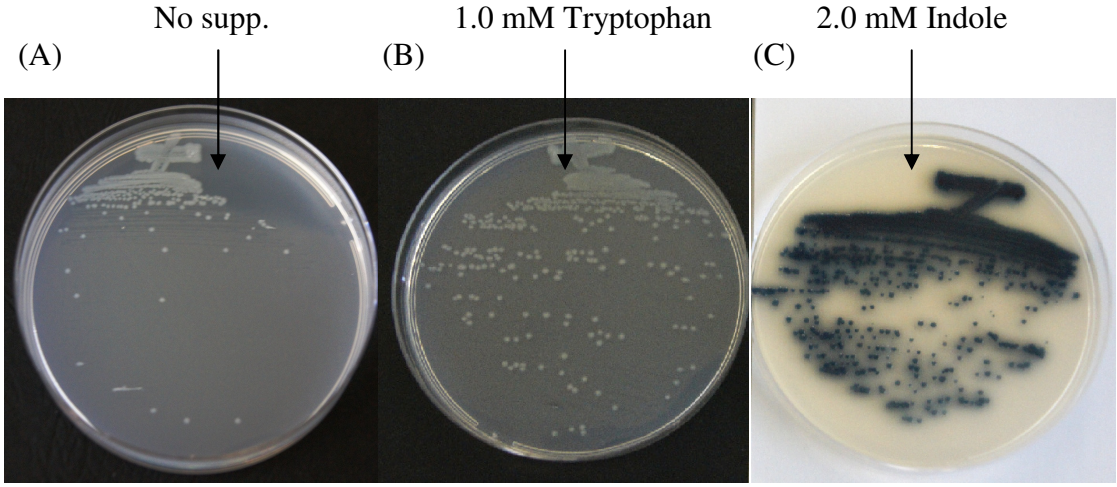


Figure 3.26 *E. coli*W3110*trp*A33*tna*A/INO transformed with IS and plated on M9 minimal media with indole or tryptophan supplementation. Indigo accumulation is evident when the media is supplemented with indole but not with tryptophan.

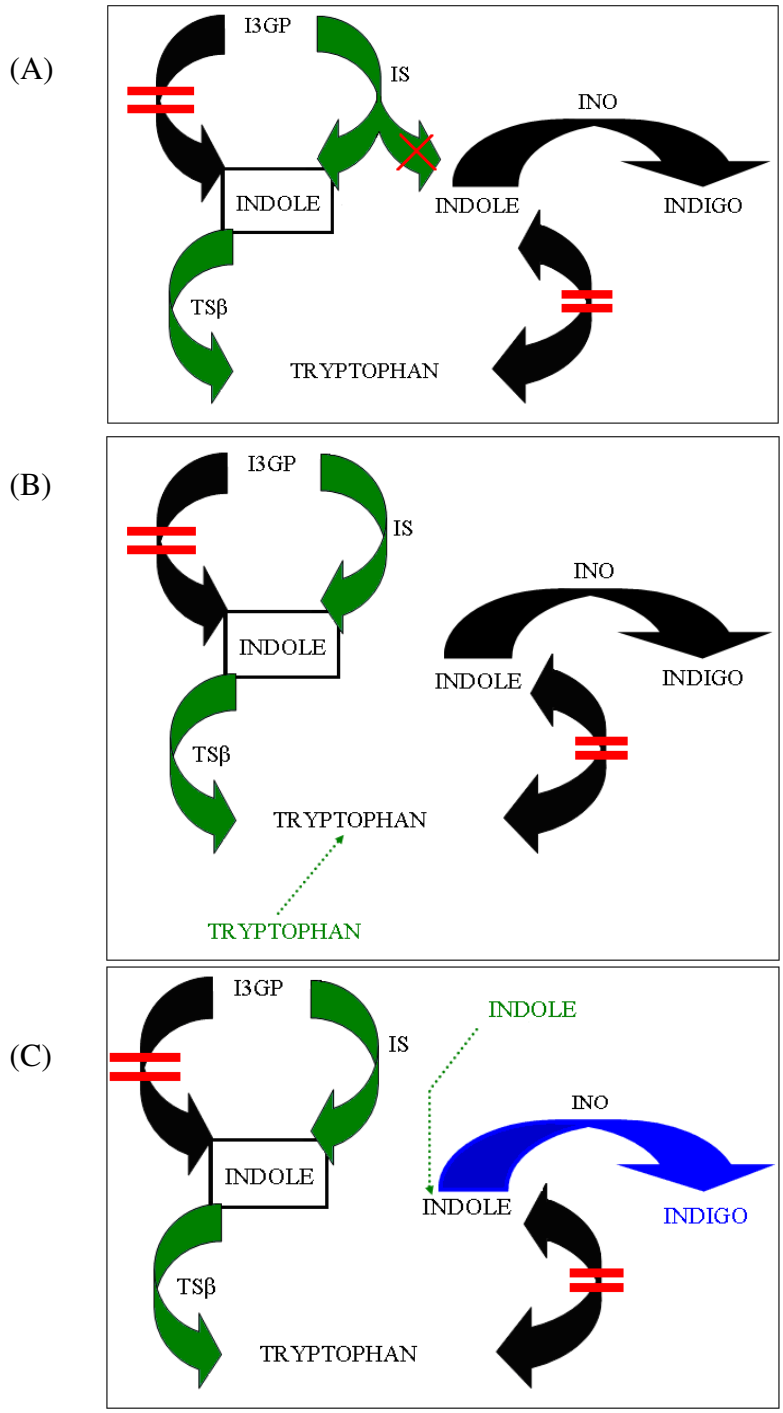


Figure 3.27 Pathways when *E. coli* W3110*trp*A33*tna*A/INO is transformed with IS. The double red lines indicate deficient TSα and Tna activity due to mutation of the *trp*A and *tna*A genes. Bold black arrows indicate inactive mechanisms. The bold green arrows represent the probable mode of tryptophan biosynthesis. The bold blue arrows represent the mechanism responsible for indigo accumulation; the dotted green arrows indicate which chemical is supplemented.

3.4 Discussion

The IS from maize, complements a TSα deficient *E. coli* strain, in that it would appear to perform the *trp*A function in the mutated W3110*trp*A33 strain leading to tryptophan biosynthesis and normal growth. This is consistent with the findings of Kramer and Koziel, (1995) who report the complementation of a *trp*A deficient strain of *E. coli* with the IS gene.

Considering that this maize IS appears to catalyse the synthesis of indole intended for DIMBOA synthesis in germinating maize seedlings and not for tryptophan biosynthesis (Melanson *et al*, 1997), the mode by which tryptophan is produced in the *trp*A mutant is a point for discussion. Is the plant indole synthase capable of forming a functional enzyme complex with an *E. coli* TSβ sub-unit, or is the indole generated by the maize enzyme utilised for tryptophan biosynthesis via another enzymatic mechanism?

The existence in gram negative bacteria of the catabolic enzyme tryptophanase provides a different means by which tryptophan can be produced from indole in *E. coli*. Considering that the maize IS independently catalyses the synthesis of indole (Melanson *et al*, 1997), tryptophan biosynthesis catalysed by tryptophanase activity is a possibility. When tryptophan levels exceed those required by the organism for metabolism, induction of the tryptophanase operon is the stimulus for the degradation of tryptophan catalysed by tryptophanase resulting in the synthesis of indole. However, the tryptophan-indigo reaction is active in both directions (Watanabe and Snell, 1972; Gong *et al*, 2001). Therefore the synthesis of tryptophan as a consequence of tryptophanase utilising indole as substrate is a distinct possibility.

The potential of maize IS/*E. coli* TSβ enzyme complex catalysing the biosynthesis of tryptophan from I3GP would entail each enzyme functioning in concert by channelling

indole between respective catalytic domains. This would imitate the catalytic process obsereved in a typical tryptophan synthase complex with an intramolecular tunnel connecting the α and β active sites (Anderson *et al*, 1991). Although a TSα sub-unit cloned from *A. thaliana* which is phylogenetically related to the maize IS could not function when co-expressed with an *E. coli* TSβ (Radwanski *et al*, 1995), the formation of a functional maize IS/TSβ sub-unit complex cannot be ruled out. Independent sub-unit interaction of tryptophan biosynthetic pathway enzymes is undefined thus both indole synthase and tryptophanase must be considered.

The creation of an indole synthase/indole oxygenase sequential enzyme system was performed to assess the potential generatation of indigo in *E. coli* W3110*trp*A33 using only glucose as the starting substrate. The elimination of tryptophanase as a means for indole production was necessary to accurately assess the indigo producing capabilities of the IS/INO enzyme system. If the tryptophanase enzyme of *E. coli* W3110*trp*A33 remained functional, indigo synthesis from precursor indole could not be definitively attributed to either IS activity or the degradation of tryptophan to indole via tryptophanase.

When indole synthase was co-expressed with the *Rhodococcus* INO in the constructed double mutant (W3110*trp*A33*tna*A), complementation occurred, however, no apparent indigo accumulation was observed. From the results obtained, it is evident that the maize IS and an *E. coli* TSβ sub-unit appear to form a functional enzyme complex. It was also deduced that 2.0 mM indole supplementation permitted indigo synthesis. This confirms that although the indole synthase can catalyse the synthesis of indole, either it cannot produce indole in a form that is utilised by the INO enzyme for indigo synthesis, or all of the indole generated by IS activity is being utilised for tryptophan biosynthesis. Exogenous indole is required for indigo biosynthesis.

This data signifies the ability of the maize indole synthase to emulate a typical TSα sub-unit. These results are significant considering the fact that in maize, this IS enzyme catalyses the synthesis of indole that is not destined for the tryptophan biosynthetic pathway. If this enzyme is able to perform the *trp*A function in *E. coli* then perhaps phylogenetically related TSα enzymes or putative indole synthases in indigo producing plants have the same enzymatic properties as the maize indole synthase and are able to synthesise 'free' indole. The identification of such enzymes in plants may allow further elucidation of the plant indigo pathway.

4 Chapter Four: The pursuit of an indole synthase in indigo producing plants

4.1 Introduction

The discovery of a maize indole synthase, capable of performing the *trp*A function in *E. coli* in the absence of any plant TSβ sub-unit (Kramer and Koziel, 1995) has led to further investigation as to whether there are other similar enzymes within plants. This maize IS is capable of catalysing the production of indole that is not confined to the tryptophan biosynthetic pathway. As previously discussed in Section 1.4.1, this 'TSα like' enzyme was renamed Indole Synthase (IS) after studies of the mutant *BxBx* maize line confirmed it as a key enzyme withinin the DIMBOA defense pathway. The synthesis of DIMBOA is dependant on the presence of vast amounts of precursor indole (Melanson *et al*, 1997; Frey *et al*, 2000).

TSα enzymes have been described as 'formal indole synthases' (Melanson *et al*, 1997) due to their ability to cleave I3GP and form indole, albeit dedicated to tryptophan biosynthesis. TSα's from a variety of organisms share amino acid residues with the maize IS that are essential to catalytic activity (Salvini *et al*, 2008). If TSα's and the maize IS display homology in the conserved regions, then perhaps they possess the same enzymatic characteristics, specifically the ability to generate 'unbound' indole. A cloned TSα sub-unit from *I. tinctoria* has recently been reported to possess the same enzymatic properties as the maize IS, the ability to catalyse reactions yielding indole without binding the TSβ subunit (Salvini *et al*, 2008). Analysis at the amino acid level has determined differences in amino acid residues within the catalytic domain between the *I. tinctoria* TSα and the maize indole synthase that directly affect α/β sub-unit association (Salvini *et al*, 2008). The replacement of two glutamic acid residues in the maize IS active site appears to have been the catalyst for independant sub-unit function and is likely to have occurred through evolutionary sequence divergence of the gene encoding TSα activity (*trp*A) (Kulik *et al*, 2005). Considering

sequence analysis of *I. tinctoria* TSα has identified these differences in amino acid substitutes within the domain responsible for β subunit interaction compared with the maize IS (Salvini *et al*, 2008), further investigation into whether the TSα from *I. tinctoria* can actually function in the absence of its TSβ sub-unit counterpart is necessary.

Although little information regarding the TS complex from *P. tinctorium* is available, the prominent indigo producing capability of the plant and the previously described conservation between TSα subunits implies that a TSα subunit from *P. tinctorium* may display the same tendencies as the characterised TSα from *I. tinctoria*, making the identification of the gene controlling such enzymatic activity of great interest. This chapter will pursue the isolation of a *P. tinctorium* TSα as well as the characterised TSα from *I. tinctoria* and investigate whether these specific genes can emulate the maize IS and complement an *E. coli trp*A auxotroph. The *trp*A cDNA (IS) characterised by Kramer and Koziel will act as a positive control for complementation experiments.

When the genes and pathway intermediates vital to tryptophan biosynthesis within *A.thaliana* were characterised and associated mutants were identified in the mid 1990's, a wealth of information was added to the scarce database regarding plant tryptophan biosynthesis. In contrast to the IS from maize that was able to perform independently of the maize TSβ sub-unit, Radwanski and Last, (1995) isolated a *trp*A gene from *A. thaliana* encoding a TSα sub-unit (TSα1) that did not function in the absence of the *A. thaliana* TSβ sub-unit. This gene will also be isolated and will act as the negative control for complementation experiments with the *trp*A auxotroph. Using BLASTx analysis (NCBI website) the only other copy of a *trp*A gene in *A. thaliana* is a 'TSα like' gene showing 72% homology at amino acid level to the TSα1 discovered by Radwanski and Last (1995). This gene was also isolated and used in complementation experiments to assess whether it can function as the IS from maize does when expressed in an *E. coli trp*A mutant.

The pursuit of an indole synthase within indigo producing plants is fundamental to identifying how these plants can generate large amounts of the blue pigment. Due to the inferior vibrancy of indigo dye from *I. tinctoria*, the search for an indole synthase like enzyme within *P. tinctorium* was preferred. The construction of a *P. tinctorium* cDNA library was performed with the aim to identify an indole synthase or indole synthase like gene within the library. In screening the library, the cloned maize indole synthase gene was used a probe to ascertain any potential indole synthase homologs at the nucleotide level.

The maize enzyme, indole-3-glycerol phosphate lyase (IGL) encoded by the *IGL* gene, displays the same enzymatic properties as IS, in that it also catalyses the conversion of I3GP to 'free' indole in maize. This IGL enzyme is able to perform the *trp*A function in *E. coli* (Frey *et al*, 2000) however, IGL is ten times less efficient at catalysing the cleavage of I3GP and the indole that is generated is of a volatile and odorous nature, designed to be released later in development and only in response to pathogen attack (Frey *et al*, 2000). The maize IS is fundamental in the synthesis of indole for accumulation of DIMBOA in germinating seeds, therefore the maize IS is the focus of this thesis.

A functional analysis of the cDNA library was also be performed to search for a clone capable of performing the *trp*A function in *E. coli* and thus rescuing a *trp*A auxotroph. This area of research is devoted to identifying enzymes with the potential to generate vast amounts of indole committed to indigo synthesis and not for tryptophan biosynthesis in indigo producing plants.

4.2 Materials and Methods

4.2.1 Seeds and growth conditions

A. thaliana seeds were obtained from the Department of Primary Industries (DPI), Victoria, Australia. *I. tinctoria* seeds were obtained from Plant Hunters Nursery, Tasmania, Australia. *P. tinctorium* seeds were obtained from Companion Plants, Athens, Ohio, U.S.A. Seeds were surfaced sterilised using 70% EtOH and exported to RMIT, Melbourne, Australia. *Z. mays* seeds were obtained from Greensborough Grain and Feed Store, Melbourne, Australia. *A. thaliana* seeds were planted in seed raising mix and stored at 4°C for 2 days. All other seeds were planted in seed raising mix and transferred to a temperature controlled growth room. Seeds were cultivated at 20°C ± 2.5°C with a 16 h light cycle. Leaf tissue was excised and snap frozen in liquid nitrogen. The plant tissue was then stored at -80°C.

4.2.2 Amplification of target genes

PCR was pursued to amplify genes of interest, the *Bx1* gene encoding the IS from *Z. mays*, and the *trp*A genes from *I. tinctoria*, *P. tinctorium* and *A. thaliana* encoding the respective TSα sub-units. RNA was extracted from *Z. mays*, *I. tinctoria*, *P. tinctorium* and *A. thaliana* leaf tissue via small scale RNA extraction (Section 2.4.2.1) and converted to cDNA via MMLV reverse transcriptase activity. Amplification of target genes was performed using respective cDNA extracts as templates. Published sequences obtained from the NCBI website (National Centre for Biotechnology Information) were used to design primers that amplify the IS gene (*Bx1*) from maize and the respective *trp*A genes from *I. tinctoria* and *A. thaliana.* Due to the conservative nature of TSα's, the *I. tinctoria* TSα primers were also used on *P. tinctorium* cDNA. All primers used for initial amplification of target genes were designed with a *Bam*H1 restriction site embedded in the sequence at the 5' end of the forward primer, while a *Hind*III restriction site was embedded in the sequence at the 5' end of the reverse primer. Primers were designed at the start of the mature peptide and

thus the chloroplast transit peptide that was in the initial translation product was omitted from PCR amplification.

The PCR conditions are outlined in Section 2.4.7, while the annealing temperature for each gene of interest is dictated by the melting temperature of each primer as detailed in Table 2.3. Cycling conditions are listed in Table 2.4. (NB: Primers designed for amplification of *I. tinctoria* TSα were used for the amplification of the *P. tinctorium* TSα, successfully amplifying a PCR product).

4.2.3 Verification of PCR amplification

PCR products were viewed by transilluminating agarose gels with UV light after gel electrophoresis detailed in Section 2.4.5. Purification of PCR products (Section 2.6) was performed prior to ligation into restriction enzyme digested plasmid.

