Chapter 1

Introduction

The rose is steeped in mystery and legend. Since ancient times roses were symbols of earthly passion and of heavenly perfection, symbols of the beauty of love and of the agony of unrequited affection, even symbols of wisdom unfolding in initiates of esoteric societies. Mythology tells of how the first red roses grew from the blood of Venus, and of how the thorns serve as a reminder of the expulsion from Paradise. But throughout the ages there were two mythical rose varieties that could be found only in the realms of fairytales and legends, namely blue roses and black roses...

After centuries of fruitless attempts to find a blue rose, it has entered popular imagination as the very symbol of the magical and the unattainable. Many modern fairytales, especially those based on the *Arabian Nights*, are woven around blue roses[1]. And in David Eddings' *Tumuli* novels, the fate of the world rests upon an ancient, magical artefact called the Blue Rose.

The black rose, in contrast, is seldom found in fairytales, but often in tales of gothic horror. There it evokes images of death and dark magics, epitomized by the evil Lord Soth, Knight of the Black Rose, in the *Dragonlance Legends*. Understandably, the black rose is seen by many as the pre-eminent symbol of the entire gothic genre.

If, some day, a blue or black rose is created, it will truly be the bringing to life of fantasy and legend. Moreover, such a discovery will even be important for people who are more interested in mundane things such as money than in fairytales and legends. The rose is the most important ornamental crop in the world, with the rose cut-flower market being worth several milliard US$ per year at retail level (Holton and Tanaka, 1994). Novelty is a very important factor in the cut-flower market, even more than improved production characteristics, so the creation of blue or black roses is expected to capture a niche market of considerable value (Chandler, 1996). Furthermore, blue and black flowering varieties are also missing from other important ornamental flowers such as tulip, chrysanthemum, carnation and gerbera (Elomaa and Holton, 1994). These flowers are not reputed to be imbued with such legendary magical powers as the rose, but the creation of blue or black varieties will still be financially most worthwhile.

The quest for blue and black roses has been disappointing to date. Countless ages of mutation, selection and hybridization have led to the development of modern rose varieties

[1] Contrary to what is claimed in several articles and books, the blue rose is neither mentioned in Rimsky-Korsakov's fairy tale opera, *Sadko*, nor in the *Arabian Nights*, *Alf Layla wa Layla*.

such as 'Black Pearl', 'Black Madonna' and 'Blue Moon', but the names are a bit of an exaggeration (Figure 1.1). The flowers of the so-called black varieties are indeed nearly black during early budding stages, but then turn a dark, rusty red. The flowers of the so-called blue roses are actually lilac or mauve. Presumably the occasional development of new varieties with novel colours kept the hope alive that true blue or black flowering varieties could be produced through breeding. Only in 1929, for example, was a new variety found that produced orange flowers (Courtney-Gutterson, 1994). Before 1929, garden roses produced only flowers in a range of red shades. But all hopes of breeding blue roses, chrysanthemums or carnations through traditional practices were dashed when it was discovered that these flowers do not even possess the gene necessary for the formation of blue pigments (Elomaa and Holton, 1994).

Figure 1.1: *Rosa hybrida* varieties Deep Secret (left, http://www.justourpictures.com/roses/ALroses.html) and Blue Moon (right, http://www.swanes.com/guide-1-bush.shtml) as examples of so-called black roses and blue roses.

Hope soon reappeared on the horizon. With the advent of genetic engineering came previously undreamed-of possibilities in plant breeding. Using the techniques of genetic engineering, plant breeders can now move beyond the traditional breeding practices of hybridization and selection, and instead change the genetic make-up of plants directly. This application of genetic engineering to flower breeding was termed 'molecular flower breeding' to distinguish it from the traditional breeding practices. So, with renewed enthusiasm, researchers endeavored to create flowers with novel colours, prolonged vase life and improved disease resistance. And there were undoubtedly some spectacular success stories, such as the creation of orange petunias (Meyer *et al.*, 1987) and mauve carnations (Holton, 2000). But, alas, the blue or black roses, carnations, tulips and chrysanthemums remained as elusive as ever.

All these past attempts to create novel-coloured flowers have focused on the manipulation of the flavonoid pigmentation pathways that exist in flowers (Elomaa and Holton, 1994; Forkmann and Martens, 2001). To create black flowers, for example, researchers tried to increase the concentrations of flavonoid pigments to very high levels. Few of these attempts gave any encouraging results, and many actually seem to point to insurmountable problems due to gene silencing (van der Krol *et al.*, 1990; Jorgensen *et al.*, 1996; Koes *et al.*, 2000). To create blue flowers, researchers planned to transfer the genes responsible for the formation of blue pigments to species lacking those genes. After many years of study, the researchers finally succeeded in finding these genes, and then introduced them into the rose and other important ornamental crops (Holton and Tanaka, 1999; Holton, 2000). Only to discover to their utter dismay that the pigment appeared purple instead of blue! This was found to be due to the relatively acidic vacuolar pH in the flowers of these plants. However, the intrepid researchers were not discouraged, and are now valiantly trying to increase the vacuolar pH of these ornamental plants. Whether this new strategy will succeed and how long it will take, only time will tell.