4.2.4 Plasmid construction

PCR products were purified, digested with *Bam*H1 and *Hind*III and ligated into similarly digested cloning vector *pUC18*. After ligation, electrocompetent DH5α cells were transformed with each construct via electroporation and resultant cultures were plated on LB^{AMP} and left to incubate at 37°C overnight. Supplementation with IPTG and X-gal permitted analysis by blue/white screening. A single white colony was selected for each gene transformant and 5 mL cultures were prepared with liquid LB^{AMP} then incubated overnight at 37°C. Plasmid constructs were analysed via restriction enzyme digestion (Section 2.5.2) and viewed on an agarose gel (Section 2.4.5). Cloned genes were sequenced at the AgGenomics sequencing facility at the Victorian AgriBiosciences Centre.

4.2.5 Complementation of *E. coli* W3110trpA33 with genes of interest

E. coli W3110trpA33 strain #7977 from the *E. coli* Genetic Stock Centre, Yale University, USA containing chromosomal markers, *LAM*-, *trp*A33, *IN*(*rrnD-rrnE*) and *rph*-1 was transformed with each cloned gene and with *pUC18* as a control measure. The method in which transformation was performed is detailed in Section 2.5.4. All strains were grown on M9 minimal media with and without tryptophan supplementation. Further selection was pursued via the addition of ampicillin (100 µg/mL) corresponding with the Amp resistance marker on the *pUC18* cloning vector. Successful transformants were determined by the ability to grow on M9 minimal media without any supplemented tryptophan.

4.2.6 Large scale RNA extraction

Exactly 0.574 g of frozen *P. tinctorium* leaf tissue was ground in liquid nitrogen to which 5 mL of RNA extraction buffer [Guanidium thiocyanate, ddH_20, 1M tri-sodium Citrate, 20% Sarcosine β-Mercaptoethanol (v/v)] was added followed by immediate homogenisation with a polytron homogeniser. The RNA slurry was transferred to a 15mL Flacon tube and left to rotate overnight at room temperature on an orbital shaker. Samples were centrifuged at 9,000 rpm for 15 mins and filtered through miracloth to collect the supernatant. Caesium chloride was added to the supernatant (0.2 g/mL) and dissolved, 0.8 mL cushion solution (5.7M CsCl, 0.1M EDTA) was then added to a Polyallomer Bell-Top Quickseal centrifuge tube to which 2.5 mL of RNA extract was carefully layered on top with a needle and syringe. Samples were heat sealed within the tubes using a Beckman Sealer and placed in a TLA100.3 rotor and spun at 100,000 rpm for 2 h at 25°C in a Beckman TL-100 Ultracentrifuge. After centrifugation, the top of the tubes were severed with scissors and the top layer of the slurry was extracted with a glass Pasteur pipette.

The tube was then cut in half and the clear phase was decanted off the RNA pellet at the bottom of the tube, 100 µL of RNA resuspension solution [2mL STE-TES, 40µL 10% SDS

(w/v), 20µL 20mg/mL Proteinase K] was added to dissolve the pellet which was then incubated at 37°C for 5 mins. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added to the extract and spun for 5 mins at 14,000 rpm. The top layer was then removed and transferred to a new 1.5 mL microcentrifuge tube to which an equal volume of chloroform was added. Once again the extract was spun at 14,000 rpm for 5 mins after which the clear phase was supplemented with 1/10 volume of 3 M Sodium acetate and 2.5 volumes of 95% EtOH (v/v). Samples were stored at -80°C.

4.2.6.1 Construction of the cDNA library

A cDNA library for *P. tinctorium* was constructed using the Creator™ SMART™ cDNA Library Construction Kit from Clontech. The protocol for cDNA library construction is detailed as per the manufacturers instructions (Creator™ SMART™ cDNA Library Construction Kit user manual). Total RNA was required for initiation of cDNA library construction. All primers used for cDNA amplification are detailed in Table 2.3 (Section 2.4.7).

4.2.7 Screening the cDNA library

Once the amplified library was titred, approximately 500,000 independent clones were screened with the aid of *E. coli* strain XL1-Blue. Plaque forming independent clones or Plaque forming units (Pfus) were plated as per the manufacturers instructions (Creator™ SMART™ cDNA Library Construction Kit user manual). Plaques were transferred to nylon membrane (ROCHE) as per the manufacturer's instructions (Nylon membranes for Colony and Plaque Hybridisation). Membranes were stored at room temperature until further use. Primary and secondary screens were performed on 140 mm petri dishes, while tertiary screens were performed on 90 mm petri dishes. Plaques were excised from agar and eluted in suspension media at 4°C in preparation for further screening or phage DNA conversion.

4.2.7.1 Labelling the probe

Purified *Z. mays* indole synthase PCR product was labelled with dCTP[α^{32}P] (Perkin Elmer) as per the manufacturers instructions (DECAprime Random Primed DNA Labelling Kit, Ambion, U.S.A.). Purified *P. tinctorium* β-glucosidase PCR product was also labelled and used as a positive control.

4.2.7.2 Pre-hybridisation/hybridisation

Membranes were subject to pre-hybridisation in 10% NEN (v/v) at 42°C for at least 30 mins in a rotating hybridisation oven. 10,000,000 DPM of labelled probe was heated for 5 mins at boiling point and chilled on ice, then was added to 10 mL of the pre warmed hybridisation solution. Membranes were then subject to hybridisation at 42°C overnight in a rotating hybridisation oven.

4.2.7.3 Stringency washes

After hybridisation, the probe/hybridisation solution was decanted and membranes were subject to stringency washes. Membranes were subject to two 20 mins low stringency washes in 2 X SSC + 0.1% SDS (w/v) buffer at 65°C. Membranes were then subject to two 20 min high stringency washes in 0.2 X SSC + 0.1% SDS (w/v) or 0.1 X SSC + 0.1% SDS (w/v) at 60°C or 65°C. Membranes were stored in 2 X SSC to avoid drying out.

4.2.7.4 Exposure and development of X-ray film

Membranes were exposed to X-ray film in complete darkness at -80°C for a period of no less than 16 h. X ray film was then developed using X-ray grade developer and fixer in a darkroom.

4.2.7.5 Analysis of cDNA inserts

Conversion of *λTriplEx2* to *pTriplEx2* was performed to analyse the cDNA inserts. Eluted positive plaques were subject to conversion protocol using *E. coli* strain BM25.8 as per the manufacturers instructions (Creator™ SMART™ cDNA Library Construction Kit user manual). Converted clones were sequenced at the AgGenomics sequencing facility at the Victorian AgriBiosciences Centre. The sequencing primers detailed in Table 2.3 (5' seq and 3' seq) were used when screening and sequencing the inserts.

4.2.7.6 Southern blot analysis

Plasmid extracts (*pTriplEx2* containing inserts) were digested with *Sfi*1 to excise inserts. Products were subject to electrophoresis on a 0.7% TAE gel (w/v) and viewed via transillumination. Transfer of DNA to a nylon membrane was performed using the iBlot™ Gel Transfer System as per the manufacturers instructions (Invirogen™). Membranes were subject to analysis as detailed in Sections 3.2.7.1-3.2.7.5

4.2.7.7 Functional analysis of the cDNA library

To assess if the cDNA library contained any clones able to complement *E. coli* W3110*trp*A33, 100,000 plaque forming units were converted to plasmids containing cDNA inserts. This was done by bulk rescuing the plasmids and pooling them creating a 1 mL culture, this was performed as per the manufacturers instructions (CLONTECH®). A 50 mL bacterial culture was prepared by adding the pooled plasmid culture to 50 mL of liquid LB supplemented with ampicillin (100 µg/mL) and cultivating overnight at 37°C. This was followed by subsequent midi-prep DNA extraction (QIAGEN®) from the culture, DNA was eluted in 100 µL of sterile water. This multi-plasmid DNA was used to transform *E. coli* W3110*trp*A33*tna*A for subsequent plating onto M9 minimal media^{amp} (-trp) to assess if the cDNA library could rescue the strain through the expression of a gene able to complement the *trp*A33 mutation.

4.3 Results

4.3.1 Amplification of target genes

PCR products were purified by gel excision (QIAGEN) and subject to gel electrophoresis (Figure 4.1). Indole synthase from *Z. mays* and the genes encoding TSα from *I. tinctoria*, TSα from *P. tinctorium*, TSα1 from *A. thaliana* and TSα2 from *A. thaliana* displaying sizes of 762, 822, 822, 822 and 765 bp pairs respectively. These results are consistent with the sequences obtained from the NCBI gene database. After purification of the PCR products, quantification of DNA by absorbance readings at 260nm and analysis via gel electrophoresis confirmed each PCR product to be approximately 30-40ng/µL.

4.3.2 Synthesis of plasmid constructs

Purified PCR generated fragments were digested with restriction enzymes *Bam*H1 and *Hind*III and ligated into the similarly digested cloning vector *pUC18* in preparation for sequence analysis and complementation experiments. Plasmid DNA was extracted from cultures using small scale plasmid DNA extraction (section 2.4.1.1) and the presence of the insert was screened for by restriction enzyme digestion. Each newly synthesised construct was renamed *pISP* and designated a number to identify the insert (Figure 4.2). Plasmid maps were also constructed (Appendix).

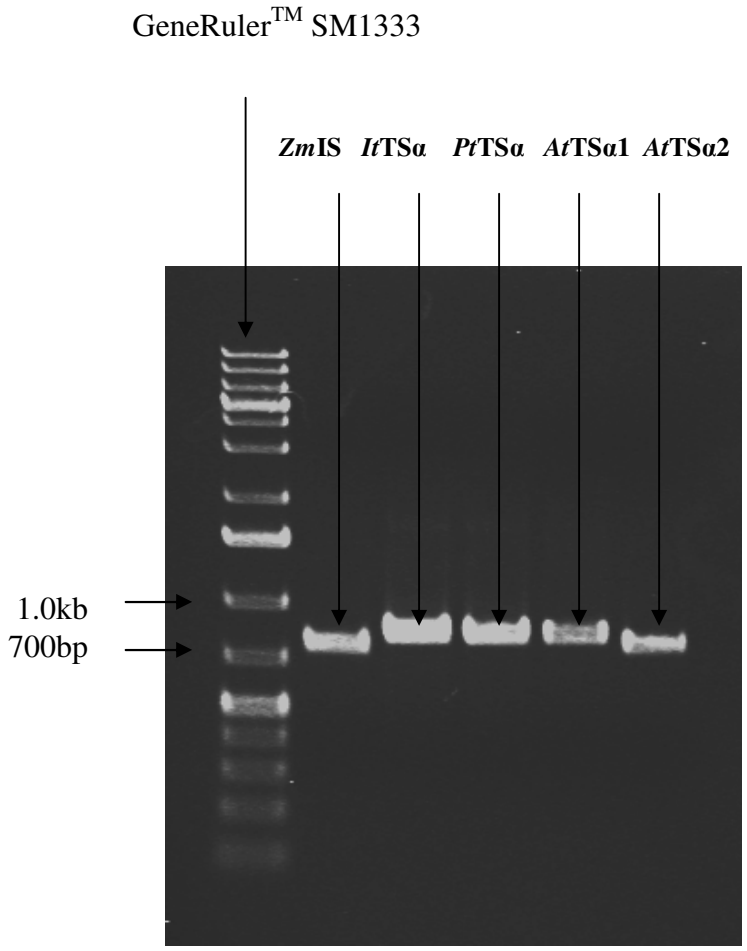


Figure 4.1 Gel electrophoresis of purified PCR products. Lane 2: *Z. mays* IS (*Bx1*) (762bp), lane 3: *I. tinctoria* TSα (*trp*A) (822bp), lane 4: *P. tinctorium* TSα (*trp*A) (822bp), lane 5: *A. thaliana* TSα1 (*trp*A) (822bp), lane 6: *A. thaliana* TSα2 (*trp*A) (765bp) (10µL of each purified PCR product loaded)

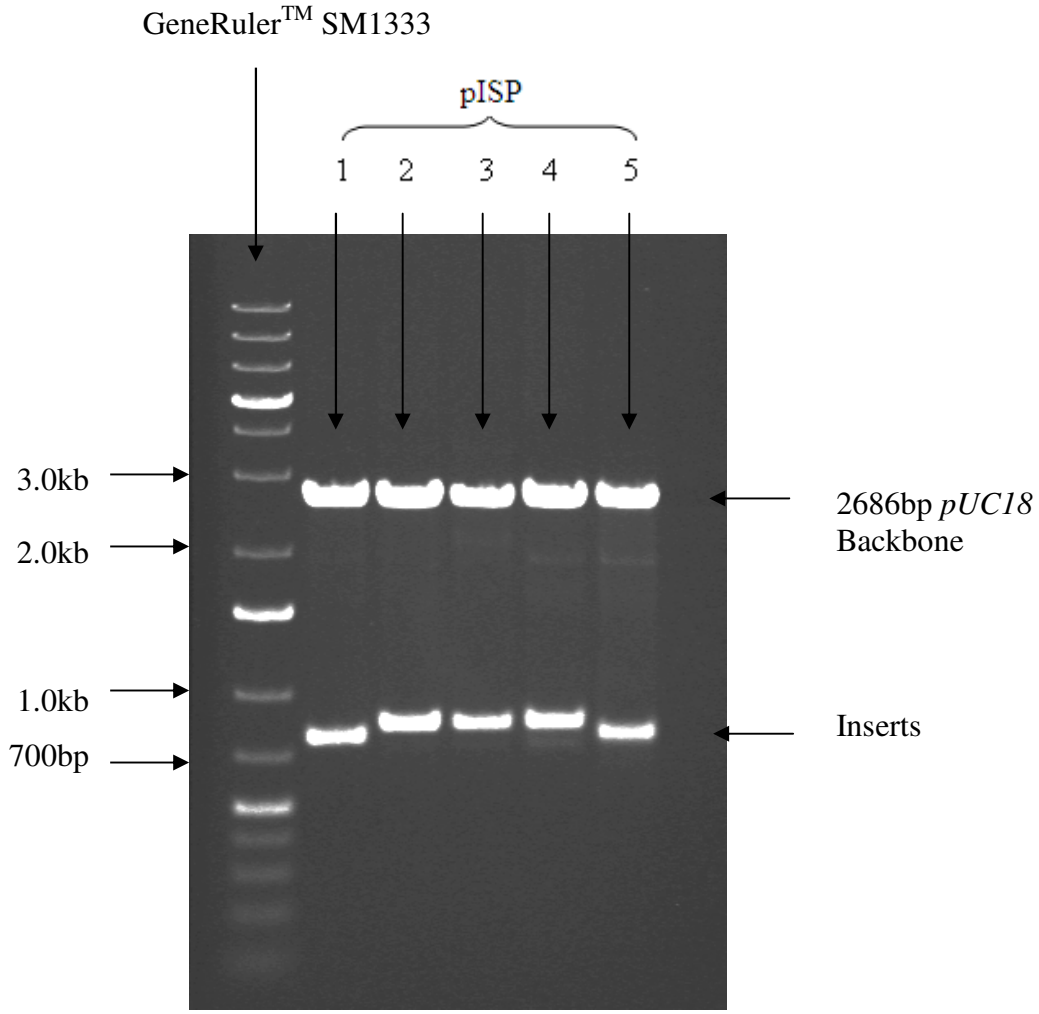


Figure 4.2 Restriction enzyme digestion of plasmid constructs. Lane 2: *Z. mays* IS/*pUC18* renamed *pISP1*, Lane 3: *I. tinctoria* TSα/*pUC18* renamed *pISP2*, Lane 4: *P. tinctorium* TSα/*pUC18* renamed *pISP3*, Lane 5: *A. thaliana* TSα1/*pUC18* renamed *pISP4*, Lane 6: *A. thaliana* TSα2/*pUC18* renamed *pISP5*. NB: Inserts are the same size denoted in Figure 4.1.

4.3.3 Sequence analysis

Multiple sequence alignment of nucleotides from the sequenced inserts was performed. A similarity matrix table (Table 4.1) displays the percentage identity between nucleotides resulting from the alignment. Multiple sequence alignment of the amino acid residues from the sequenced inserts showed that identities between the *Z. mays* indole synthase and *I. tinctoria* TSα, *P. tinctorium* TSα, *A. thaliana* TSα1 and *A. thaliana* TSα2 were 56.2%, 56.3%, 56.2% and 58.8% respectively (Table 4.2). Amino acid identity between *I. tinctoria* TSα and *P. tinctorium* TSα is 98.3%, whilst the TSα from *I. tinctoria* and the TSα1 from *A. thaliana* share 93.6% amino acid residues. The TSα2 from *A. thaliana* is 68.0% identical to both *I. tinctoria* and *P. tinctorium* TSα and only 71.2% identical to the TSα1 from *A. thaliana* indicating that perhaps the TSα2 is a 'TSα like' enzyme.

Although the sequence alignment displayed a similarity of less than 60% at amino acid level between *Z. mays* indole synthase and the respective *trp*A genes, the amino acid residues required for TSα catalytic activity previously identified (Miles, 1991; Radwanski *et al*, 1995) are conserved between each protein (Figure 4.3). The conserved amino acids are glutamic acid, aspartic acid, tyrosine, threonine, three glycine residues and a serine at positions 41, 52, 125, 126, 166, 174, 202, 204, 225 and 226 for *Z. mays* indole synthase. These essential residues are located at positions 57, 68, 141, 142, 182, 190, 218, 220, 241 and 242 for *I. tinctoria* TSα, *P. tinctorium* TSα and *A. thaliana* TSα1 (Figure 4.3). Although being less than 72% identical at the amino acid level, the AtTSα2 also contains the necessary residues for TSα catalytic activity at positions 38, 49, 122, 123, 163, 171, 199, 201, 222 and 223 (Figure 4.3). Multiple sequence alignment of the amino acids also displayed the absence of two specific glutamic acid residues in the indole synthase, present in the TSα genes. These glutamic acid residues reside in the domain responsible for interaction with the TSβ sub-unit (Salvini *et al*, 2008).

Table 4.1 Similarity matrix displaying the identity (%) between sequenced inserts at nucleotide level

	ZmIS	**ItTSα**	**PtTSα**	**AtTSα1**	**AtTSα2**
ZmIS (***Bx1***)	-	58.7	59.3	58.2	60.7
ItTSα (***trp*A**)	-	-	99.0	92.5	68.7
PtTSα (***trp*A**)	-	-	-	92.4	68.8
AtTSα1 (***trp*A**)	-	-	-	-	68.4
AtTSα2 (***trp*A**)	-	-	-	-	-

Table 4.2 Similarity matrix displaying the identity (%) between sequenced inserts at amino acid level

	ZmIS	**ItTSα**	**PtTSα**	**AtTSα1**	**AtTSα2**
ZmIS	-	56.2	56.3	56.2	58.8
ItTSα	-	-	98.3	93.6	68.0
PtTSα	-	-	-	93.8	68.0
AtTSα1	-	-	-	-	71.2
AtTSα2	-	-	-	-	-

```
                                                                                  ■
ZmIS     1 ----------------AALMAKGKTAFIPYITAGDPDLATTAEALRLLDGCGADVIELGV
ItTSα      AALSTSSPTLGLADTFKE.KKQ..V.............S......KV.AAS.S.I.....
PtTSα      AALSTSSPTLGLADTFKE.KKQ..V.............S......KV..A..S.I.....
AtTSα1     ASLSTSSPTLGLADTFTQ.KKQ..V.............S......KV..A..S.I.....
AtTSα2     .................----SQ..V.L...........S...K..KV..S..S.I.....

                  ■
ZmIS    61 PCSDPYIDGPIIQASVARALASGTTMDAVLEMLREVTPELSCPVVLLSYYKPIMSR----
ItTSα      .Y...LA...V...AAT...EN..NL.NI.D..DK.L.QV....S.FT..N..LK.GLGK
PtTSα      .Y...LA...V...AAT.S.EN..NL.NI.D..DK.L.QV....S.FT..N..LK.GLGK
AtTSα1     .Y...LA...V...AAT.S.ER..NL.SI....DK.V.QI...IS.FT..N..LK.GLGK
AtTSα2     .Y...LA...A...AAR.S.LK..NFNSIIS..K..I.Q....IA.FT..N..LR.GVEN

ZmIS   121 SLAEMKEAGVHGLIVPDLPYVAAHSLWSEAKNNNLELVLLTTPAIPEDRMKEITKASEGF
ItTSα      FMSSIRDV..Q..V...V.LEETEM.RK..L...I........TT.TE...R.VD.....
PtTSα      FMSSIRDV..Q..VE..V.LEETEM.RK..L...I........TT.TE...R.VD.....
AtTSα1     FMSSIRAV..Q..V...V.LEETEM.RK..L..DI........TT.TE...L.VD.....
AtTSα2     YMTVI.N......L...V.LEETET.RN..RKHQI........TT.KE..NA.VE.....

           ■        ■                              ■ ■
ZmIS   181 VYLVSVNGVTGPRANVNPRVESLIQEVKKVTNKPVAVGFGISKPEHVKQIAQWGADGVII
ItTSα      I....SI....A..S.SGK.Q..LKDI.EA.D..........Q........G.......V
PtTSα      I....SI....A..S.SGK.Q..LKDI.EA.D..........Q........G.......V
AtTSα1     I....SI....A.SS.SGK.Q..LKDI.EA.D....................G.......V
AtTSα2     I....SV....T.ES..EK.Q..L.QI.EA.S..................V.E.......V

           ■■
ZmIS   241 GSAMVRQLGEAASPKQGLRRLEEYARGMKNALP-
ItTSα      ......L..D.K..TE..KE..RLTKSL.S..L.
PtTSα      ......L..D.K..TE..KE..RLTKSL.S..L.
AtTSα1     .....KL..D.K..TE..KE..KLTKSL.S..L.
AtTSα2     .....KI...SE..E...KE..FFTKSL.S..VS
```

Figure 4.3 Alignment of mature peptides from *Z. mays*, *I. tinctoria*, *P. tinctorium* and *A. thaliana*. Amino acid sequences have been translated from the sequenced inserts generated by PCR amplification. Dots represent conserved amino acids for all sequences. Numbering is relative to the *I. tinctoria* TSα, *P. tinctorium* TSα and *A. thaliana* TSα1. Bold black squares above the residues represent amino acids essential to catalytic activity (Radwanski *et al*, 1995). The yellow highlighted region indicates two glutamic acid residues residing in the domain responsible for interaction with the TSβ sub-unit. In the indole synthase these amino acids are replaced by a valine and alanine (highlighted in green). Gaps that maximise alignment are represented by a hyphen.

The highly conservative nature of the TSα sub-unit in plants is reflected in Figure 4.4, displaying at least a 92% similarity between all cloned inserts at amino acid level.

4.3.4 Complementation of *E. coli* W3110*trp*A33 with cloned genes

Each cloned gene was used to transform the *trp*A auxotroph, W3110*trp*A33. Colonies transformed with the indole synthase from *Z. mays* had the ability to grow on M9 minimal media without tryptophan (Figure 4.5). Transformants containing only the *pUC18* plasmid were unable to grow in the absence of added tryptophan. Transformants containing the TSα's from *I. tinctoria* and *P. tinctorium* or the TSα1 and the TSα2 from *A. thaliana* were also unable to grow without tryptophan supplementation (Figure 4.6). The complementation of W3110*trp*A33 with the indole synthase indicates that this gene encodes the *trp*A function in *E. coli*. Non-complementation of the *trp*A auxotroph by clones isolated from the cDNA library indicates an inability to emulate the maize indole synthase and function with the *E. coli* TSβ sub-unit to synthesise tryptophan. These results indicate the failure of the TSα's from *I. tinctoria*, *P. tinctorium* and *A. thaliana* to function in the absence of their plant TSβ sub-unit counterpart.

4.3.5 Screening the cDNA library

4.3.5.1 Titering the cDNA library

The *P. tinctorium* cDNA library was amplified as per the manufacturer's instructions (Clontech) to determine the total pfu in the library lysate. A total of 240mL of lysate was generated. After packaging, recombination efficiency was calculated to be > 80%. The total number of independent clones in the library was calculated to be 9.9 X 10^{11} pfu/μL of lysate.

```
ItTSα   1  AALSTSSPTLGLADTFKELKKQGKVAFIPYITAGDPDLSTTAEALKVLAASGSDIIELGV
PtTSα      ..................................................D.C.........
AtTSα1     .S...............TQ...............................D.C.........

ItTSα  61  PYSDPLADGPVIQAAATRALENGTNLDNILDMLDKVLPQVSCPVSLFTYYNPILKRGLGK
PtTSα      ...................S..............................................
AtTSα1     ...................S..R.....S..E.....V..I...I................

ItTSα 121  FMSSIRDVGVQGLVVPDVPLEETEMLRKEALNNNIELVLLTTPTTPTERMKRIVDASEGF
PtTSα      ...............E...............................................
AtTSα1     ......A..............................D..................L........

ItTSα 181  IYLVSSIGVTGARASVSGKVQSLLKDIKEATDKPVAVGFGISQPEHVKQIAGWGADGVIV
PtTSα      ............................................................
AtTSα1     .............S..............................K..................

ItTSα 241  GSAMVRLLGDAKSPTEGLKELERLTKSLKSALL
PtTSα      ...............................
AtTSα1     .....K................K..........
```

Figure 4.4 Multiple sequence alignment of amino acids between between *I. tinctoria*, *P. tinctorium* and *A. thaliana* TSα's. Amino acid sequences have been translated from the sequenced inserts generated by PCR amplification. Dots represent conserved amino acids whilst bold black squares above the residues represent amino acids required for TSα catalytic activity (Radwanski *et al*, 1995). The yellow highlighted region indicates two glutamic acid residues that reside in the domain responsible for interaction with the TSβ sub-unit (Salvini *et al*, 2008).

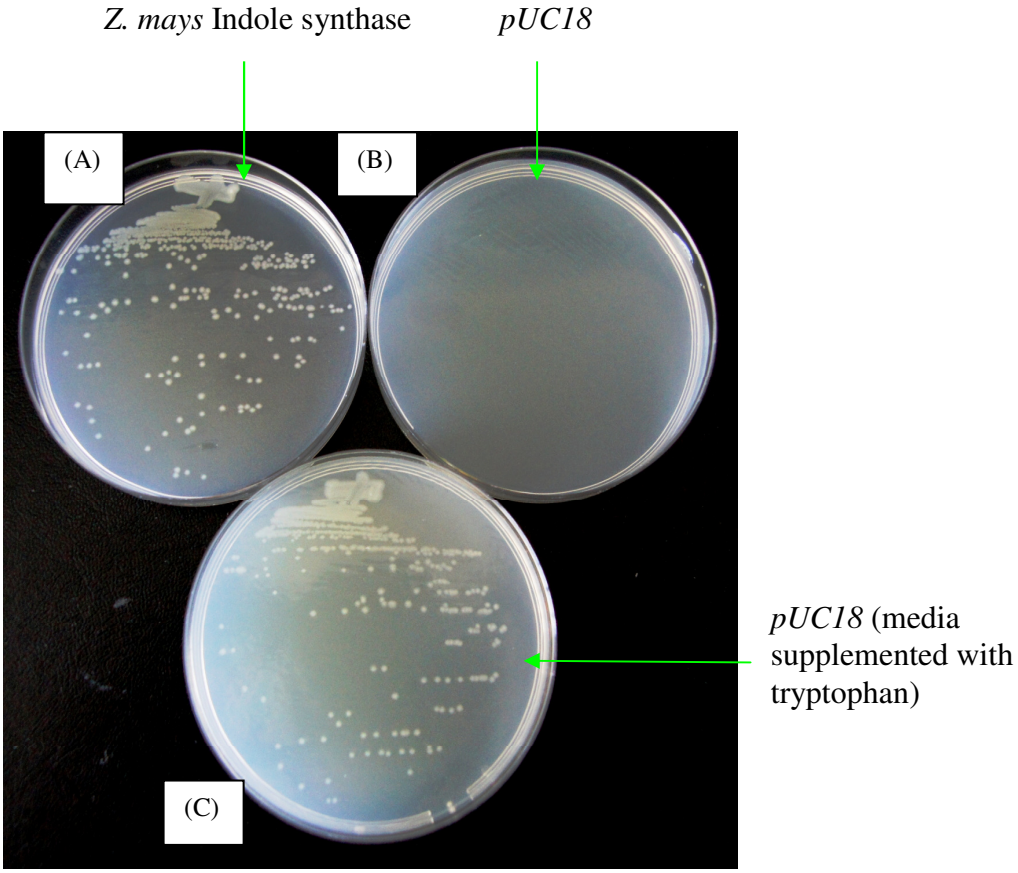


Figure 4.5 (A) Complementation of *E. coli* W3110*trp*A33 with *Z. mays* indole synthase. (B) Negative control: transformation of *E. coli* W3110*trp*A33 with cloning vector *pUC18*. (C) Positive control: transformation of *E. coli* W3110*trp*A33 with *pUC18* (M9 minimal media is supplemented with tryptophan).

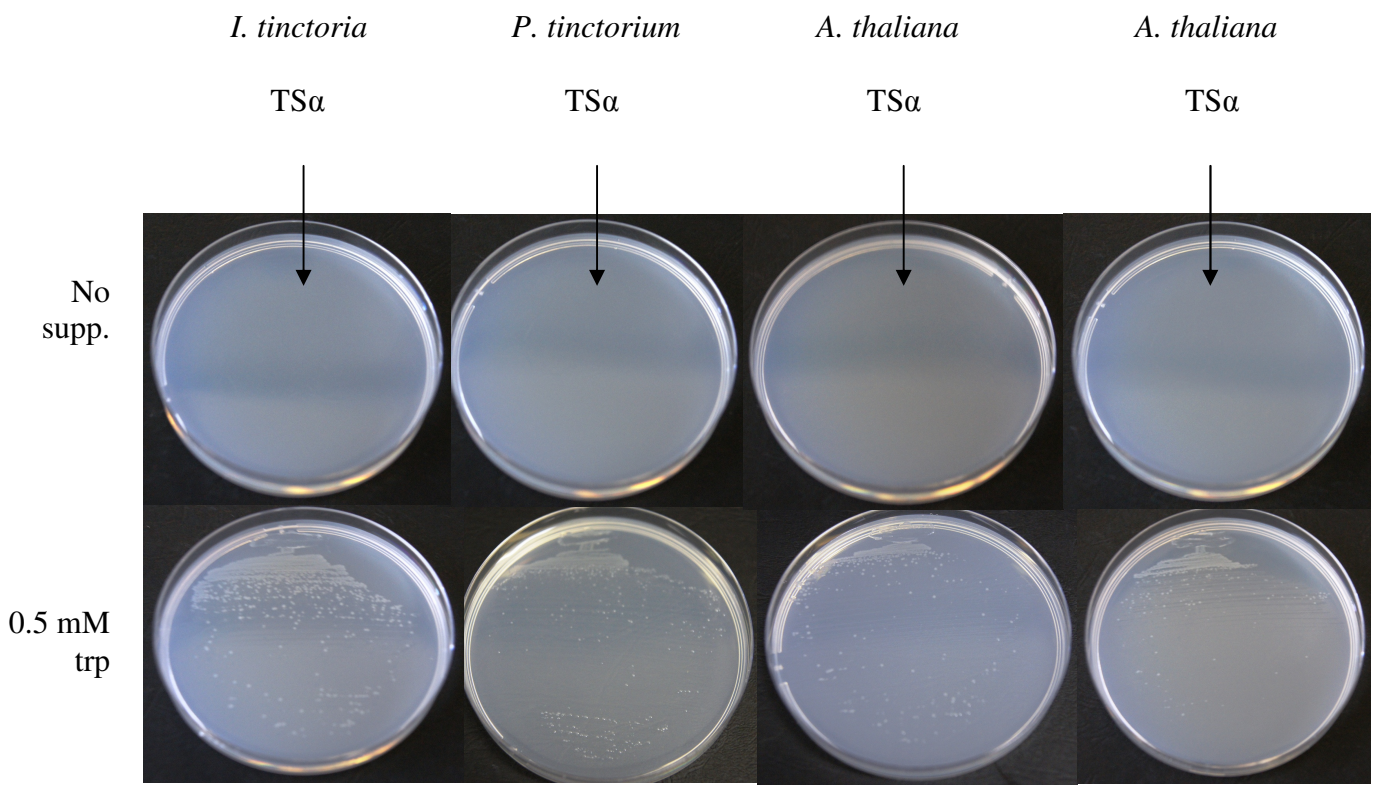


Figure 4.6 Complementation of W3110*trp*A33 with the cloned *trp*A genes from *I. tinctoria, P. tinctorium* and *A. thaliana*.

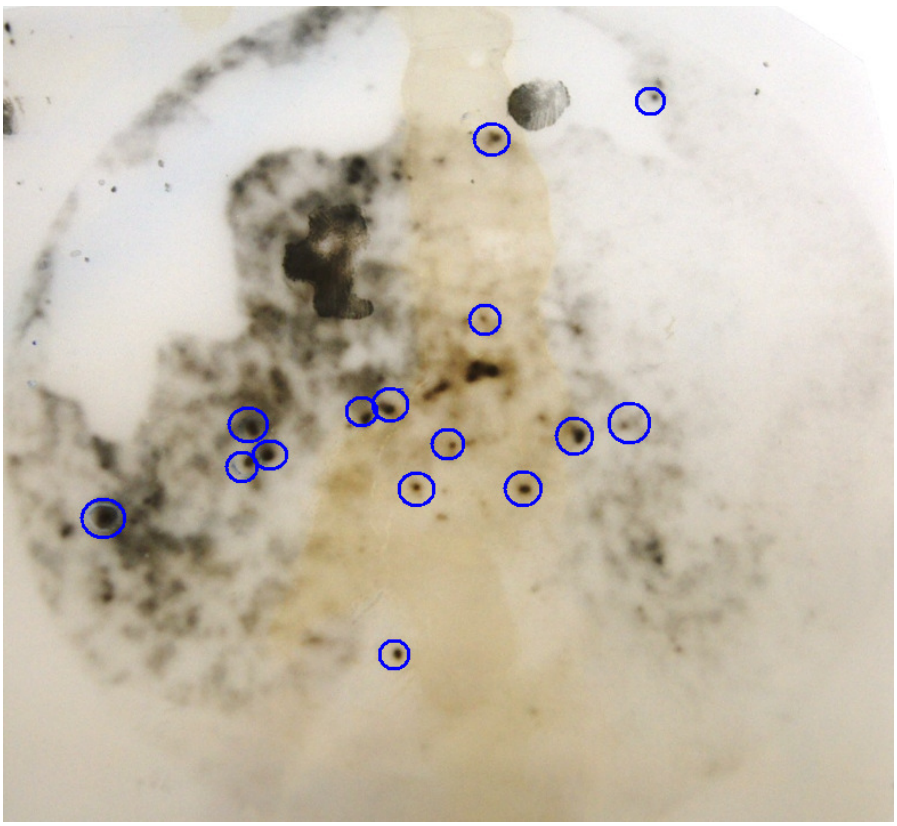
4.3.5.2 cDNA library positive control

Using the calculated figure for pfu/µL of lysate, approximately 10,000 independent clones were screened using a labelled *P. tinctorium* β-glucosidase PCR product as a probe. Plaques were grown on 5 separate 140mm petri dishes (approximately 2500 plaques/plate) and transferred to nylon membranes. After probing and dual low stringency washes, membranes were subject to dual high stringency washes with 0.1 X SSC + 0.1% SDS (w/v) at 65°C. After exposure (Figure 4.7), the X-ray films displayed 15 positives that aligned to plaques on the corresponding agar plate and matching plaques were excised and eluted in suspension media then converted from phage DNA to plasmid DNA (*pTriplEx2* containing insert). These phage converts were used as a template for amplification of cDNA inserts (Figure 4.7) using the 5' and 3' sequencing primers provided by the kit (Clontech). Inserts were purified and sent for sequencing. Sequencing results determined that all inserts screened except for sample 14 corresponded to the published *P. tinctorium β*-glucosidase gene using the BLAST tool on the NCBI website (Appendix).