But perhaps the key to the creation of blue or black flowers does not lie in the manipulation of flavonoid pigmentation pathways. One of the great advantages of molecular flower breeding over traditional breeding is that plant breeders are no longer fettered to only those biosynthetic pathways already existing in flowers. Instead, novel pigmentation pathways can be established, using a combination of enzymes found in various diverse, unrelated organisms. These novel pigmentation pathways can then lead to the expression of blue or black pigments that have never before been found in any flower.

The aim of this study is, therefore, to take the first small steps in the investigation of the use of such novel pigmentation pathways. Firstly, various pigmentation pathways will be discussed and evaluated. The two most promising pigmentation pathways can then be chosen – one for blue flowers and one for black flowers. Of course, even the most promising pathways do not always work as intended. The only way to know with certainty that these pathways will indeed produce brilliant blue or pitch-black flowers is to express the pathways in actual rose, tulip or carnation flowers. However, the tissue culture and physiology of these popular garden flowers are not yet as well-known as that of model plants such as *Arabidopsis*, tobacco or petunia. So, in this study, the activity of the novel pathways will only be studied in tobacco. This should give a good indication of the feasibility of the use of these novel pathways in plants, while potential problems should be less troublesome to sort out. All the genes involved in the pathways will therefore be characterized, placed in plant expression

vectors and introduced into tobacco. Lastly, the expression of the genes and activity of the gene products will be evaluated.

hapter 2

Literature review

2.1 The heroic past of molecular flower breeding: flavonoid pigments

2.1.1 The four classes of flower pigments

The wondrous variety of colours displayed by flowers is due to four structurally distinct classes of pigments, namely flavonoids, carotenoids, betalains and chlorophyll. Chlorophyll is very rare as a flower pigment, and is only present in the handful of flowers that show a slight green hue (Eugster and Märki-Fischer, 1991). Betalains also have a rather limited distribution, as they are found exclusively in the Caryophyllales, a group of angiosperms that include beetroot (Stafford, 1994). Despite this limited distribution, betalains can assume a wide range of colours from ivory and yellow to orange, red and violet. The other two classes of pigments are more widespread and often occur together in the same flower, in many cases producing very similar colours. Carotenoids are yellow, orange or red, and account for the majority of yellow and orange hues in flowers. Flavonoids are the most widespread and important pigments, found in all vascular plants. They contribute the greatest variety of colours ranging from creamy yellow to red, purple and – very importantly – also blue. Furthermore, extremely high expression of certain flavonoid pigments can even cause flowers to appear pitch black (Elomaa and Holton, 1994).

2.1.2 Biosynthesis of flavonoid pigments

The flavonoid biosynthetic pathway has received the most attention in genetic engineering of flower colour, and is much better known than that of other flower pigments (reviewed in Gerats and Martin, 1992; Courtney-Gutterson, 1994; Elomaa and Holton, 1994). Besides contributing to flower pigmentation, flavonoids also serve a number of other unrelated functions, such as signaling in symbiotic plant-microbe interactions, modulation of auxin responses, and protection against UV light and pathogens. Despite this wide variety of functions, all flavonoids have the same basic structure consisting of two aromatic C_6 rings (A and B) connected by a heterocycle (C) (figure 2.1).

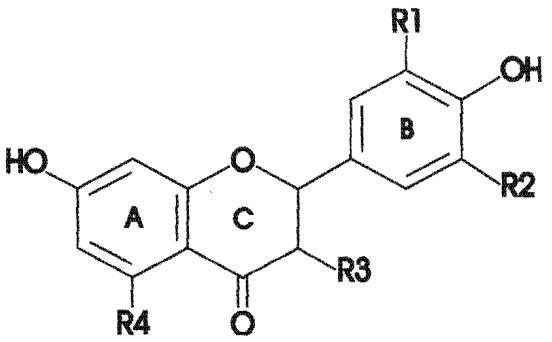

Figure 2.1: Basic structure of flavonoids, consisting of two aromatic rings (A and B) and a central heterocycle (C). Positions at which various additions on the basic skeleton take place are indicated by R1-R4. (van Tunen and Mol, 1991)

On the basic C_{15} skeleton various oxidations, additions and rearrangements occur, which determine the flavonoid class formed. In this way anthocyanins, aurones, flavonones, flavones, isoflavonoids and flavonols can be distinguished. In flowers the main flow of flavonoid biosynthesis is usually towards the anthocyanin pigments, which include the orange to brick red pelargonidin-3-glucoside, the red cyanidin-3-glucoside, and the purple to blue delphinidin-3-glucoside. However, flavonoids belonging to other classes are produced in a number of cases, and they can also play a role in flower pigmentation.

Flavonoids are synthesized via the complex phenylpropanoid pathway, starting from phenylalanine, as shown in figure 2.2. The first reaction committed to flavonoid biosynthesis is catalyzed by chalcone synthase (CHS), which condenses three malonyl-CoA residues with the CoA ester of a hydroxycinnamic acid such as coumaroyl-CoA, caffeoyl-CoA or feruloyl-CoA, and leads to the formation of yellow chalcones. The enzyme chalcone-flavanone isomerase (CHI) then catalyzes the isomerization of chalcones to colourless flavanones. The flavanones are converted to dihydroflavonols, also colourless, by flavanone 3-hydroxylase (F3H). Although several different dihydroflavonols can be synthesized by F3H, only dihydrokaempferol, which is derived from coumaroyl-CoA, serves as the precursor to the anthocyanin pathway.