4.3.5.3 Probing with Z. mays IS

Using the calculated figure for pfu/µL of lysate, approximately 500,000 independent clones were screened at a time using the labelled maize IS PCR product as a probe. Plaques were grown on 8 separate 140mm petri dishes (approx. 62,500 plaques/plate) and transferred to nylon membranes. After probing and dual low stringency washes, membranes were subject to dual high stringency washes with 0.2 X SSC + 0.1% SDS (w/v) at 60°C. After exposure (Figure 4.8), the X-ray films displaying positives were aligned with corresponding agar plates and matching plaques were excised and eluted in suspension media in preparation for secondary screening.

(A)



(B)

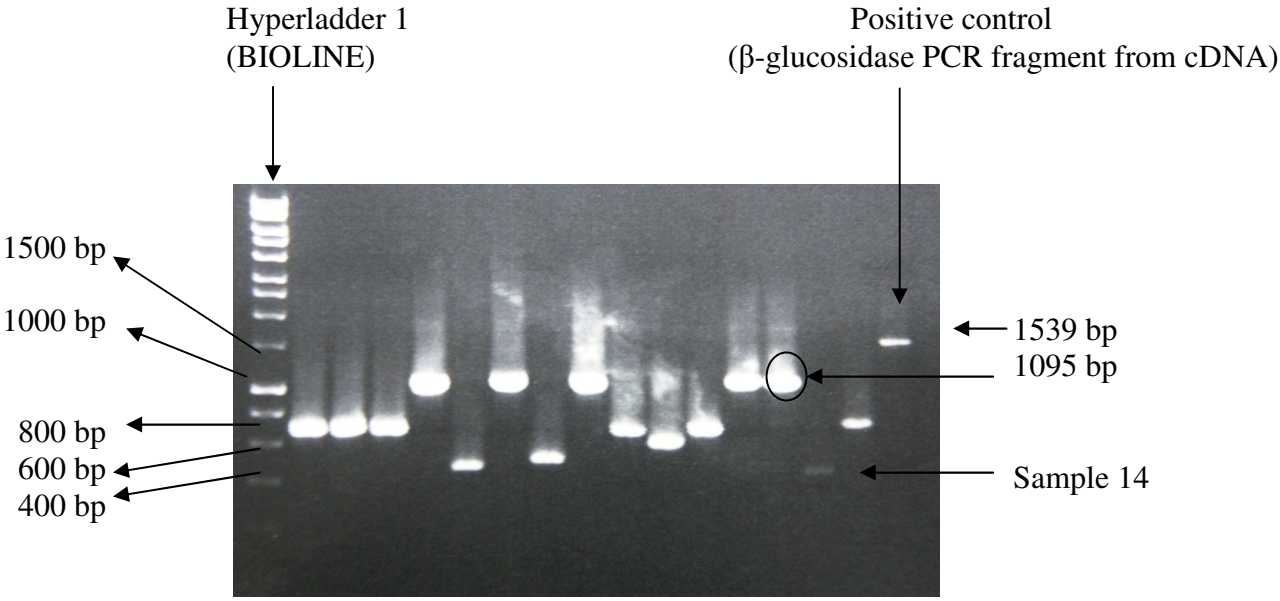


Figure 4.7 (A) Developed X-ray film of *P. tinctorium* cDNA library probed with labelled β-glucosidase gene from *P. tinctorium*. Blue circles indicate positive plaques. (B) cDNA inserts amplified using 5' and 3' Seq. primers (Table 2.3). Sequenced inserts corresponded to the published β-glucosidase gene (Lanes 2-16: cDNA inserts, Lane 17: Positive control-β-glucosidase).

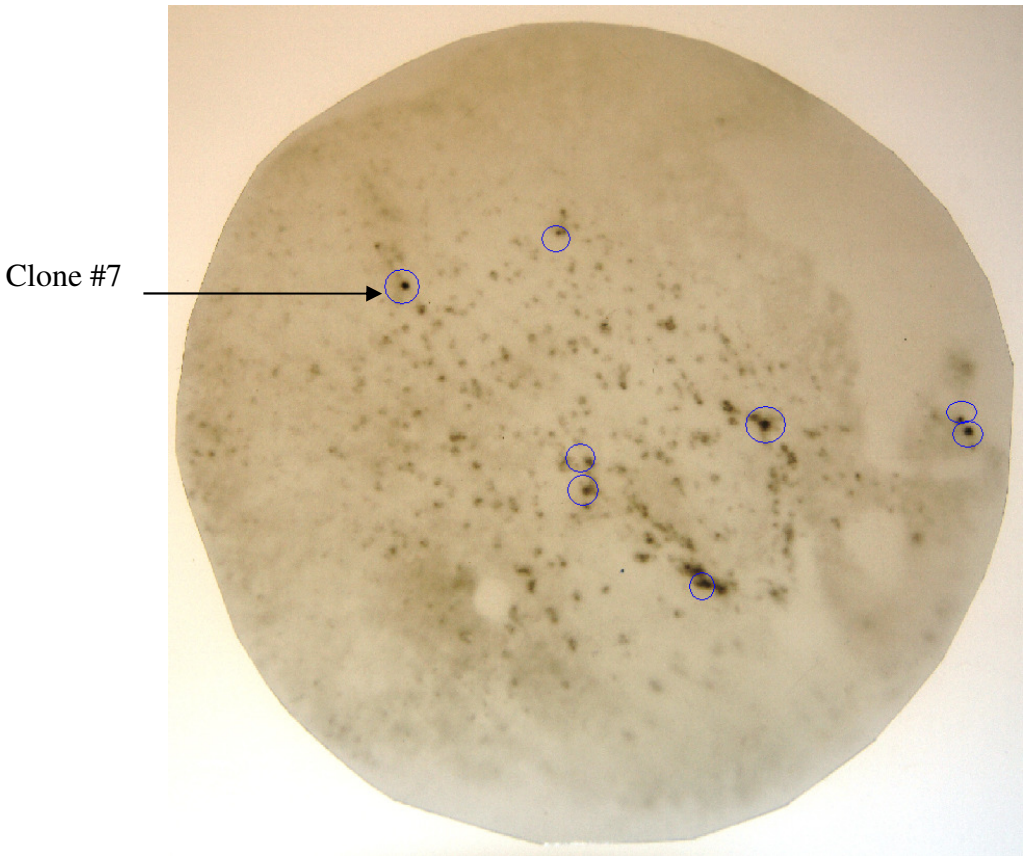


Figure 4.8 Primary screen of *P. tinctorium* cDNA library using maize IS (*Bx1*) as a probe. Blue circles indicate plaques that were eluted in preparation for secondary screening.

Eluted plaques were titered to determine the pfu/µL. Approximately 1000 potential positive clones, per eluted phage from the primary screen were re-screened using the labelled maize IS PCR product as a probe. Each eluted phage sample was grown on a separate 140mm petri dishes and transferred to nylon membranes. After probing and dual low stringency washes, membranes were subject to dual high stringency washes with 0.2 X SSC + 0.1% SDS (w/v) at 60°C. After exposure (Figure 4.9), the X-ray films displaying positives were aligned with corresponding agar plates and matching plaques were eluted in suspension media in preparation for tertiary screening.

Eluted plaques were once again titered to determine the pfu/µL. Approximately 50-100 potential positive clones per eluted phage from the secondary screen were re-screened using the labelled maize IS PCR product as a probe. Each eluted phage sample was grown on a separate 90mm petri dishes and transferred to nylon membranes. After probing and low stringency washes, membranes were subject to high stringency washes with 0.2 X SSC + 0.1% SDS (w/v) at 60°C. After exposure (Figure 4.10), the X-ray films dislaying positive plaques were aligned with corresponding agar plates and matching plaques were converted from phage DNA to plasmid DNA (*pTriplEx2* containing insert) in preparation for restriction enzyme digestion.

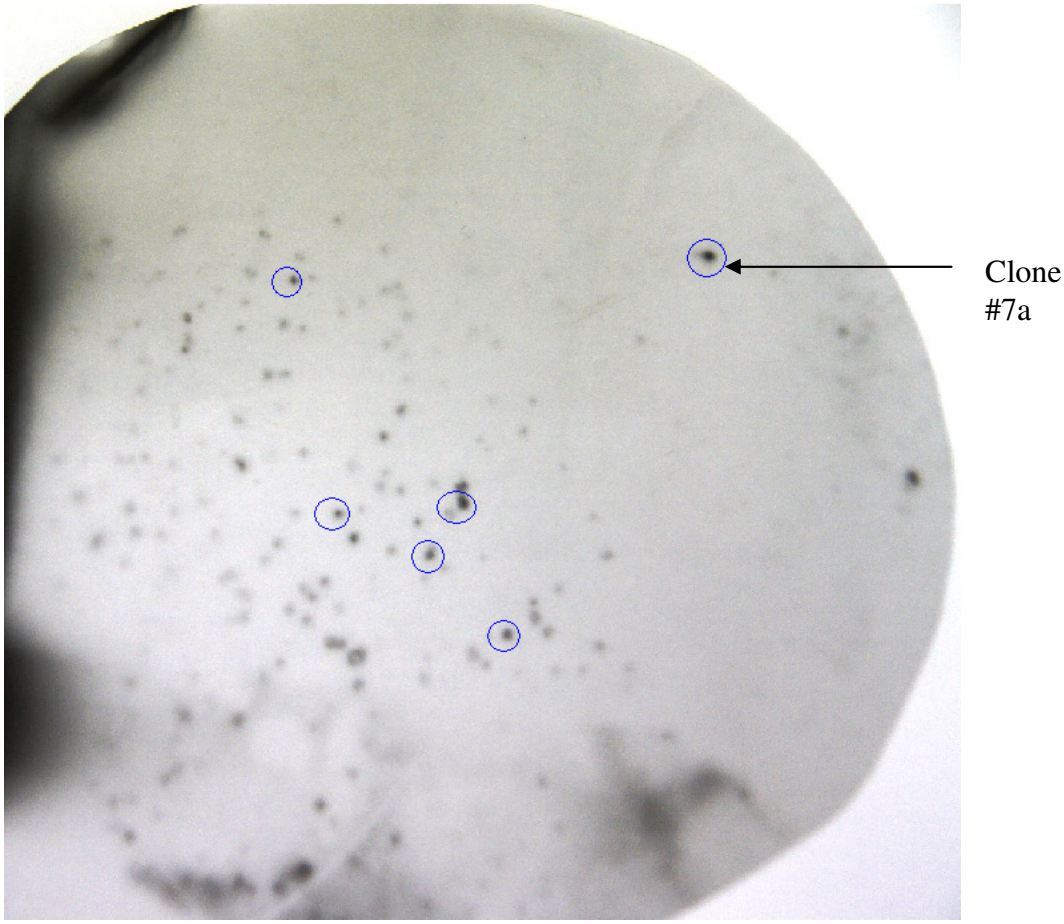


Figure 4.9 Secondary screen of *P. tinctorium* cDNA library using maize IS (*Bx1*) as a probe. Blue circles indicate plaques that were eluted in preparation for tertiary screening.

4.3.5.4 Southern blot analysis

A total of 8 potential positive plaques (Figure 4.10) were converted from phage to plasmid DNA. Plasmids (*pTriplEx2* containing inserts #7a 1-8) were digested with *Sfi*1 to excise the insert. Products were run on a DNA electrophoresis gel displaying inserts of approximately 600bp in length (Figure 4.11). DNA was then transferred to a nylon membrane which was probed with a labelled maize indole synthase. The membranes were then subject to dual low and high stringency washes [0.2 X SSC + 0.1% SDS (w/v) at 60°C].

After exposure, X-ray films showed positive hybridisation for insert #7a2 and #7a5. Plaques #7a2 and #7a5 were traced back to plaque #7a from the secondary screen then traced back to plaque #7a from the initial primary screen. The positive control (diluted maize IS PCR product) displayed prominent hybridisation with the probe, whilst the negative control (linearised *pTriplEx2* backbone) did not display any hybridisation with the labelled maize indole synthase (Figure 4.12). 'Positive' clones #7a2 and #7a5 were converted to *pTriplEx2* and sent for sequencing, however the returned sequences, when subject to the BLASTx tool on the NCBI website returned completely unrelated gene sequences. None of the resultant BLAST products were from *Z. mays* or *P. tinctorium*. Clone #7a5 was unable to be matched to any sequences of significant homology and therefore no BLAST results were obtained. This indicated that that results were in fact of a 'false positive' nature.

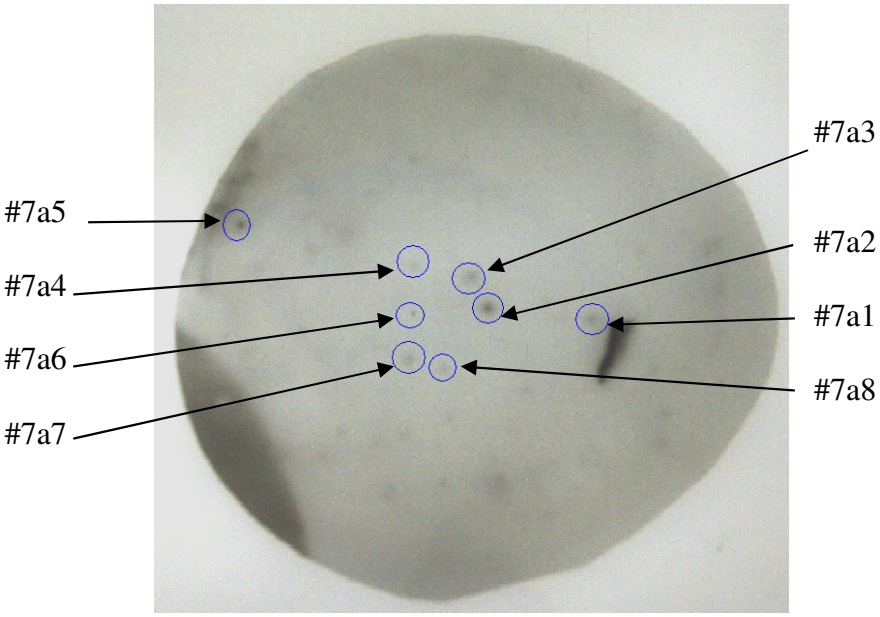


Figure 4.10Tertiary screen of *P. tinctorium* cDNA library using maize IS (*Bx1*) as a probe. Blue circles indicate plaques that were converted to plasmid DNA (*pTriplEx2* containing insert).

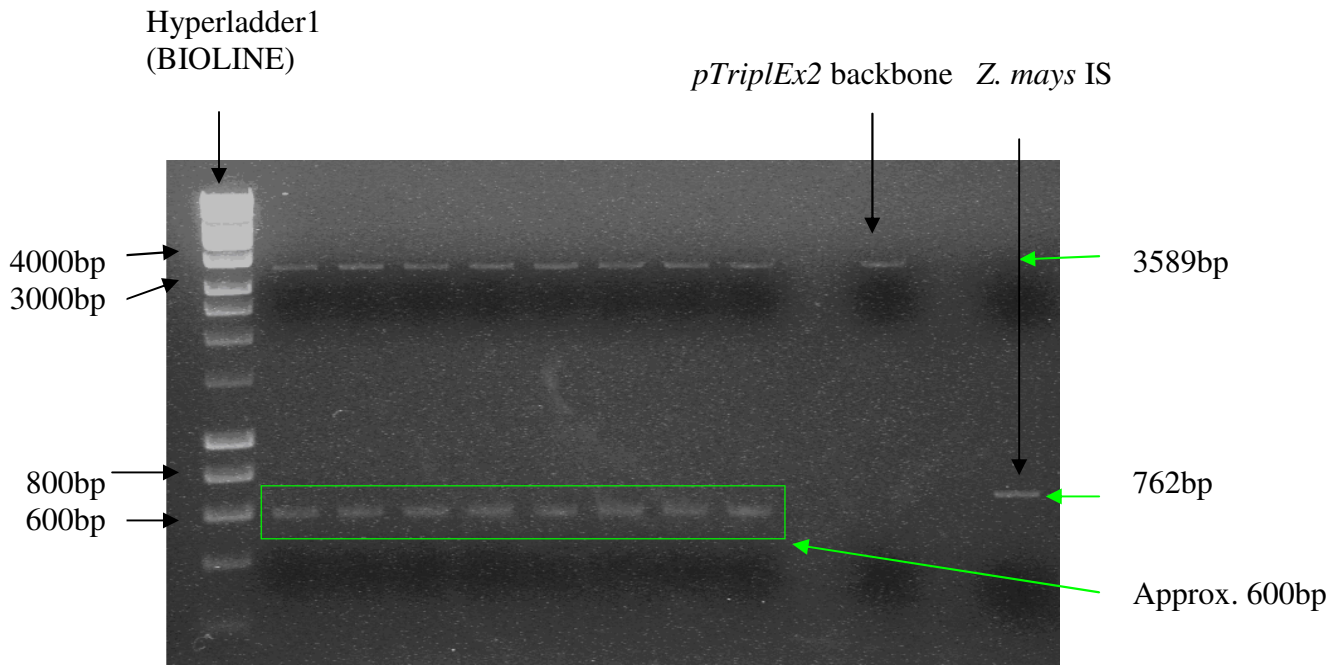


Figure 4.11 Restriction enzyme digests of potential positive clones from tertiary screen. Inserts were excised from 3589bp *pTriplEx2* via digestion with *Sfi*1. Linearised *pTriplEx2* is included as a negative control, whilst diluted maize IS PCR product is included as a positive control.