The next reaction is one of the key steps in anthocyanin synthesis, as it determines the type of anthocyanin that can ultimately be produced. If flavonoid 3'-hydroxylase (F3'H) is present, dihydrokaempferol will be hydroxylated at the 3' position of the B-ring to form another dihydroflavonol, dihydroquercitin, which may eventually become cyanidin-3-glucoside. However, if flavonoid 3',5'-hydroxylase (F3'5'H) is present, dihydrokaempferol – and also dihydroquercitin – will be hydroxylated at both the 3' and 5' positions to produce dihydromyricetin, which can be converted to delphinidin-3-glucoside. If neither F3'H nor

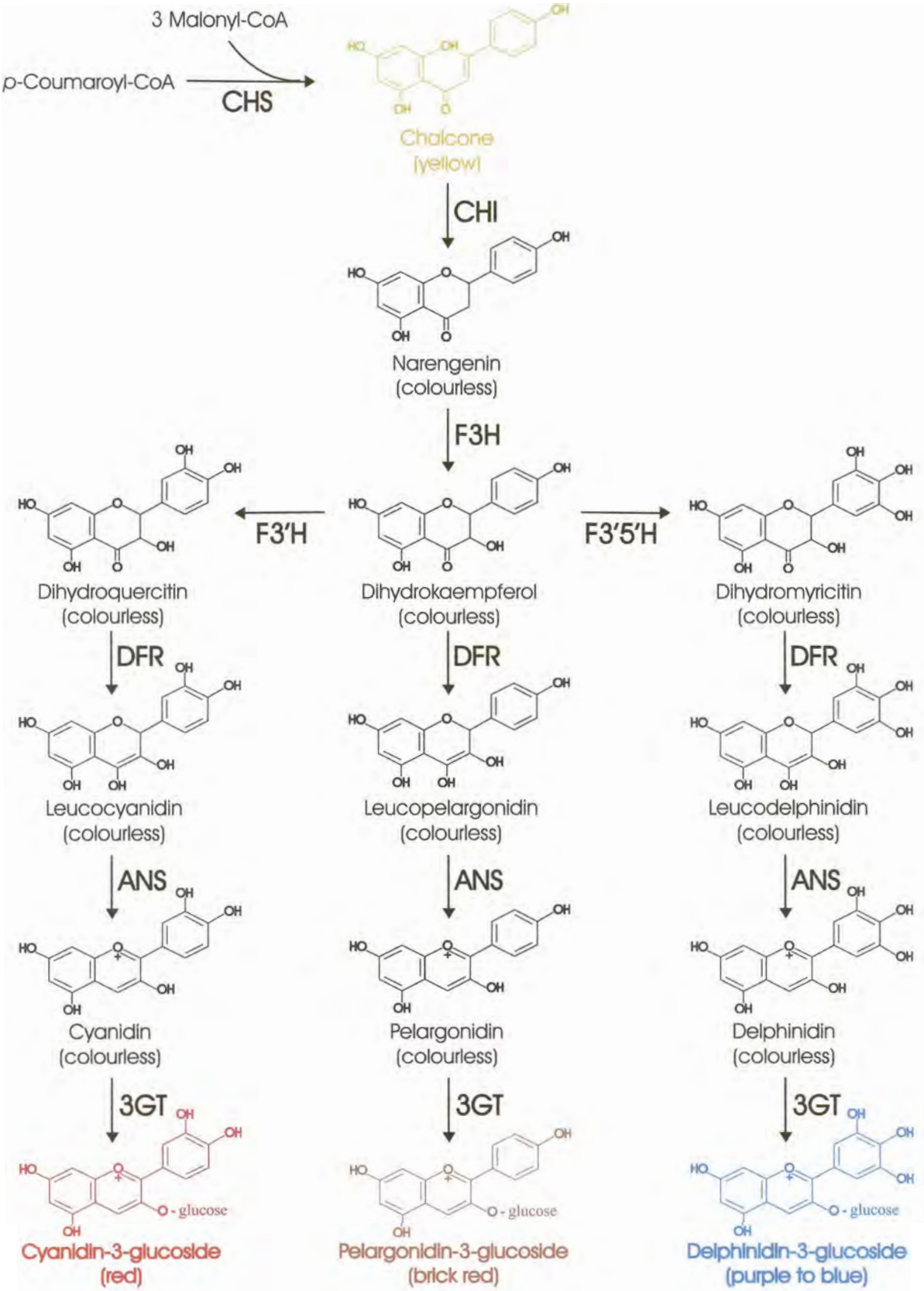

Figure 2.2: Outline of the anthocyanin biosynthetic pathway (based on Elomaa and Holton, 1994; Holton and Tanaka, 1994)

F3'5'H are present, dihydrokaempferol will remain unhydroxylated and become pelargonidin-3-glucoside.