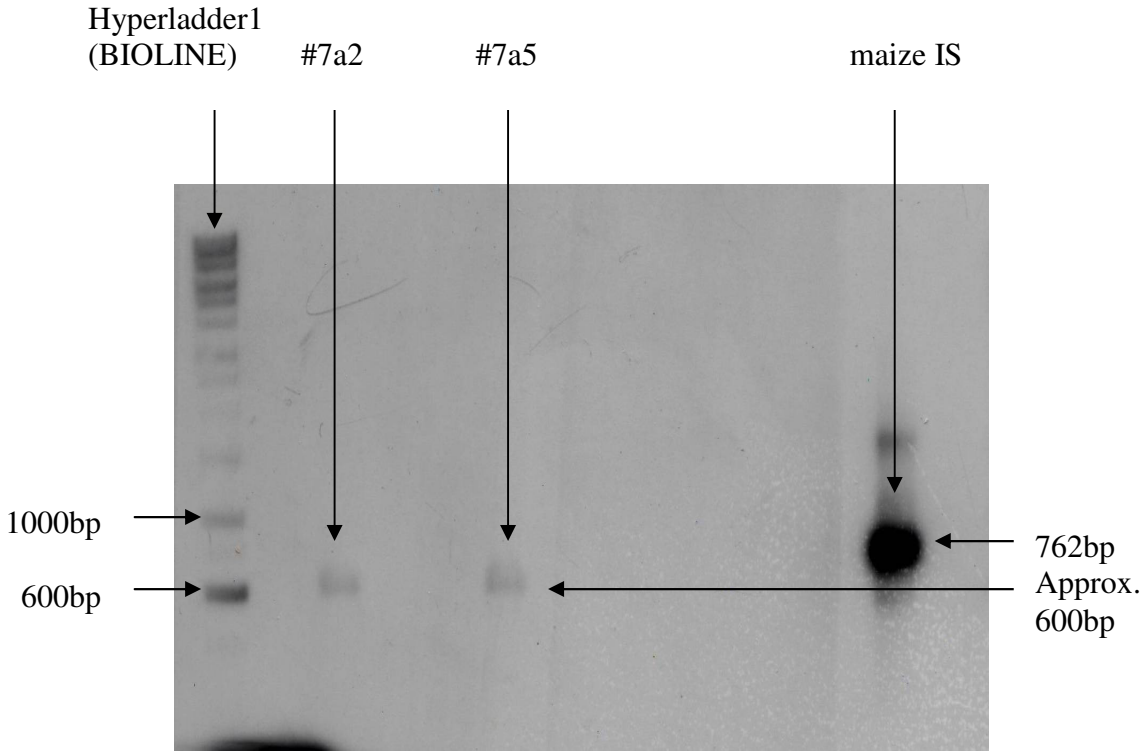


Figure 4.12 Southern blot analysis of inserts from tertiary screen.

4.3.5.5 Functional analysis of the cDNA library

Approximately 100,000 pfu's were bulk rescued and converted to plasmid DNA. The resulting mixed plasmid DNA prep was used to transform *E. coli* W3110*trp*A33 to assess if any clones within the cDNA library were capable of rescuing the strain in the absence of tryptophan, thus emulating the maize IS (Figure 4.13). Plating the transformed cultures on tryptophan rich LB ^{amp} and incubating at 37°C overnight enabled the observation that approximately 140,000 transformants were within the entire transformed culture. However, after the entire transformation was plated onto M9 minimal media, plates were incubated overnight at 37°C. There were no bacterial colonies indicating that of the 140,000 transformants, none were able to rescue the *trp*A auxotroph and as such function as the maize IS does by complementing the *trp*A mutation in *E. coli* W3110*trp*A33.

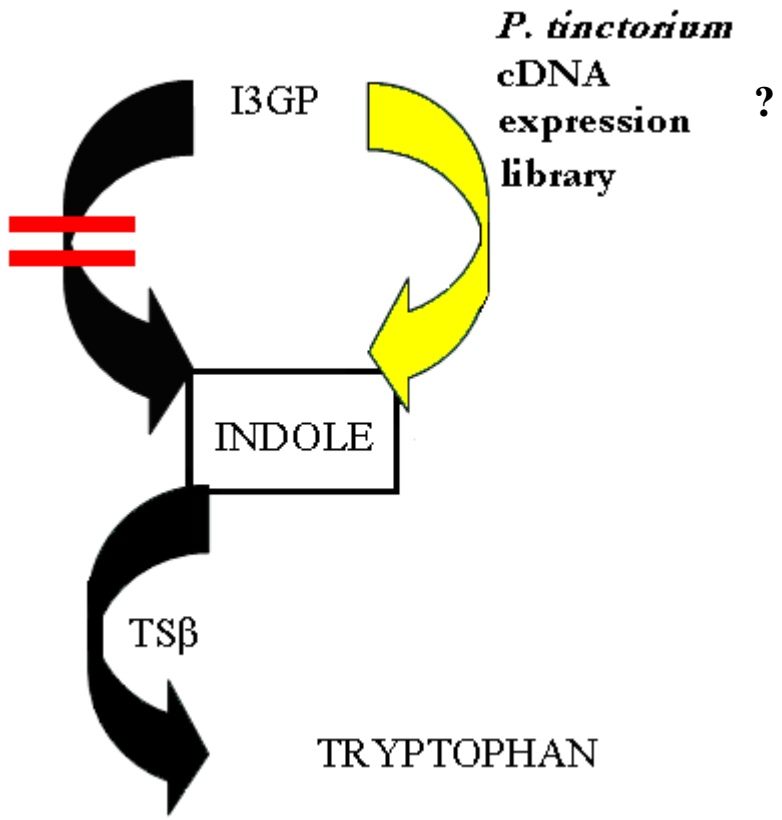


Figure 4.13 Potential scheme for rescuing *E. coli* W3110*trp*A with the *P. tinctorium* cDNA library. The double red lines indicate deficient TSα activity due to the *trp*A mutation. The bold yellow arrow represents a putative clone or clones within the library that can perform the *trp*A function in *E. coli* in the absence of a *P. tinctorium* TSβ sub-unit and in conjunction with an *E. coli* TSβ sub-unit.

4.4 Discussion

The *trp*A genes from *I. tinctoria* and *P. tinctorium* were selected for examination on the premise that they encoded enzymes with potentially the same catalytic properties as the maize IS, being the ability to perform the *trp*A function in the TSα deficient *E. coli* strain, W3110*trp*A33. However, the failure of these genes to rescue the strain in the absence of supplemented tryptophan indicated the inability to emulate the maize IS. These findings were supported by the high homology at amino acid level of of 93.6% and 93.8% for the *I. tinctoria* and *P. tinctorium* TSαs compared to *A. thaliana* TSα1, the negative control. This was in direct contrast to the significantly lower amino acid homology of the *I. tinctoria* and *P. tinctorium* TSαs compared to the mazie IS of 56.2% and 56.3% respectively. Due to the sequences being quite divergent at amino acid level, it was reasonable to expect different properties between the TSα sub-units and the maize IS due to those differences. The TSα2 from *A. thaliana*, believed to be a 'TSα like sub-unit' displayed all the required amino acids for TSα catalytic activity, however, it showed only a 71.2% homology with the *A. thaliana* TSα1.

The research findings within this thesis are inconsistent with the findings of Salvini *et al*, (2008) in which they claim the *I. tinctoria* TSα can complement a *trp*A auxotroph in a similar fashion to the maize IS. This thesis details complementation experiments using the *trp*A mutant *E. coli* W3110*trp*A33. This strain is deficient in TSα activity due to the replacement of the glutamic acid codon with a methionine codon at position 48 of the TSα protein resulting from at least two nucleotide substitutions and cannot be reverted back to prototrophy (Drapeau *et al*, 1968). In the current study, *I. tinctoria* TSα failed to rescue *trp*A deficient W3110*trp*A33 in the absence of exogenous tryptophan, contrary to the findings detailed by Salvini *et al*, (2008). The current study is strengthened by the inclusion of both negative (*A. thaliana* TSα1) and positive (*Z. mays* IS) controls for the complementation experiment, both absent in the research presented by Salvini *et al*, (2008). Furthermore, this

current study details sequence analysis highlighting greater percentage similarity, at both nucleotide and amino acid level, of the *I. tinctoria* TSα to *A. thaliana* TSα1 compared to the maize IS (Frey *et al*, 2000).

The limited amount of information regarding indigo synthesis by plants makes the screening of a *P. tinctorium* cDNA vital for the identification of genes with the ability to encode enzymes capable of catalysing indole not exclusive to tryptophan biosynthesis. It was calculated that the primary cDNA library contained approximately 2.5 X 10^6 independent clones. The library was probed with the radiolabelled maize IS gene approximately 5.0 X 10^5 independent clones at a time. Considering the library was screened a total of 4 times, approximately 2.0 X 10^6 independent clones were screened for the presence of a gene homologous to the maize IS (*Bx1*).

Screening the library with the heterologous maize IS probe was performed to identify any clones displaying sequence identity at nucleotide level. Using the high stringency hybridisation and wash conditions of 0.2 X SSC + 0.1% SDS (w/v) at 60°C should still allow clones that are at least 70-80% homologous to hybridise (Sambrook and Russell, 2001), however no clones were detected using these conditions. From these results it can be concluded that there is no gene within the *P. tinctorium* cDNA library that shows greater than 70-80% homology to the maize IS (*Bx1*) at nucleotide level. Given the results achieved using the β-glucosidase positive control gene, the quality of the cDNA library appeared to be of a high quality and the screening process seemed reliable.

Considering the indigo producing properties of *P. tinctorium* and its capacity to synthesise the indigo precursor indole, the search for an enzyme capable of performing the *trp*A function in *E. coli* in the absence of any plant TSβ sub-unit is of great interest. For this reason a functional screen was attempted using the *P. tinctorium* cDNA library. The failure

of the cDNA library to rescue *E. coli* W3110*trp*A33 indicated the absence of any clone able to complement the *trp*A deficient strain. Therefore it can be concluded from these results that there is no functional clone within the *P. tinctorium* cDNA library that is capable of catalysing the direct synthesis of indole nor is there a clone encoding a protein that can combine with an *E. coli* TSβ sub-unit for tryptophan biosynthesis.

The non-existance of a functional clone within the cDNA library may be attributed to two potential factors. Either the library does not contain any clones capable of performing the *trp*A function when expressed in *E. coli*, or there is a clone within the libray that may encode an IS like protein but given the chloroplast targeting of this protein retention of the chloroplast transit peptide is inhibiting the clone from functioning within the *E. coli*. If there was a full length putative indole synthase within the cDNA library, presumably the transit peptide would be retained. Cleavage of the transit peptide may be required before the IS protein can function, the retention of the transit peptide may inhibit enzyme function when expressed in *E. coli*, and as a result fail to rescue a *trp*A deficient strain of *E. coli*.

5 Chapter Five: General discussion and conclusions

Indole is a fundamental precursor not only for indigo biosynthesis, but for the generation of defense compounds and amino acids. The research within this thesis examines the role of indole in plants and bacteria and how indigo producing plants are able to utilise it to generate vast quantities of indigo.

This thesis detailed the function of the maize IS gene when expressed in *E. coli* as well as utilising the gene as a heterologous probe to screen a *P. tinctorium* cDNA library via hybridisation. This gene encodes an indole synthase enzyme that catalyses the synthesis of indole for subsequent production of the hydroxamic defense compound DIMBOA in maize (Melanson *et al*, 1997; Frey *et al*, 2000). Indole is also a fundamental intermediate in the tryptophan biosynthetic pathway. The last two steps of the tryptophan biosynthetic pathway in both plants and bacteria are mediated by the tryptophan synthase $\alpha_2\beta_2$ heterotetramer enzyme complex, The primary function of the α sub-unit within this complex, encoded by the *trp*A gene, is to cleave I3GP and form indole for subsequent tryptophan biosynthesis by β sub-unit activity. In a typical $TS\alpha_2\beta_2$ complex, indole is confined to an intramolecular tunnel connecting the active site of each sub-unit. Whilst the maize IS and TSα sub-units appear to perform the same catalytic function, in maize the IS however appears to be fully functional in homomeric form in the absence of a TSβ sub-unit whilst the TSα is only active when bound to the TSβ sub-unit. Furthermore, the IS is reported to be up to 30 fold times more efficient at cleaving I3GP than a typical $TS\alpha_2\beta_2$ enzyme complex (Frey *et al*, 2000; Gierl and Frey, 2001).

TSα sub-units are conserved across all kingdoms (Radwanski and Last, 1995). It is evident that the gene encoding TSα activity (*trp*A) in plants was duplicated and has evolved to play a role in secondary metabolism, probably involved in defense mechanisms. As evolution

progressed, expression of maize *Bx1* was altered for the gene to function in secondary metabolism with the properties of the encoded enzyme modified to the specific catalytic function (Gierl and Frey, 2001). The maize IS is divergent from TSαs, but it has been suggested that the amino acid residues essential to TSα activity have been retained (Radwanski *et al*, 1995). However, the IS and TSαs display amino acid differences, in particular, within the domain responsible for interation with the TSβ sub-unit. These amino acid differences may be responsible for the differences observed between the IS and TSα enzymes (Frey *et al*, 2000; Salvini *et al*, 2008).

The results in this thesis indicate that the maize IS forms a functional complex with an *E. coli* TSβ sub-unit. The successful knockout of the *tna*A and elimination of tryptophanase activity then subsequent complementation of the *trp*A*tna*A mutant indicate that this gene is able to perform the *trp*A function. The indole produced as a part of this process is then utilised by an *E. coli* TSβ sub-unit for ensuing tryptophan biosynthesis. This is in contrast to the inability of a plant TSα sub-unit to rescue an *E. coli trp*A auxotroph in the absence of the specific plant TSβ counterpart. In light of these results, it is necessary to examine the differences between TSα and TSβ sub-units from both plants and *E. coli*. When an alignment of the amino acid sequence from the plant TSα's used in this study and the published *E. coli* TSα (Figure 5.1) is performed, there is only a 43% similarity between the *E. coli* protein and each plant α sub-unit. However, each of the catalytic residues essential to TSα activity are conserved. Sequence alignment between the IS and the *E. coli* TSα (Figure 5.2) allows the observation that there is only a 42% identity. It is also evident that in the domain responsible for β sub-unit interaction, there are specific amino acid differences. Additionally, it is also necessary to examine the TSβ sub-units of both plants and *E. coli*. At only 53.5% similar at amino acid level (Figure 5.3) the respective TSβ sub-units from *E. coli* and *A. thaliana* are vastly different.

```
AtTSα1      ASLSTSSPTLGLADTFTQLKKQGKVAFIPYITAGDPDLSTTAEALKVLDACGSDIIELGV 60
It TSα      .A...............KE.............................A.S.........
Pt TSα      .A...............KE.............................A.S.........
E. coli TSα --------MERYESL.A...ERKEG..V.FV.L...GIEQSLKIIDT.IEA.A.AL...I

AtTSα1      PYSDPLADGPVIQAAATRSLERGTNLDSILEMLDKVVP-QISCPISLFTYYNPILKRGLG 120
It TSα      ...................A..N.....N..D.....L...V...V...............
Pt TSα      ...................A..N.....N..D.....L...V...V...............
E. coli TSα .F........T..N.TL.AFAA.VTPAQCF...ALIRQKHPTI..G.LM.A.LVFNK.ID

AtTSα1      KFMSSIRAVGVQGLVVPDVPLEETEMLRKEALNNDIELVLLTTPTTPTERMKLIVDASEG 180
It TSα      .......D..............N..................R.......
Pt TSα      .......D.......E.....N..................R.......
E. coli TSα E.YAQCEK...DSVL.A...V..SAPF.QA..RHNVAPIFICP.NADDDLLRQ.ASYGR.

AtTSα1      FIYLVSSIGVTGARSSVSGKVQSLLKDIKEATDKPVAVGFGISKPEHVKQIAGWGADGVI 240
It TSα      ...............A.............................Q.................
Pt TSα      ...............A.............................Q.................
E. coli TSα YT..L.RA.....ENRAALPLNH.VAKL..YNAA.PLQ.....A.DQ..AAIDA..A.A.

AtTSα1      VGSAMVKLLGDAKS-PTEGLKELEKLTKSLKSALL-
It TSα      ......R..................R...........
Pt TSα      ......R..................R...........
E. coli TSα S...I..IIEQHINE.EKM.AA.KVFVQPM.A.TRS
```

Figure 5.1 Alignment of TSα mature peptides from *I. tinctoria*, *P. tinctorium*, *A. thaliana*. and *E. coli*. Dots represent conserved amino acids for all sequences. Bold black squares above the residues represent amino acids essential to catalytic activity (Radwanski *et al*, 1995). The yellow highlighted region indicates two glutamic acid residues residing in the domain reported to be responsible for interaction with the TSβ sub-unit (Frey *et al*, 2000). Gaps that maximise alignment are represented by a hyphen.

```
E. coli TSα  MERYESLFAQLKERKEGAFVPFVTLGDPGIEQSLKIIDTLIEAGADALELGIPFSDPLAD 60
ZmIS         --------.A.MAKGKT..I.YI.A...DLATTAEALRL.DGC...VI...V.C...YI.

E. coli TSα  GPTIQNATLRAFAAGVTPAQCFEMLALIRQKHPTIPIGLLMYANLVFNKGIDEFYAQCEK 120
ZmIS         ..I..ASVA..L.S.T.MDAVL...REVTP-ELSC.VV..S.YKPIMSR----SL.EMKE

E. coli TSα  VGVDSVLVADVPVEESAPFRQAALRHNVAPIFICPPNADDDLLRQIASYGRGYTYLLSRA 180
ZmIS         A..HGLI.P.L.YVAAHSLWSE.KNN.LELVLLTT.AIPE.RMKE.TKASE.FV..V.VN

E. coli TSα  GVTGAENRAALPLNHLVAKLKEYNAAPPLQGFGISAPDQVKAAIDAGAAGAISGSAIVKI 240
ZmIS         ....PRANVNPRVES.IQEV.KVTNK.VAV.....K.EH..QIAQW..D.V.I...M.RQ

E. coli TSα  IEQHINEPEKMLAALKVFVQPMKAATRS
ZmIS         LGEAAS-.KQG.RR.EEYARG..N.LP-
```

Figure 5.2 Alignment of mature peptides from sequenced maize IS and TSα from *E. coli*. Dots represent conserved amino acids for all sequences. Bold black squares above the residues represent amino acids essential to catalytic activity (Radwanski *et al*, 1995). The yellow highlighted region indicates glutamic acid residues within the domain reported to be responsible for interaction with the TSβ sub-unit (Frey *et al*, 2000). Gaps that maximise alignment are represented by a hyphen.

```
AtTSβ      MAASGTSATFRASVSSAPSSSSQLTHLKSPFKAVKYTPLPSSRSKSSSFSVSCTIAKDPP 60
E. coli TSβ ------------------------M.T.LN.Y------------------------------

AtTSβ      VLMAAGSDPALWQRPDSFGRFGKFGGKYVPETLMHALSELESAFYALATDDDFQRELAGI 120
E. coli TSβ ---------------------..E...M...QI..P..RQ..E..VSAQK.PE..AQFNDL

AtTSβ      LKDYVGRESPLYFAERLTEHYRRENGEGPLIYLKREDLNHTGAHKINNAVAQALLAKRLG 180
E. coli TSβ ..N.A..PTA.TKCQNI.A------.TNTTL.......L.G....T.QVLG.......M.

AtTSβ      KKRIIAETGAGQHGVATATVCARFGLECIIYMGAQDMERQALNVFRMRLLGAEVRGVHSG 240
E. coli TSβ .TE.............S.LAS.LL..K.R.....K.V...SP.......M....IP....

AtTSβ      TATLKDATSEAIRDWVTNVETTHYILGSVAGPHPYPMMVRDFHAVIGKETRKQALEKWGG 300
E. coli TSβ S......CN..L...SGSY..A..M..TA.......TI..E.QRM..E..KA.I..RE.R

AtTSβ      KPDVLVACVGGGSNAMGLFHEFVNDTEVRMIGVEAAGFGLDSGKHAATLTKGDVGVLHGA 360
E. coli TSβ L..AVI.........I.M.AD.I.E.N.GL....PG.H.IET.E.G.P.KH.R..IYF.M

AtTSβ      MSYLLQDDDGQIIEPHSISAGLDYPGVGPEHSFFKDMGRAEYYSITDEEALEAFKRVSRL 420
E. coli TSβ KAPMM.TE....E.SY.......F.S...Q.AYLNST...D.V....D.......TLCLH

AtTSβ      EGIIPALETSHALAYLEKLCPTLSDG-TRVVLNFSGRGDKDVQTVAKYLDV----
E. coli TSβ ........S.....HAL.MMRENP.KEQLL.V.L.......IF..HDI.KARGEI
```

Figure 5.3 Alignment of peptides from published TSβ sequences from *A. thaliana* and *E. coli*. Dots represent conserved amino acids for all sequences. Gaps that maximise alignment are represented by a hyphen.

The contrasting differences in amino acid sequence between the *E. coli* TSα and the plant TSαs may be the point of difference between the binding capacities of each sub-unit. If the primary structure of a protein dictates the secondary structure of a protein (Dill *et al*, 1995) then it is logical to expect a significant difference in protein assembly between the *E. coli* TSα and the plant TSα's. This is also true for the β sub-units of the plant and *E. coli* tryptophan synthase complex. This may be the crucial difference regarding the different binding properties of each protein; therefore it is not surprising that a plant TSα sub-unit cannot form a functional enzyme complex with an *E. coli* TSβ sub-unit.

The secondary structure of the *E. coli* TSα has been predicted (Crawford *et al*, 1987), however, there is yet to be a prediction of any plant TSα or the maize IS secondary structure an area that warrants further investigation. This may aid in the elucidation of why the maize IS can form a functional enzyme complex with an *E. coli* TSβ.

The IS is an enzyme specific to the DIMBOA defense pathway in maize localised to the chloroplast and responsible for cleaving I3GP for indole synthesis. A range of duplicated cytochrome P450's then mediate the conversion of indole to DIMBOA, which is then glycosylated and sequestered in the vacuole (Gierl and Frey, 2001). The DIMBOA defense pathway is common to grasses, which are members of the monocotyledon class of angiosperms. It is likely that grasses such as maize developed this defense mechanism as evolution progressed in response to specific herbivore attack. The DIMBOA defense pathway is mainly limited to monocots (Pratt *et al*, 1995), therefore it is improbable that there is a gene within plants of the dicotyledon class of angiosperms, such as *P. tinctorium* that displays significant nucleotide homology with the maize IS gene (*Bx1*).

In healthy indigo producing plants such as *P. tinctorium*, it is speculated that indigo has no known function, however it has been reported that indigo dye resulting from plant maceration

may have antimicrobial properties (Iwaki *et al*, 2006). Considering the links that indigo precursors have with plant defense mechanisms it is plausible that there is a chloroplast targeted gene within indigo producing plants displaying analogous properties to those of the maize IS. This putative IS would likely utilise I3GP branching off the tryptophan pathway for indole synthesis.

In 2008 there was a report of a cytosol localised indole synthase gene in *A. thaliana* (INS) that had the ability to complement a *trp*A auxotroph (Zhang *et al*, 2008). Alignment of the translated INS nucleotide sequence with the *A. thaliana* TSα1 and maize IS peptides display percentage identities of 74% and 55% respectively suggesting that the INS is more likely to have similar properties to the TSα1 than the IS. Despite these deficiences, Zhang *et al*, (2008) claim that this INS enzyme is tryptophan biosynthetic pathway independent that catalyses the production of indole from I3GP in the cytosol for subsequent indole-3-acetic acid (IAA). The delocalisation of INS is consistent with the absence of a chloroplast transit peptide. The failure to include both a positive (maize IS) and negative (*A. thaliana* TSα1) control in the complementation experiment as well as failing to detail the bacterial strain and nature of the *trp*A mutation within the strain make drawing conclusions from that study an issue.

Considering that I3GP is located in the chloroplast as an intermediate in the tryptophan biosynthetic pathway, it would have to enter the cytosol either by diffusion through the chloroplast membrane or transported via active transport. However, there is contrasting evidence suggesting that such indole producing enzymes are indeed localised to the chloroplast (Kramer and Koziel, 1995; Frey *et al*, 2000; Gierl and Frey, 2001) thus investigation of the specific role of the INS in plants is also an area that warrants further examination. The presence of duplicate tryptophan biosynthetic pathway genes in plants (Niyogi and Fink, 1992; Last *et al*, 1995; Gierl and Frey, 2001) strengthens the argument that

the biosynthesis of secondary metabolites associated with indigo production is linked with the tryptophan biosynthetic pathway. These duplicate genes may be expressed in response to pathogen attack and tryptophan may be produced via an alternative mechanism (Last *et al*, 1995). The mechanism for how indigo producing plants produce such large quantities of indigo precursors remains unclear.

The data within this thesis permits the conclusions that:

- The maize indole synthase forms a functional enzyme complex with an *E. coli* TSβ sub-unit.

Elimination of tryptophanase based tryptophan biosynthesis is confirmation that the indole generated by IS catalysed cleavage of I3GP is subsequently utilised by the TSβ as substrate for tryptophan biosynthesis, mimicking a typical TSαβ enzyme complex.

- The indole generated by IS activity cannot be utilised by the INO for indigo production in an *E. coli trp*A mutant.

This is confirmed by the requirement of exogenous indole either by direct supplementation into the media or by tryptophanase based tryptophan degradation for indigo accumulation to occur. This result also suggests that indole is confined within the IS/TSβ enzyme complex. It is unknown, however, if indole is confined to an intramolecular tunnel connecting both catalytic domains.

- A TSα sub-unit from *I. tinctoria* and *P. tinctorium* does not function in the absence of its plant TSβ counterpart.

The plant TSα's are unable to form a heterodimer complex with an *E. coli* TSβ sub-unit, therefore they do not have the same properties as the maize IS.

- There is no gene in the *P. tinctorium* cDNA library that is strongly homologous at the nucleotide level to the maize IS gene sequence (*Bx1*).

The IS from maize is too divergent from any gene with *P. tinctorium* and there may not be anything homologous to this gene within this indigo producing plant.

- There is no functional clone within the *P. tinctorium* cDNA that can perform the *trp*A function in a TSα deficient *E. coli* strain.

This can be attributed to the lack of a full length clone within the cDNA library. Alternatively, if any putative indole synthase gene is chloroplast localised, a full length clone within the cDNA library would not have had the transit peptide cleaved off when expressed in *E. coli*, therefore the clone is unlikely to function.

Future research associated with this thesis could entail the further examination of how such large quantities of indigo precursors are generated in indigo producing plants. At the time of submission of this thesis, work is underway to analyse the transcriptome of *P. tinctorium* as well as sequencing the entire cDNA library that was constructed as part of the current project. Future research may also involve isolating the INS gene describe by Zhang *et al*, (2008) and using this gene to probe the *P. tinctorium* cDNA library. This gene could also be subjected to attempt complementation experiments with the *trp*A*tna*A double mutant generated in this current project with the inclusion of both positive (maize IS) and negative (any plant TSα sub-unit) controls to substantiate the results. Although the maize

IGL enzyme is ten fold less efficient than the IS and cleaving I3GP for indole synthesis (Frey *et al*, 2000), this gene could also be isolated and subject to the same experiments as the IS.

The nature of how the maize IS forms a functional enzyme complex with an *E. coli* TSβ sub-unit could also be examined. Prediction of the secondary structure of the IS may aid in the elucidation of the binding properties of the protein and identify the specific mode by which a heterodimer is formed. Further experiments with *E. coli* mutants should also be undertaken. The co-expression of the maize IS and the *Rhodococcus* INO in an *E. coli* strain solely deficient in tryptophanase activity could be performed to examine whether the maize IS is capable of producing indole for potential indigo accumulation in bacteria when the TSαβ enzyme complex remained intact.

In conclusion this thesis reports the examination of indole biosynthesis in indigo producing plants. Specifically, this study examined the possible role of an Indole Synthase like enzyme in *P. tinctorium* and performed a detailed analysis of the interactions of the *Z. mays* IS with bacterial TSβ sub-units for tryptophan biosynthesis. This work has provided a more thorough understanding of the interactions of different components of the tryptophan and indigo biosynthetic pathways in plants and *E. coli*. This current study provides a platform to move forward with further research that will increase our understanding of these processes.

6 Bibliography

Anderson, K. S. M., E. W., and K. A. Johnson. 1991. Serine modulates substrate channeling in tryptophan synthase. The Journal of Biological Chemistry **266:**8020-8033.

Angelini, L. G., S. Tozzi, and N. Di Nassio. 2004. Environmental factors affecting yield productivity, indican content and indigo yield in *Polygonum tinctorium* Ait., A subtropical crop grown under temperate conditions. Agricultural and Food Chemistry **52:**7541-7547.

Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori. 2006. Construction of *Eschericia coli* K-12 in-frame single-gene knockout mutants: the Keio collection. Molecular Systems Biology **2:**1-11.

Balfour-Paul, J. 1998. Indigo. British Museum Press, London.

Banik, U., S. A. Ahmed, P. McPhie, and E. W. Miles. 1995 Subunit assembly in the tryptophan synthase alpha 2 beta 2 complex. Stabilization by pyridoxal phosphate aldimine intermediates. The Journal of Biological Chemistry. **270:**7944-9.

Bechtold, T., A. Turcanu, S. Geissler, and E. Ganglberger. 2002. Process balance and product quality in the production of natural indigo from *Polygonum tinctorium* Ait. Applying low-technology methods. Bioresource Technology **81:**171-177.

Bentley, R. 1990. The shikimate pathway-a metabolic tree with many branches. CRC Critical Reviews in. Biochemistry and Molecular Biology **25:**307-384.

Berlyn, M., R. L. Last, and G. R. Fink. 1989. A gene encoding the tryptophan synthase beta subunit of *Arabidopsis thaliana*. Genetics **86:**46044-4608.

Bhattacharya, S. 1977. The indigo revolt of Bengal. Social Scientist **5:**13-23.

Bilezikian, J. P., R. O. R. Kaempfer, and B. Magasanik. 1967. Mechanism of tryptophanase induction in Escherichia coli. Journal of Molecular Biology **27:**495-506.

Bunting, D. 2000. Wild indigo: *Baptisia tinctoria*. *In* R. Gladstar and P. Hirsch (ed.), Planting the Future. Healing Arts Press, Rochester, Vermont.

Burns, J., T. Yokota, H. Ashihara, M. E. J. Lean, and A. Crozier. 2002. Plant foods and herbal sources of resveratrol. Journal of Agricultural and Food Chemistry **50:**3337-3340.

Campeol, E., L. G. Angelini, S. Tozzi, and M. Bertolacci. 2006. Seasonal variation of indigo precursors in *Isatis tinctoria* L. and *Polygonum tinctorium* Ait. as affected by water deficit. Environmental and Experimental Botany **58:**223-233.

Celik, A., R. E. Speight, and N. Turner. 2005. Identification of broad specificity $P450_{CAM}$ variants by primary screening against indole as substrate. Chemical Communications.**:**3652-3654.

Chanayath, N., S. Lhieochaiphant, and S. Phutrakul. 2002. Pigment extraction techniques from the leaves of *Indigofera tinctoria Linn*. and *Baphicacanthus cusia Brem*. and chemical structure analysis of their major components. Chiang Mai University Journal **1:**149-160.

Cheah, E., K. MacPherson, D. Quiggin, P. Keese, and D. L. Ollis. 1998. Crystallization and preliminary X-ray analysis of IND, an enzyme with indole oxygenase activity from *Chromobacterium violaceum*. Acta Crystallographica. **54:**657-658.

Clark, R. J. H., C. J. Cooksey, M. A. M. Daniels, and R. Withnall. 1993. Indigo, woad, and tyrian purple: important vat dyes from antiquity to the present. Endeavour **17:**191-199.

Crawford, I. P. 1979. Gene rearrangements in the evolution of the tryptophan synthetic pathway. Bacteriological Reviews **39:**87-120.

Crawford, I.P., T. Niermann, and K. Kirschner. 1987. Prediction of secondary structure by evolutionary comparison: Application to the α sub-unit of tryptophan synthase. Proteins. **2:** 118-129.

Crawford, I. P. 1989. Evolution of a biosynthetic pathway: the tryptophan paradigm. Annual Review of Microbiology **43:**567-600.

Datsenko, K. A., and B. L. Wanner. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Science U S A **97:**6640-6645.

Davies, K. M. 2004. Plant pigments and their manipulation, vol. 14. Blackwell Publishing, Oxford.

Deeley, M. C., and C. Yanofsky. 1981. Nucleotide sequence of the structural gene for tryptophanase of *Escherichia coli* K-12. Journal of Bacteriology **147:**787-796.

Dill, K. A., S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yee, P. D. Thomas, H. S. Chan. 1995. Principles of protein folding-A perspective from simple exact models. Protein Science **4:** 561-602.

Doukyu, N., K. Toyoda, and R. R. Aono. 2003. Indigo production by *Escherichia coli* carrying the phenol hydroxylase gene from *Acinetobacter* sp. strain ST-550 in a water–organic solvent two-phase system. Applied Microbiology and Biotechnology **60:**720-725.

Drapeau, G. R., W. J. Brammar, and C. Yanofsky. 1968. Amino acid replacements of the glutamic acid residue at position 48 in the tryptophan synthetase a protein of *Escherichia coli* Journal of Molecular Biology **35:**357-367.

Dunn, M. F., V. Aguilar, P. Brzović, W. F. Drewe, K. F. Houben, C. A. Leja, and M. Roy. 1990. The tryptophan synthase bienzyme complex transfers indole between the alpha- and beta-sites via a 25-30 A long tunnel. Biochemistry **29:**8598-607.

Edwards, J. M., and L. M. Jackman. 1965. Chorismic acid. A branch point in aromatic biosynthesis. Australian Journal of Chemistry. **18:**1227-1239.

Ensley, B. D., B. J. Ratzkin, T. D. Osslund, M. J. Simon, L. P. Wackett, and D. T. Gibson. 1983 Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. Science **222:**167-169.

Evans, W. C., W. Richard, C. Handley, and F. C. Happold. 1941. The tryptophanase-indole reaction: some observations on the production of tryptophanase by *Esch. coli*; in particular the effect of the presence of glucose and amino acids on the formation of tryptophanase. Biochemistry Journal. **35:**207–212.

Farber, G.K., and G.A. Petsko. 1990. The evolution of α/β barrel enzymes. Trends in Biochemical Science. **15:** 228-234.

Fawcett, C. H. 1961. Indole auxins. Annual Review of Plant Physiology **12:**345-368.

Ferreira, E. S., A. N. Hulme, H. McNab, and A. Quye. 2004. The natural constituents of historical textile dyes. Chemical Society Reviews **33:**329 - 336.

Frey, M., P. Chomet, E. Glawischnig, C. Stettner, S. Grun, A. Winklmair, W. Eisenreich, A. Bacher, R. B. Meeley, S. P. Briggs, K. Simcox, and A. Gierl. 1997. Analysis of a chemical plant defense mechanism in grasses. Science **277:**696-699.

Frey, M., C. Stettner, P. W. Pare, E. A. Schmelz, J. H. Tumlinson, and A. Gierl. 2000 An herbivore elicitor activates the gene for indole emission in maize Proceedings of the National Academy of Science USA. **97:**14801-14806.