The conversion of the colourless dihydroflavonols into coloured anthocyanins requires at least three further enzymatic modifications. Any of these enzymes may show substrate specificity, favouring the precursors of one anthocyanin and thereby inhibiting the synthesis of other anthocyanins. The presence or absence of F3'H and F3'5'H is therefore not the only factor that determines which of the anthocyanins are actually produced. In the first step committed to anthocyanin synthesis, the dihydroflavonols are reduced to the corresponding leucoanthocyanidins by dihydroflavonol 4-reductase (DFR). The leucoanthocyanidins are converted by anthocyanidin synthase (ANS) to unstable anthocyanidins. These are then stabilized by a glycosylation at the 3 position of the C-ring, catalyzed by UDPG-flavonoid-3-*O*-glucosyltransferase (3GT), to finally become the anthocyanins pelargonidin-3-glucoside, cyanidin-3-glucoside or delphinidin-3-glucoside.

The anthocyanins are often modified even further by glycosylation, acylation and methylation. Among the sugars that are often added to anthocyanins during glycosylation are rhamnose, xylose, glucose and galactose. It has been proposed that the formation of a 3,5-bisglucoside might actually be an essential prerequisite for blue colour, although the shade of blue will be different for different modifications (Courtney-Gutterson, 1994). Methylation, on the other hand, has a reddening effect on colour and results in peonidin, petunidin and malvidin from cyanidin, pelargonidin and delphinidin, respectively (Forkmann, 1991). The extent of modification, and the types of glucosides and acyl groups attached vary among different species and even varieties, determining the exact type of anthocyanin derivative that is eventually produced (van Tunen and Mol, 1991).

The chemical environment in which the anthocyanins are found can also cause considerable structural changes in the anthocyanins, resulting in altered colours. These conditions vary from cell to cell, are subject to developmental programmes, and include vacuolar pH values, co-pigmentation and interactions with metals. In general a higher pH leads to a bluer colour (van Tunen and Mol, 1991). This effect is especially striking if delphinidin-3-glucoside is present: at a pH near or higher than neutrality, delphinidin-3-glucoside gives a blue colour, while a purple colour is found at a pH of 5.5 or less (de Vlaming *et al.*, 1983). Secondly, metal ions such as iron, aluminium and magnesium can form complexes with anthocyanins and modify colour. Although the formation of such complexes is necessary for blue colour in a few species, it does not generally occur in blue-flowered species (Goto and Kondo, 1991; Harborne and Williams, 2000). Lastly, the presence of co-pigments enhances the stability of

the anthocyanin chromophore and may also modify its colour, even though the co-pigment is itself colourless (Aida *et al.*, 2000; Harborne and Williams, 2000). Various substances are known to serve as co-pigments, many of them being polyphenols or flavonoids. When compared over a range of pH values, co-pigmented anthocyanins are usually bluer than anthocyanins alone, and the relative amounts of anthocyanin to co-pigments can therefore have a strong effect on colour. In fact, certain blue-flowered species such as cornflower and Morning Glory do not even contain delphinidin derivatives. Their blue colour is largely due to intensive co-pigmentation of cyanidin and peonidin glucosides, combined with high vacuolar pH and metal chelation (Harborne and Williams, 2000).

Although all the enzymes of anthocyanin biosynthesis are cytosolic, floral anthocyanins are known to accumulate in the vacuoles of petal epidermal cells. Conditions inside the vacuole are therefore of utmost importance in determining the exact structure and colour of anthocyanins. Recent evidence suggests that anthocyanins are tagged for transport into the vacuole by glutathione S-transferases, which conjugate the glutathione tripeptide (γ-Glu-Cys-Glu) to a broad variety of herbicides and other toxic heterocyclic compounds (Marrs *et al.*, 1995; Alfenito *et al.*, 1998). The phenolic group of anthocyanins is potentially toxic (Gerats and Martin, 1992), explaining why anthocyanins are among the few endogenous substrates of plant glutathione S-transferases. The glutathionated anthocyanins are then sequestered in the vacuole via a tonoplast Mg-ATP-requiring glutathione pump, which actively transports glutathionated compounds into the vacuole. Finally, the glutathione moiety is clipped off as soon as the anthocyanins enter the vacuole, but the details of this reaction are still unclear (Mol *et al.*, 1998).

In addition to anthocyanins, other flavonoids can also play a role in floral pigmentation (van Tunen and Mol, 1991; Markham *et al.*, 2001). Yellow aurones are synthesized from chalcones, but little is known of the enzymatic reaction involved. Multiple enzymatic steps are involved in the formation of isoflavonoids, but again, these have only been poorly characterized. The flavonols kaempferol, quercitin and myricetin are synthesized from the corresponding dihydroflavonols by flavonol synthase and are important co-pigments. Flavones are synthesized from flavanones by flavone synthase and are known to be a substantial part of protective plant waxes, but also function as co-pigments. Flavonols and flavones are generally colourless, but can appear yellow under certain conditions. Some of the yellow colour seen in flowers is therefore caused by yellow-coloured aurones, chalcones, flavonols and flavones. The other flavonoids are colourless to man's eye but can play a role in the attraction of some insects that perceive ultraviolet colours (Kevan *et al.*, 2001).

2.1.3 Genetic engineering of the flavonoid pathway

As can be gathered from the above discussion, the flavonoid pathway is immensely complex and delicately intertwined with the surrounding conditions inside the cell. Some parts of the pathway, such as the formation of co-pigments and the modification of anthocyanins, are still known in only the barest outlines. These lesser-known parts do not form the most important, key steps of the pathway, but can still cause genes encoding the key steps to have somewhat unpredictable, pleiotropic effects. This makes the engineering of specific single, desirable traits more complicated than the random transformation attempts in which any new, unusual colour is deemed a success.

Yet there have been many reports of the successful development of novel-coloured flowers through genetic engineering. Several different strategies have been employed, some with more success than others. Most of the successful attempts involved a loss-of-function strategy in which the expression of a gene of the anthocyanin pathway is blocked, using either antisense or sense suppression. By suppressing these genes, anthocyanin synthesis can be inhibited and a range of paler colours produced (Figure 2.3). To date, suppression of gene expression at several different steps of the pathway has been successful in important ornamental species such as rose (Firoozabady *et al.*, 1994), carnation (Elomaa and Holton, 1994), gerbera (Elomaa *et al.*, 1993; Figure 2.3A) and chrysanthemum (Courtney-Gutterson *et al.*, 1994; Figure 2.3B), and has become almost routine in model plants such as petunia and tobacco (van der Krol *et al.*, 1990; Jorgensen *et al.*, 1996 and the references therein). The results of these studies were not entirely predictable and the flowers of the transformants often had random blotches of colour. Still, the colours of some of the transformants were stable and attractive enough to be commercially viable, and Plant Variety Rights have been granted for several of these (Elomaa and Holton, 1994).

Another possible strategy is to introduce genes of heterologous origin, encoding enzyme activities missing from the target plant. This would allow formation of novel pigments not naturally found in this species. The very first case of flower colour manipulation by genetic engineering involved such an approach, in which a maize DFR was transferred into petunia (Meyer *et al.*, 1987). It has been found that the petunia DFR does not accept dihydrokaempferol as a substrate, explaining why orange, pelargonidin-containing petunias do not occur naturally. The maize DFR, however, is capable of producing pelargonidin by reducing dihydrokaempferol. The maize DFR was therefore transferred into a petunia variety

that accumulated dihydrokaempferol, and pelargonidin-containing, brick red flowers were obtained (Figure 2.4).

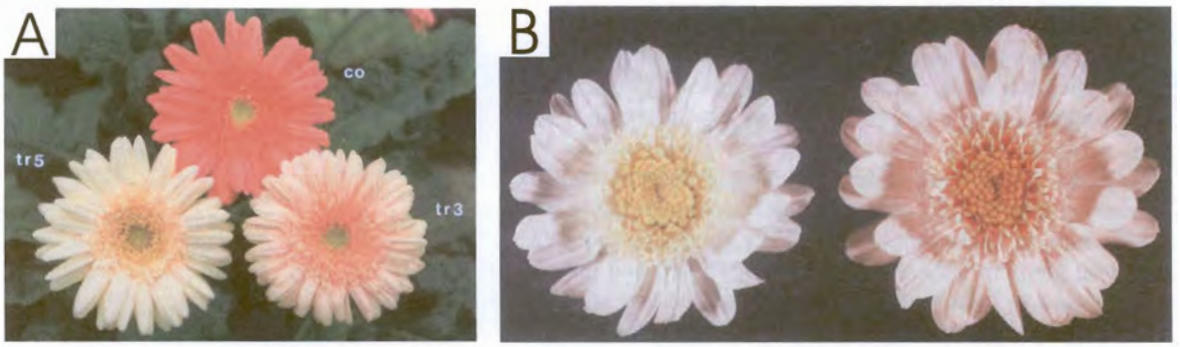

Figure 2.3: Examples of successful inhibition of pigmentation in ornamental species. (A) Comparison of gerbera var. Terra Regina (top) with transgenic flowers in which *chs* has been suppressed (tr3 and tr5) (Elomaa *et al.*, 1993). (B) Comparison of chrysanthemum var. Moneymaker (right) with a transgenic flower in which *chs* has been suppressed (left) (Courtney-Gutterson *et al.*, 1994).

Figure 2.4: The very first example of flower colour manipulation by genetic engineering. Maize *dfr* was transferred into a pale pink petunia variety that accumulated dihydrokaempferol (left), giving pelargonidin-containing, brick red flowers (right) (Meyer *et al.*, 1987).

An interesting variation on this strategy would be to introduce genes that encode regulatory proteins such as C1 and R (Gerats and Martin, 1992). When these genes are placed under tissue-specific promotors, the complete anthocyanin biosynthesis route may be turned on in previously acyanic plant tissues. This may even be used to produce fruits and vegetables with unusual new colours.

2.1.3.1 Molecular flower breeding for black flowers

Black flowers may be developed through genetic engineering if a way can be found to raise the concentration of anthocyanin pigments sufficiently high. Although the colours of the anthocyanin pigments actually range from blue to red and orange, they appear black if present in very high concentrations, as can be seen in the black areas of flowers such as pansy and cornflower. In maize it has been found that high concentrations of either cyanidin or pelargonidin glucosides can cause kernels to appear black (Figure 2.5) (Elomaa and Holton, 1994). It is not yet known which of the other flavonoid pigments can also appear black when present in sufficiently high concentrations.