Garcia-Macias, P., and P. John. 2004. Formation of indigo derived from woad (*Isatis tinctoria* L.) in relation to product purity. Journal of Agricultural and Food Chemistry **52:**7891-7896.

Gibson, F., and J. Pittard. 1968. Pathways of biosynthesis of aromatic amino acids and vitamins and their control in microorganisms. Bacteriological Reviews **32:**465-492.

Gierl, A., and M. Frey. 2001. Evolution of benzoxazinone biosynthesis and indole production in maize. Planta **213:** 493-498.

Gilbert, K. G., and D. T. Cooke. 2001. Dyes from plants: past usage, present understanding and potential. Plant Growth Regulation **34:**57–69.

Gilbert, K. G., H. G. Maule, B. Rudolph, M. Lewis, H. Vandenburg, E. Sales, S. Tozzi, and D. T. Cooke. 2004 Quantitative analysis of indigo and indigo precursors in leaves of *Isatis* spp. and *Polygonum tinctorium* Biotechnology Progress. **20:**1289-1292.

Gillam, E. J. M., A. M. A. Aguinaldo, L. M. Notley, D. Kim, R. D. Mundkowski, A. A. Volkov, F. H. Arnold, P. Soucek, J. J. DeVoss, and F. P. Guengerich. 1999. Formation of indigo by recombinant mammalian cytochrome P4501. Biochemical and Biophysical Research Communications **265:**469–472.

Gillam, E. M., L. M. Notley, C. Hongliang, J. J. De Voss, and F. P. Guengerich. 2000. Oxidation of indole by cytochrome P450 enzymes. Biochemistry **39:**13817-13824.

Gillam, E. M. J., and F. P. Guengerich. 2001. Exploiting the versatility of human cytochrome P450 enzymes: The promise of blue roses from biotechnology. Life **52:**271-277.

Gish, K., and C. Yanofsky. 1993. Inhibition of expression of the tryptophanase operon in *Escherichia coli* by extrachromosomal copies of the tna leader region. Journal of Bacteriology **175:**3380-3387.

Gong, F., K. Ito, Y. Nakamura, and C. Yanofsky. 2001. The mechanism of tryptophan induction of tryptophanase operon expression: tryptophan inhibits release factor-mediated cleavage of TnaC-peptidyl-tRNAPro. Proceedings of the National Academy of Science USA. **98:**8997-9001.

Grant, S. G., J. Jessee, F. R. Bloom, and D. Hanahan. 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proceedings of the National Academy of Science USA. **87:**4645-4649.

Han, X., W. Wang, and X. Xiao. 2008. Microbial biosynthesis and biotransformation of indigo and indigo-like pigments. Chinese Journal of Biotechnology **24:**921-926.

Hanioka, N., N. Tatarazako, H. Jinno, K. Arizono, and M. Ando. 2000. Determination of cytochrome P450 1A activities in mammalian liver microsomes by high-performance liquid chromatography with fluorescence detection. Journal of Chromatography B: Biomedical Sciences and Applications **744:**399-406.

Hardwood, L. M., C. J. Moody, and J. M. Percy. 1999. Experimental organic chemistry, standard and microscale, 2nd ed. Blackwell Science.

Hart, S., R. Kirby, and D. R. Woods. 1990. Structure of a *Rhodococcus* gene encoding blue pigment production in *Escherichia coli*. Journal of General Microbiology **136:**1357-1363.

Hart, S., K. R. Koch, and D. R. Woods. 1992. Identification of indigo related pigments produced by *Escherichia coli* containing a cloned *Rhodococcus* gene. Journal of General Microbiology **138:**211-216.

Hayashi, K. 1978. Chemical procedure for the determination of plant dyes in ancient Japanese textiles. , International Symposium on the Conservation and Restoration of Cultural Property. Tokyo National Research Institute of Cultural Properties Tokyo.

Herrmann, K. M. 1995. The shikimate pathway: early Steps in the biosynthesis of aromatic compounds. The Plant Cell, Vol. 7, 907-919, July 1995 O 1995 American Society of Plant Physiologists **7:**907-919.

Hommel, U., A. Lustig, and K. Kirschner. 1989. Purification and characterization of yeast anthranilate phosphoribosyltransferase. European Journal of Biochemistry **180:**33-40.

Hsiao, P., M. Sanjaya, S. Ruey-Chhih, J. A. Teixeira, and C. Ming-Tsair. 2006. Plant native tryptophan synthase beta 1 gene is a non-antibiotic selection marker for plant transformation. Planta **225:**897-906.

Hull, A. K., R. Vij, and J. L. Celenza. 2000. *Arabidopsis* cytochrome P450's that catalyze the first step of tryptophan-dependant indole-3-acetic acid biosynthesis. Proceedings of the National Academy of Science USA **97:**2379-2384.

Hurry, J. B. 1973. The woad plant and its dye Oxford University Press, London.

Hutter, R., and P. Niederberger. 1986. Tryptophan biosynthetic genes in eukaryotic organisms. Annual Review of Microbiology **40:**55-77.

Hyde, C. C., S. A. Ahmed, E. A. Padlan, E. W. Miles, and D. R. Davies. 1988 Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from *Salmonella typhimurium*. The Journal of Biological Chemistry **263:**17857-17871.

Ibel, K., R. P. May, K. Kirschner, A. N. Lane, H. Szadkowski, M. T. Dauvergne, and M. Zulauf. 1985. The domain structure of tryptophan synthase. A neutron scattering study. European Journal of Biochemistry **151:**505-514.

Isin, E. M., and F. P. Guengerich. 2007. Complex reactions catalysed by cytochrome P450 enzymes. Biochimica et Biophysica Acta **1770:**314-329.

Ito, J., and I. P. Crawford. 1965. Regulation of the enzymes of the tryptophan pathway in *Eschericia coli*. Genetics **52:**1303–1316.

Iwaki, K., S. Koya-Miyata, and K. Kohno. 2006. Antimicrobial activity of *Polygonum tinctorium* lour: extract against oral pathogenic bacteria. Journal of Natural Medicine **60:** 121-125.

Jing, Y., D. Cui, F. Bao, Z. Hu, Z. Qin, and Y. Hu. 2009. Tryptophan deficiency affects organ growth by retarding cell expansion in *Arabidopsis*. The Plant Journal **57:**511-521.

Karapetyan, G., K. Chakrabarty, M. Hein, and P. Langer. 2011. Synthesis and bioactivity of carbohydrate derivatives of indigo, Its isomers and heteroanalogues. ChemMedChem **6:**25-37.

Kelly, S. L., D. E. Kelly, C. J. Jackson, A. G. S. Warrilow, and D. C. Lamb. 2005. The diversity and importance of microbial cytochromes P450. *In* P. R. O. d. Montellano (ed.), Cytochrome P450: Structure, Mechanism and Biochemistry. Plenum Publishers, New York.

Kende, H., and J. Zeevaart. 1997. The five "classical" plant hormones. Plant Cell. 1997 July; 9(7): 1197–1210 **9:**1197-1210.

Kim, J., K. Lee, Y. Kim, C. Kim, and K. Lee. 2003. Production of dyestuffs from indole derivatives by naphthalene dioxygenase and toluene dioxygenase. Letters in Applied Microbiology **36:**343-348.

Kim, J.-Y., J.-Y. Lee, Y.-S. Shin, and G.-J. Kim. 2009. Mining and identification of a glucosidase family enzyme with high activity toward the plant extract indican. Journal of Molecular Catalysis B: Enzymatic **57:**284-291.

Kim, S., K. Song, D. Jung, Y. Chae, and H. J. Lee. 1996. Production of indoxyl derivatives in indole-supplemented tissue cultures of *Polygonum tinctorium*. Planta Medica **62:**54-56.

Kirschner, K., H. Szadkowski, T. S. Jardetzky, and V. Hager. 1987. Phosphoribosylanthranilate isomerase-indoleglycerol-phosphate synthase from *Escherichia coli*. Methods Enzymology **142:**386-397.

Kokubun, T., J. Edmonds, and P. John. 1998. Indoxyl derivatives in woad In relation to medieval indigo production. Phytochemistry **49:**79-87.

Kramer, V. C., and M. G. Koziel. 1995. Structure of a maize tryptophan synthase alpha subunit gene with pith enhanced expression. Plant Molecular Biology **27:**1183-1188.

Kulik, V., E. Hartmann, M. Weyand, M. Frey, A. Gierl, D. Niks, M.F. Dunn, I. Schilling. 2005. On the structural basis of the catalytic mechanism and regulation of the alpha sub-unit of tryptophan synthase from *Salmonella typhimurium* and Bx1 from maize, two evolutionary related enzymes. Journal of Molecular Biology. **352:** 608-620.

Lane, A. N., and K. Kirschner. 1983. The catalytic mechanism of tryptophan synthase from *Escherichia coli* kinetics of the reaction of indole with the enzyme-L-serine complexes. European Journal of Biochemistry **12:**571 -582.

Last, R. L., P. H. Bissinger, D. J. Mahoney, E. R. Radwanski, and G. R. Fink. 1991. Tryptophan mutants in *Arabidopsis*: The consequences of duplicated tryptophan synthase beta genes. The Plant Cell **3:**345-358.

Last, R. L., P. H. Bissinger, D. J. Mahoney, E. R. Radwanski, and G. R. Fink. 1991. Tryptophan mutants in *Arabidopsis*: The consequences of duplicated tryptophan synthase beta genes. The Plant Cell **3:**345-358.

Li, S.-L., R. M. Denney, and C. Yanofsky. 1973. Nucleotide sequence divergence in the alpha chain-structural genes of tryptophan synthetase from *Escherichia coli*, *Salmonella typhimurium* and *Aerobacter aerogenes*. Proceedings of the National Academy of Science U.S.A. **70**:1112-1116.

Li, J., and R. L. Last. 1996. The *Arabidopsis thaliana* trp5 mutant has a feedback-resistant anthranilate synthase and elevated soluble tryptophan. Plant Physiology **110:**51-59.

Li, J. J. 2003. Baeyer-Villiger Oxidation p. 14-15, Name reactions: a collection of detailed reaction mechanisms Third Expanded Edition ed. Springer Berlin.

Li, Q. S., J. Ogawa, R. D. Schmid, and S. Shimizu. 2005. Indole hydroxylation by bacterial cytochrome P450 BM-3 and modulation of activity by cumene hydroperoxide. Bioscience Biotechnology and Biochemistry **69:**293-300.

Liles, J. N. 1990. The art and craft of natural dyeing: traditional recipes for modern use. The University of Tennessee Press, Tennessee.

Lorenz, M. G., and W. Wackernagel. 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiological Reviews. 1994 Sep;58(3):563-602. **58:**563-602.

Ludwig-Muller, J., K. Pieper, M. Ruppel, J. D. Cohen, E. Epstein, G. Kiddle, and R. Bennett. 1999. Indole glucosinolate and auxin biosynthesis in *Arabidopsis thaliana* (L.) Heynh. glucosinolate mutants and the development of clubroot disease. Planta **208:**409-419.

Maugard, T., E. Enaud, A. De La Sayette, P. Choisy, and L. M.D. 2002. β-Glucosidase-catalyzed hydrolysis of indican from leaves of *Polygonum tinctorium*. Biotechnology Progress **18:**1104-1108.

McKee, J. R., and M. Zanger. 1991. A microscale synthesis of indigo: vat dyeing. Journal of Chemical Education **68:**242-244.

Melanson, D., M. D. Chilton, D. Masters-Moore, and W. S. Chilton. 1997. A deletion in an indole synthse gene is responsible for the DIMBOA-deficient phenotype of *BxBx* Maize. Proceedings of the National Academy of Science USA **94:**13345-13350.

Mikkelsen, M., C. Hørslev Hansen, U. Wittstock, and B. A. Halkier. 2000. Cytochrome P450 CYP79B2 from *Arabidopsis* catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a precursor of indole glucosinolates and indole-3-acetic acid. The Journal of Biological Chemistry **275:**33712–33717.

Miles, E. W. 1991. Structural basis for catalysis by tryptophan synthase. Advanced Enzymology and Related Areas of Molecular Biology **64:**93-172.

Minami, Y., T. Kanafuji, and K. Miura. 1996. Purification and characterization of a β-glucosidase from *Polygonum tinctorium*, which catalyzes preferentially the hydrolysis of indican. Bioscience, Biotechnology, and Biochemistry **60:**147-149.

Minami, Y., H. Takao, T. Kanafuji, K. Miura, M. Kondo, I. Hara-Nishimura, M. Nishimura, and H. Matsubara. 1997. Beta-glucosidase in the indigo plant: intracellular localization and tissue specific expression in leaves. Plant Cell Physiology **38:**1069-1074.

Mitich, L. W. 1998. Pale smartweed (*Polygonum lapathifolium* L.) and other *Polygonums*. Weed Technology **12:**560-562.

Moore-Braun, J. 11 March 2010, posting date. BASF report 2009. BASF SE. [Online.]

Morant, M., S. Bak, B. L. Møller, and D. Werck-Reichhart. 2003. Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. Current Opinions in Biotechnology. **14:**151-162.

Nagano, H., H. Zalkin, and E. J. Henderson. 1970. The anthranilate synthetase-anthranilate-5-phosphoribosylpyrophosphate phosphoribosyltransferase aggregate on the reaction mechanism of anthranilate synthetase from *Salmonella typhimurium*. The Journal of Biological Chemistry **245:**3810-3820.

Nagendrappa, G. 2003. Chemistry triggered the first civil disobedience movement in India Resonance **8:**42-48.

Niemeyer, H. M. 1988. Hydroxamic acids (4-hydroxy-1,4-benzoxazin-3-ones), defence chemicals in the *gramineae*. Phytochemistry **27:**3349-3358.

Niyogi, K. K., and G. R. Fink. 1992. Two anthranilate synthase genes in *Arabidopsis*: Defense-regulated regulation of the tryptophan pathway. The Plant Cell **4:** 721-733.

Normanly, J., J. D. Cohen, and G. R. Fink. 1993. *Arabidopsis thaliana* auxotrophs reveal a tryptophan-independant biosynthetic pathway for indole-3-acetic acid. Proceedings of the National Academy of Science USA **90:**10355-10359.

Oberthur, C., B. Schneider, H. Graf, and M. Hamburger. 2004. The elusive indigo precursors in woad (*Isatis tinctoria* L.) - identification of the major indigo precursor, isatin A, and a structure revision of isatin B. Chemistry and Biodiversity **1:**174-182.

O'Connor, K. E., and S. Hartmans. 1998. Indigo formation by aromatic hydrocarbon-degrading bacteria. Biotechnology Letters **20:**219-223.

Palazzolo, M. J., B. A. Hamilton, D. Ding, C. H. Martin, D. A. Mead, R. C. Mierendorf, K. Vijay Raghavan, E. M. Meyerowitz, and H. D. Lipshhitz. 1990. Phage lambda cDNA cloning vectors for subtractive hybridization, fusion-protein synthesis and Cre-loxP automatic plasmid subcloning. Gene **88:**25-36.

Parales, R. E., K. Lee, S. M. Resnick, H. Jiang, D. J. Lessner, and D. T. Gibson. 2000. Substrate specificity of naphthalene dioxygenase: effect of specific amino acids at the active site of the enzyme Journal of Bacteriology **182:**1641-1649.

Parotta, J. A. 2001. Healing plants of peninsular India. CABI Publishing, Rio Piedras.

Pathak, H., and D. Madamwar. 2009. Biosynthesis of indigo dye by newly isolated naphthalene-degrading strain *Pseudomonas* sp. HOB1 and its application in dyeing cotton fabric. Applied Biochemistry and Biotechnology **160:**1616-1626.

Pollmann, S., P. Düchtinga, and E. W. Weilera. 2009. Tryptophan-dependent indole-3-acetic acid biosynthesis by 'IAA-synthase' proceeds via indole-3-acetamide. Phytochemistry **70:**523-531.

Powell, L. E. 1963. Preparation of indole extracts from plants for gas chromatography and spectrophotoflourimetry. Plant Physiology**:** 836-842.

Pratt, K. K., P. Kumar, and W. S. Chilton. 1995. Cyclic hydroxamic acids in dicotyledonous plants. Biochemical Systematics and Ecology **23:** 781-785.

Radwanski, E. R., and R. L. Last. 1995. Tryptophan biosynthesis and metabolism: biochemical and molecular genetics. The Plant Cell **7:**921-934.

Radwanski, E. R., J. Zhao, and R. L. Last. 1995. *Arabidopsis thaliana* tryptophan synthase alpha: gene cloning, expression, and subunit interaction. Molecular Genetics and Genomics **248:**657-667.

Ratledge, C. 1964. Relationship between the products of aromatic biosynthesis in *Mycobacterium smegmatis* and *Aerobacter aerogenes*. Nature **203:**428-429.

Reardon, D., and G.K. Farber. 1995. The structure and evolution of α/β barrel proteins. Federation of American Societies for Experimental Biology **9:** 497-503.

Reed, P. 1992. The British chemical industry and the indigo trade. The British Journal for the History of Science **25:**113-125.

Rose, A. B., and R. L. Last. 1994. Molecular genetics of amino acid, nucleotide and vitamin biosynthesis., p. 835-879. *In* E. M. Meyerowitz and C. R. Somerville (ed.), *Arabidopsis*. Cold Spring Harbor Press, New York.

Roychowdhury, P., and B. S. Basak. 1975. The crystal structure of indole. Acta Crystallographica **31:**1559-1563.

Russell, G. A., and G. Kaupp. 1969. Oxidation of indoxyl to indigo in basic solution. Journal of the American Chemical Society **91:**3851-3859.

Salvini, M., T. M. Boccardi, E. Sani, R. Bernardi, S. Tozzi, C. Pugliesi, and M. Durante. 2008. Alpha-tryptophan synthase of *Isatis tinctoria*: gene cloning and expression. Plant Physiology and Biochemistry **46:**715-723.