Figure 2.5: Maize kernels showing the dark colour that can be achieved through very high expression of anthocyanins (Bruckner *et al.*, 1998)

Several studies have been done to intensify the expression of anthocyanin pigments in flowers – though not necessarily with the aim of producing black flowers. Using a loss-of-function approach, pigmentation can be intensified through suppression of enzymes that subvert intermediates of the anthocyanin pathway away from anthocyanin synthesis. In one example of this approach, levels of anthocyanin three times higher than normal were achieved in petunia and tobacco by suppression of flavonol synthase (Figure 2.6A) (Holton *et al.*, 1993b).

Using a gain-of-function approach, increased pigmentation has been achieved in carnation flowers through the expression of a petunia cytochrome b_5 gene, *difF*, that has been found to enhance the activity of F3'5'H (de Vetten *et al.*, 1999; Koes *et al.*, 2000). In sad contrast, most attempts to intensify pigmentation by expression of extra copies of *chs* or *dfr* have been unsuccessful. Introduction of additional copies of such flavonoid genes into target plants has, in fact, most often led to the opposite effect – a marked reduction of pigmentation (Figure 2.7)

(van der Krol *et al.*, 1990; Jorgensen *et al.*, 1996 and references therein). The expression of both the transgene and the endogenous gene was silenced due to co-suppression. Only rarely was the transgene expressed and increased pigmentation achieved – and to date this has only occurred in pale flowers such as tobacco (Figure 2.6B) (Polashock *et al.*, 2002).

Figure 2.6: Examples of successful intensification of pigmentation. (A) Comparison of wild-type petunia (left) with transgenic petunia that shows higher levels of anthocyanin after suppression of flavonol co-pigment formation (right) (Holton *et al.*, 1993b). (B) Comparison of wild-type tobacco (left) with transgenic tobacco that expresses heterologous DFR (right) (Polashock *et al.*, 2002)

Figure 2.7: Examples of the reduction in pigmentation caused by over-expression of *chs* in the dark violet petunia variety shown on the far left (Napoli *et al.*, 1990).

2.1.3.2 Molecular flower breeding for blue flowers

Based on the biosynthetic pathway of anthocyanins as described above, there are four main requirements for producing blue colour. First, delphinidin-3-glucoside pigments must be produced through the action of flavonoid 3',5'-hydroxylase and subsequent enzymes in the pathway. Second, the pH of the vacuolar compartment in which anthocyanins are stored must be near 6.0 or higher. Third, sufficient co-pigments such as flavonols or flavone glycosides

must be present. Finally, appropriate covalent modification of the anthocyanin base, such as acylation, glycosylation or metal chelation, is necessary.

In many rose varieties, two of these four requirements are met. The anthocyanins are bisglucosylated, and large amounts of flavonols are produced (Eugster and Märki-Fischer, 1991). Unfortunately, no rose variety expresses F3'5'H, and in wide surveys of pigments in roses, no delphinidin-derived pigments could be found (Eugster and Märki-Fischer, 1991; Mikanagi *et al.*, 2000). Most of the so-called blue roses, which are in reality lilac or purple, contained only cyanidin 3,5-diglucoside, together with large amounts of flavonols. Also, the vacuolar pH in rose flowers is too low. The pH of outer epidermal cells has been found to range from 3.56 to 5.38, with lilac-coloured flowers having the highest pH values (Biolley and Jay, 1993).

The pigment biosynthesis in other important ornamental flowers such carnation, chrysanthemum, gerbera and tulip is not as well studied as in the rose. At least it is known that delphinidin-derived pigments are found in purple tulip varieties (van Raamsdonk, 1993), but that carnation and chrysanthemum, like the rose, lack the genetic capability to produce F3'5'H and delphinidin (Elomaa and Holton, 1994). Fortunately this seems to be the only enzyme of the anthocyanin pathway that is missing from these species. Feeding experiments with rose and carnation petals have demonstrated that these flowers contain all the enzymes necessary for converting dihydromyricetin, the product of F3'5'H, to delphinidin glucosides (Stich *et al.*, 1992; Holton and Tanaka, 1994).

The obvious solution to the problem of a missing F3'5'H enzyme is to introduce and express a F3'5'H gene from another species. The expression of a heterologous F3'5'H gene in a pelargonidin- or cyanidin-producing cultivar should divert the anthocyanin flower colour towards blue – if potential problems such as gene silencing and substrate competition with endogenous DFR and flavonol synthase are overcome. In 1993, Holton *et al.* (1993a) reported the isolation of two petunia genes, *Hf1* and *Hf2*, which each encode a F3'5'H. This quickly led to one of the most important success stories in the genetic engineering of blue colour: the creation of mauve carnations. Petunia F3'5'H and DFR genes were introduced and expressed in a white carnation variety that accumulated dihydrokaempferol, resulting in the production of delphinidin (Holton, 2000). The colour of the flowers ranged from pale violet to mauve, correlating with the level of expression of the petunia genes. Several transgenic lines with differing intensity of pigmentation were selected for commercialization, including the pale mauve Moondust™ and the deeper violet Moonshadow™ (Figure 2.8).

Figure 2.8: Trangenic carnations that were created by introduction of petunia F3'5'H and DFR genes into a white-flowered variety. The pale mauve variety Moondust™ is on the left, and the deeper violet Moonshadow™ on the right (http://www.florigene.com).