Sambrook, J., and D. W. Russell. 2001. Molecular Cloning: A Laboratory Manual, Third Edition ed, vol. 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sandberg, G. 1989. Indigo textiles and history. Lark Books, Asheville, North Carolina.

Satoshi, U., and K. Yuki. 2002. Various colour dyeing using fresh leaves of Japanese indigo plant. Nature Science Series **49:**55-58.

Sequin-Frey, M. 1981. The chemistry of plant and animal dyes. Journal of Chemical Education **58:**301-305.

Sharrer, G. T. 1971. The indigo bonanza in South Carolina. Technology and Culture **12:**447-455.

Shim, J. Y., Y. J. Chang, and S. U. Kim. 1998. Indigo and indirubin derivatives from indoles in *Polygonum tinctorium* tissue cultures. Biotechnology Letters **20:**1139-1143.

Smith, G. R. 1988. Homologous recombination in procaryotes. Microbiolical Reviews **52:**1-28.

Smolen, G., and J. Bender. 2002. *Arabidopsis* cytochrome P450 *cyp83B1* mutations activate the tryptophan biosynthetic pathway. Genetics **160:**323-332.

Solomons, G. T. W., and C. Fryhle. 2007. Organic chemistry, 9th ed. John Wiley and Sons, Miami.

Srinivasan, P. R. 1965. The biosynthesis of anthranilate from (3,4-^{14}C) glucose in *Eschericia coli*. Biochemistry **4:**2860-2865.

Steingruber, E. 2004. Indigo and indigo colorants, Ullmann's Encyclopedia of Industrial Chemistry. Wiley-VCH, Weinheim.

Stewart, V., R. Landick, and C. Yanofsky. 1986. Rho-dependent transcription termination in the tryptophanase operon leader region of *Escherichia coli* K-12. Journal of Bacteriology **166:**217-223.

Sun, H. C., J. K. Kim, E. H. Cho, Y. C. Kim, J. I. Kim, and S. W. Kim. 2003. A novel flavin-containing monooxygenase from *Methylophaga* sp. strain SK1 and its indigo synthesis in *Escherichia coli*. Biochemical and Biophysical Research Communications **306:**930-936.

Warzecha, H., A. Frank, M. Peer, E. M. Gillam, F. P. Guengerich, and M. Unger. 2007. Formation of the indigo precurser indican in genetically engineered tobacco plants and cell cultures. Plant Biotechnology **5:**185-191.

Watanabe, T., and E. E. Snell. 1972. Reversibility of the tryptophanase reaction: synthesis of tryptophan from indole, pyruvate, and ammonia. Proceedings of the National Academy of Science USA **69** 1086-1090.

Wilhelm, P., B. Friguet, Djavadi-Ohaniance., I. Pilz, and M. Goldberg. 1987. Epitope localization in antigen-monoclonal-antibody complexes by small-angle X-ray scattering: An approach to domain organization in the β2 subunit of *Escherichia coli* tryptophan synthase. European Journal of Biochemistry **164:**103-109.

Winans, S. C., S. J. Elledge, J. H. Krueger, and G. C. Walker. 1985. Site-directed insertion and deletion mutagenesis with cloned fragments in *Escherichia coli*. Journal of Bacteriology **161:**1219-1221.

Wright, A. D., C. A. Moehlenkamp, G. H. Perrot, M. G. Neuffer, and K. C. Cone. 1992. The maize auxotrophic mutant orange pericarp is defective in duplicate genes for tryptophan Synthase beta. The Plant Cell **4:**711-719.

Xia, Z., and M. H. Zenk. 1992. Biosynthesis of indigo recursors in higher plants. Phytochemistry **31:**2695-2697.

Yamada, T., Y. Kambara, H. Imaishi, and H. Ohkawa. 2000. Molecular cloning of novel cytochrome P450 species induced by chemical treatments in cultured tobacco. Pesticide Biochemistry and Physiology **68:**11-25.

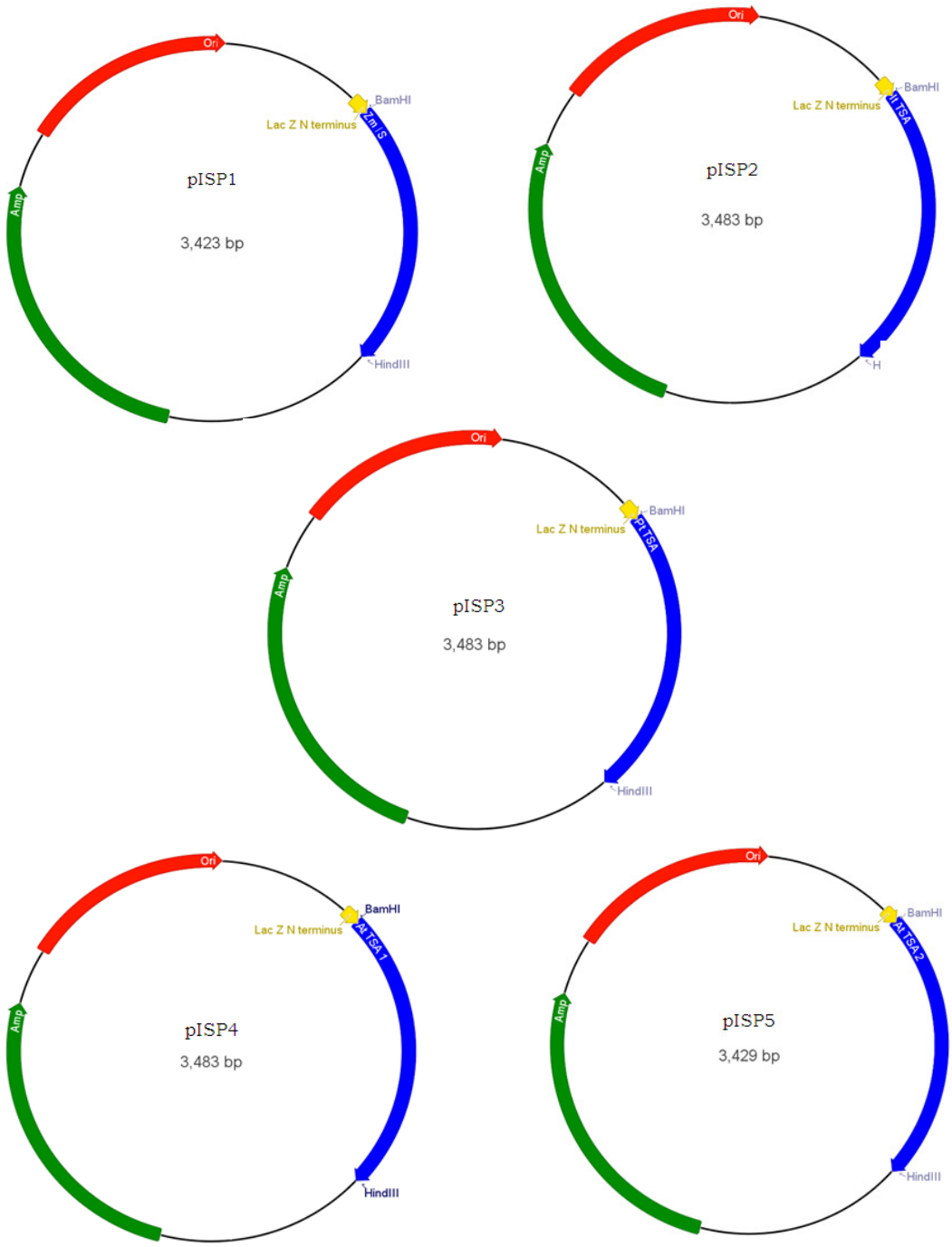
Zhang, R., B. Wang, J. Ouyang, J. Li, and Y. Wang. 2008. *Arabidopsis* indole synthase, a homolog of tryptophan synthase alpha, is an enzyme involved in the trp-independent indole containing metabolite biosynthesis. Journal of Integrative Plant Biology. **50:** 1070-1077.

Zhao, J., and R. L. Last. 1995. Immunological characterization and chloroplast localization of the tryptophan biosynthetic enzymes of the flowering plant *Arabidopsis thaliana*. Journal of Biological Chemistry. **270:**6081-7.

Zhao, J., and R. L. Last. 1995. Coordinate regulation of the tryptophan biosynthetic pathway and indolic phytoalexin accumulation in *Arabidopsis*. The Plant Cell. **8:** 2235-2244.

Zou, P., and H. L. Koh. 2007. Determination of indican, isatin, indirubin and indigotin in *Isatis indigotica* by liquid chromatography/electrosparay ionization tandem mass spectrometry. Rapid Communications in Mass Spectroscopy **21:**1239-1246.

7 Appendix

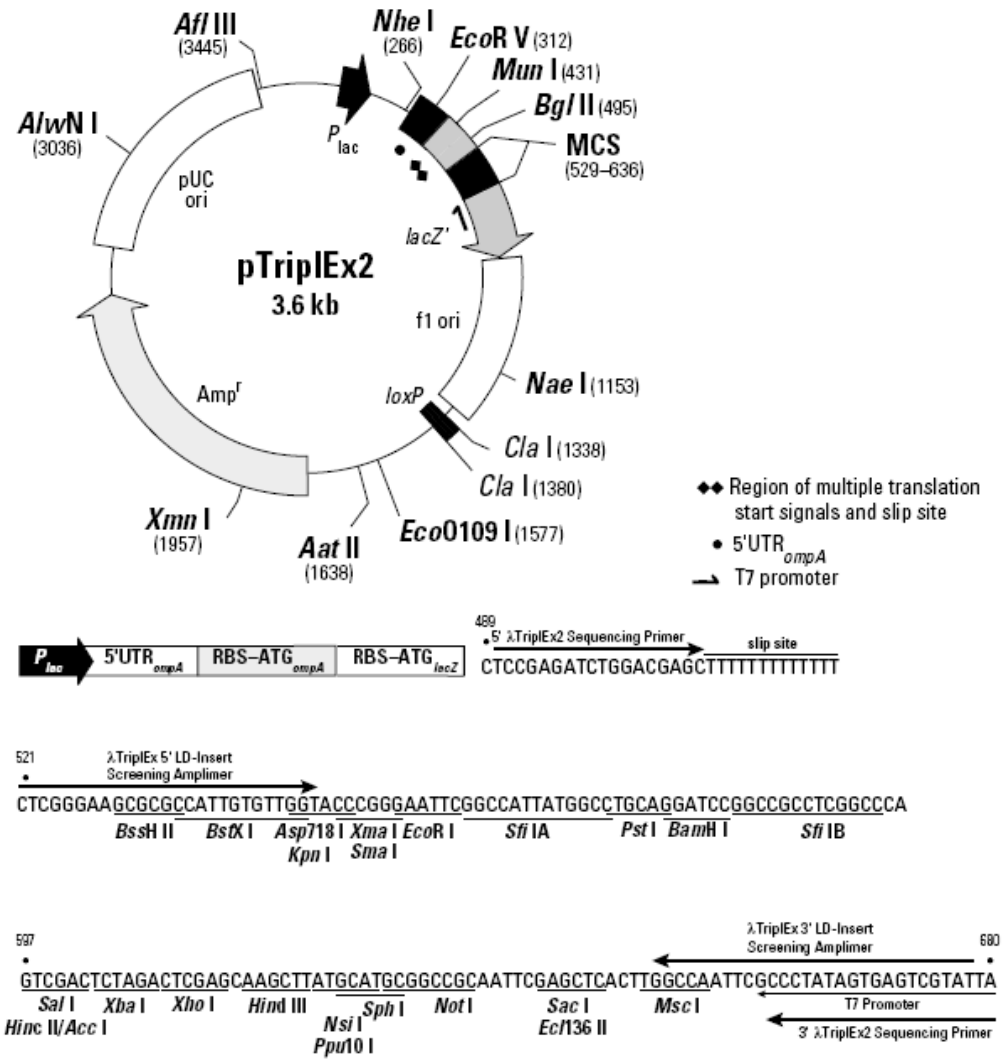


Plasmid maps of pISP1 - pISP5

Example of BLASTn analysis for cDNA library positive control. Sample 13 displayed 99% homology at nucleotide level to the published β-glucosidase mRNA from *Polygonum tinctorium*

```
Sample 13-Polygonum tinctorium mRNA for beta-glucosidase, complete cds
Length=1794

 Score = 1977 bits (1070),  Expect = 0.0
 Identities = 1087/1095 (99%), Gaps = 2/1095 (0%)
 Strand=Plus/Plus

Query  1     AATGGAATCAAGCCACTTGTCACTATCTACCATTGGGACCTTCCACAAGCACTTCAAGAC  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  480   AATGGAATCAAGCCACTTGTCACTATCTACCATTGGGACCTTCCACAAGCACTTCAAGAC  539

Query  61    GAATATGGAGGGTTCTTGAGCCCCAAAATCGTGGATGACTTTCTGGAATATGCAAACCTA  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  540   GAATATGGAGGGTTCTTGAGCCCCAAAATCGTGGATGACTTTCTGGAATATGCAAACCTA  599

Query  121   GTTTTTAAGGAGTTCGGGGATAGGGTTAAGCATTGGGCGACACTGACCGAACCCAATATA  180
             ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Sbjct  600   GTTTTTAAGGAGTTCGGGGATAGGGTTAAGCATTGGGCGACACTGAATGAACCCAATATA  659

Query  181   ATGACCCAACAAGGGTACGTATTTGGGGCACATGCACCCGGACGATGTTCTCACTTCGAA  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  660   ATGACCCAACAAGGGTACGTATTTGGGGCACATGCACCCGGACGATGTTCTCACTTCGAA  719

Query  241   TGGAACTGCCCGGCTGGAAACTCCGGCACCGAGCCTTATATAGTTGGTCACCACCTCCTC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  720   TGGAACTGCCCGGCTGGAAACTCCGGCACCGAGCCTTATATAGTTGGTCACCACCTCCTC  779

Query  301   CTATGTCATGCTGCAGCTTTTCAACTATACAAACAAAAGTATAAGGATGATCAAAAGGGT  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  780   CTATGTCATGCTGCAGCTTTTCAACTATACAAACAAAAGTATAAGGATGATCAAAAGGGT  839

Query  361   ATAATCGGAATAACAACCGCGACACAGATGGCCATACCGTTAAACGACAACGTTGCCAAC  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  840   ATAATCGGAATAACAACCGCGACACAGATGGCCATACCGTTAAACGACAACGTTGCCAAC  899

Query  421   CTCTTGGCAGCGTCACGATCCATCGATTTCAACATTGGATGGTTTTTGCATCCGGTTGTT  480
             |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Sbjct  900   CTCTTGGCAGCGTCACGAGCCATCGATTTCAACATTGGATGGTTTTTGCATCCGGTTGTT  959

Query  481   TACGGCGAGTATCCACAGACGATGAGGGAGCGGTTGGGAAGTCGACTGCCAAAATTCACA  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  960   TACGGCGAGTATCCACAGACGATGAGGGAGCGGTTGGGAAGTCGACTGCCAAAATTCACA  1019

Query  541   GAAAAAGAGTCGGAGATGTTGAAACAATCGTTCGACTTTATAGGGTTGAATTACTACTCA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1020  GAAAAAGAGTCGGAGATGTTGAAACAATCGTTCGACTTTATAGGGTTGAATTACTACTCA  1079

Query  601   ACTGATTATGCAGCCGCATCATCTTTTTCAGTTGATCCAGTGAATGTCAGTTACACAACT  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1080  ACTGATTATGCAGCCGCATCATCTTTTTCAGTTGATCCAGTGAATGTCAGTTACACAACT  1139

Query  661   GATTCCCGAGCAACATTATCAGCGATAAAAGATGGGACACCTATCGGCGACCCGACATTT  720
             |||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Sbjct  1140  GATTCCCGAGCAACATTATCAGCGATAAAAGATGGGGTTCCTATCGGCGACCCGACATTT  1199
```

```
Query  721   ATGAGCTGGTTGCATATATATCCAGAGGGCATCCTAACTCTGTTGCGATACGTAAAGGAA  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1200  ATGAGCTGGTTGCATATATATCCAGAGGGCATCCTAACTCTGTTGCGATACGTAAAGGAA  1259

Query  781   AGGTACAACAATCCATTTGTCATGA-CAACTGAGAATGGGATGGCCGATGAAAACAAGGG  839
             |||||||||||||||||||||||||| | ||||||||||||||||||||||||||||||||
Sbjct  1260  AGGTACAACAATCCATTTGTCATGATC-ACTGAGAATGGGATGGCCGATGAAAACAAGGG  1318

Query  840   ATCATTAGCGGAAGATCCGATGGCTTTAAAAGACAACGTCAGAATTCGATATCACCGCGA  899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1319  ATCATTAGCGGAAGATCCGATGGCTTTAAAAGACAACGTCAGAATTCGATATCACCGCGA  1378

Query  900   ACATCTATACTATGTTCTTGAAGCTATAAAGGAGGGTGTGAACGTGGGAGGATACTACGC  959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1379  ACATCTATACTATGTTCTTGAAGCTATAAAGGAGGGTGTGAACGTGGGAGGATACTACGC  1438

Query  960   ATGGACATGGATGGATGATTTCGAGTGGGGTTCTGGATATACTCCTCGATTCGGTCTCAA  1019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1439  ATGGACATGGATGGATGATTTCGAGTGGGGTTCTGGATATACTCCTCGATTCGGTCTCAA  1498

Query  1020  CTTTGTGGATTTCGACAATGATTTGAAGAGAACCCCCAAGGATTCTTACTTCTGGTTCAA  1079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1499  CTTTGTGGATTTCGACAATGATTTGAAGAGAACCCCCAAGGATTCTTACTTCTGGTTCAA  1558

Query  1080  GGACTTCCTTGCAAA  1094
             |||||||||||||||
Sbjct  1559  GGACTTCCTTGCAAA  1573
```



Plasmid map of *pTriplEx2* with MCS.