Solving the problem of a too acidic vacuolar pH may be more difficult. In a last-ditch effort to create blue roses, researchers are trying to find ways to raise vacuolar pH. In petunia, the pH of the vacuole is regulated by the action of at least seven *pH* genes (Mol *et al.*, 1998; van Houwelingen *et al.*, 1998). If the genes *ph1* – *ph7* are recessive, the pH of the vacuole is increased from 5.3 to 6.2, shifting the flower colour from red towards blue (van Houwelingen *et al.*, 1998; figure 2.9). The gene from one of these loci, *ph6*, has been cloned (Chuck *et al.*, 1993). Inactivation of this gene by mutation resulted in an increase of the pH of petal-cell extracts by 0.4 units. The mechanisms by which these *ph* genes regulate pH are still unknown, but they are already employed in strategies to produce blue flowers (Holton and Tanaka, 1994). The *ph6* gene of petunia will be used as a probe to identify its homologues in other ornamental species. The expression of *ph6*-homologous genes can then be eliminated through sense or antisense suppression, and this would hopefully raise the pH of the vacuole sufficiently for delphinidin-derived pigments to appear blue.

Figure 2.9: Comparison between wild-type petunia (A) and petunia with raised vacuolar pH (B). The higher pH causes a shift towards blue (Mol *et al.*, 1998).

2.2 The future of molecular flower breeding: novel pigmentation pathways

2.2.1 Choosing promising pigmentation pathways

Even given all the advances in genetic engineering, one cannot simply introduce any pigmentation pathway into a plant and expect to find bright blue and pitch-black flowers. There are still many potential problems that need to be kept in mind, as can easily be seen from all the failed attempts to create blue roses. For a novel pigmentation pathway to be useful, it should be applicable to situations where engineering of the flavonoid pathway have failed, such as the creation of the blue roses, or at least present fewer problems and intricacies in general situations. Based on previous studies in which transgenic plants were created, and the problems experienced in these studies, several criteria can be deduced according to which the more promising novel pigmentation pathways can be chosen.

First and foremost, the pigment and its biosynthetic pathway should be well known and characterized. Only if the pathway is well known can potential problems at each step in the biosynthesis of the pigment be evaluated. This will also avoid nasty surprises such as discovering only at the last moment that the pigment appears orange at vacuolar pH instead of blue. Furthermore, the pathway should be simple, encoded by as few genes as possible and consisting of as few steps as possible. The fewer steps in the pathway, the fewer potential places where problems can arise, and the easier it is to coordinate the expression of the genes.

Once a simple, well-characterized pigmentation pathway has been found, its potential can be evaluated in more detail. Each step in the gene expression process should be analyzed to find potential problems that might hinder successful expression of the transgenes. For example, problems may arise during transcription and translation due to the transgene DNA sequence. If a plant recognizes the transgenes as 'foreign', typically due to a different AT-content relative to the flanking DNA, the transgenes might be methylated and silenced (Stam *et al.*, 1997). Later, during post-transcriptional processing, certain sequences inside the transgene primary transcript might be incorrectly recognized as splice sites or polyadenylation sites (Perlak *et al.*, 1991). Translation is also likely to be inefficient if the codon usage of a transgene differs considerably from the preferred codons of the plant (Murray *et al.*, 1991). These problems most often arise if the transgene is from bacterial, fungal or animal origin, but may even occur if a gene from a dicot is expressed in a monocot, or *vice versa*. The transgene

sequences should therefore always be scrutinized beforehand to find potential problematic sequences. Since plant genomes typically contain many GC-rich isochores and show an overall preference for guanine or cytosine in the third position of codons (Murray *et al.*, 1991; Elomaa *et al.*, 1995), transcriptional gene silencing and inefficient translation would be minimized by choosing transgenes that also have high GC contents. As a last resort, the transgene sequences can even be altered using site-directed mutagenesis, as has been done to improve expression of *Bacillus thuringiensis* endotoxins (Perlak *et al.*, 1991; Estruch *et al.*, 1997).

Post-transcriptional gene silencing might also present a problem if the transgene is expressed at too high levels, especially if the transgene is homologous to an endogenous plant gene (Stam *et al.*, 1997). Therefore the novel enzymes should preferably not be homologous to any endogenous protein, and should be relatively active so that the transgenes do need not to be expressed too strongly.

After translation, the newly formed polypeptide should move into the same subcellular compartment as its substrate. It is essential that the subcellular localization of the substrate be known. Only then can correct targeting of the polypeptide be ensured by adding necessary targeting sequences to the transgene, or by removing unwanted targeting sequences. Of course, the enzyme must be able to function properly in the targeted organelle. This consideration is typically most relevant if the enzyme is targeted to the vacuole, which has a relatively acidic pH, or if the enzyme is normally glycosylated but is now expressed intracellularly, where it will not be glycosylated.

The last step of gene expression involves the folding of the polypeptide into the correct three-dimensional structure. Protein folding is assisted first by molecular chaperones, which bind to hydrophobic peptide stretches in nascent and translocating polypeptides to prevent premature aggregation of the protein (Feldman and Frydman, 2000). The molecular chaperones seem to act in a very similar way across all taxa and normally do not present any complications. Some proteins, however, also need assistance from another class of chaperones called chaperonins. These are large, cylindrical protein complexes that enclose proteins or protein domains within their interior cavity to facilitate correct folding of the proteins. Chaperonins are found in all organisms, but the chaperonins of the eukaryotic cytosol differ from those of bacteria, mitochondria and chloroplasts (Carrascosa *et al.*, 2001). Although the two groups of chaperonins bind to approximately the same range of proteins, they cannot always assist the folding of the same proteins (Leroux and Hartl, 2000). The proteins that need assistance from chaperonins are typically large, multi-domain proteins with

prominent hydrophobic regions (Grantcharova *et al.*, 2001). Thus, if eukaryotic cytosolic enzymes are targeted to the plastids or mitochondria, or if bacterial enzymes are expressed in the plant cell cytosol, the enzyme should preferably be small, monomeric and mostly hydrophilic.

Some enzymes might need the assistance of yet another protein for the uptake of a prosthetic group, cofactor or even electrons. So, not only must the necessary prosthetic groups or cofactors be present in the environment of the novel enzyme, but also any assisting proteins that might be needed. A case in point is the flavonoid 3',5'-hydroxylases used to create mauve carnations (Holton *et al.*, 1993a). F3'5'H is a cytochrome P450 enzyme that has to associate with electron donors such as P450 reductases for transfer of electrons from NADPH to the heme group of the cytochrome P450 enzyme. Luckily these reductases were found to be rather unspecific and the carnation P450 reductase could easily activate the petunia F3'5'H. Other assisting proteins are not necessarily as accommodating and unspecific. Some, such as copper transfer factors, can only assist very specific target enzymes (Harrison *et al.*, 2000). In these cases, the assisting protein might have to be expressed alongside the novel enzyme.

The eventual fate of the pigment and all its precursors should also be examined in detail. Some of the more important questions that should be asked include: Does the normal cellular metabolism produce sufficient pigment precursors and other necessary cofactors? Will the reaction product of one step in the biosynthetic pathway be able to move into the appropriate subcellular compartment for the next step? Might the pigment or any precursor have toxic effects on the organelle in which they find themselves? Will the hapless pigment be degraded too quickly or perhaps be conjugated to some colour-destroying moiety? Is the pigment stable enough under various conditions so that it will stay the desired colour in the subcellular compartment it is stored in?

In the end, however, it must be remembered that the criteria and consideration given here are merely timesaving guidelines for choosing promising pathways, and not at all guarantees that the promising pathway will definitely produce beautiful pigments. Very few pathways are known in such detail that each step of the entire biosynthesis can be predicted with accuracy. And even the most promising and best thought-out pigment biosynthetic pathway can meet an untimely end at the hands of some novel, hitherto unknown inhibitor in the target plant. Any attempt to establish a novel pigment biosynthetic pathway will probably have to go through several rounds of trial and error, starting with the more promising pathways, trying different enzymes and targeting different organelles, perhaps later also trying the less promising pathways, until the elusive blue and black flowers are found at last.

2.2.2 *Indigo as a novel blue pigment*

Only a few different types of blue pigment are known, other than delphinidin. Regrettably, the biosynthesis of most of these pigments has not been studied in detail. The marine microalga *Haslea ostrearia* produces at its extremities a blue hydrosoluble pigment called marennine, responsible for the greening of oysters (Robert and Hallet, 1981). *Streptomyces coelicolor* produces the pigment γ-actinorhodin when subjected to various nutrient limitations (Bystrykh *et al.*, 1996). This pigment is related to the red/blue pH indicator pigment actinorhodin, but is blue over a wider pH range. Yet another blue pigment, 4,5,4',5'-tetrahydroxy-3'-diazadiphenoquinone-(2,2'), is produced by *Arthrobacter* species during degradation of 2,6-dihydroxypyridine (O'Laughlin *et al.*, 1999).

One may even think beyond pigments and instead consider chromoproteins. Chromoproteins typically require a prosthetic group, such as a metal ion or small non-peptide molecule, that is essential for the chromogenic properties of the protein. The only known exception to this rule is a nonfluorescent homologue of green fluorescent protein that was found to determine strong purple colouration in the sea anemone *Anemonia sulcata* (Lukyanov *et al.*, 2000). The colour of this protein is due entirely to internal interactions between amino acids within the protein – no other cofactors or prosthetic groups are required.

One class of chromoprotein that seems promising for use in molecular flower breeding is the blue biliproteins. These chromoproteins utilize biliverdin, a product of heme degradation by heme oxygenase isozymes, as a chromophore (Rhie and Beale, 1992). The green camouflage of many plant-feeding caterpillars results from a combination of such blue biliproteins with yellow carotenoid pigments (Saito, 1998). Also, blue biliproteins serve as part of light-harvesting complexes in photosynthetic cyanobacteria and red algae (de Lorimier *et al.*, 1993). Another option is to use blue copper proteins (sometimes called cupredoxins) such as plantacyanin, plastocyanin and azurin. These chromoproteins have characteristic blue or blue-green colours, due to a so-called 'blue Cu centre' in the protein that binds a copper ion as cofactor (Guss *et al.*, 1996). In cases where the biological function of the blue Cu centre is known, it is invariably electron transfer.

But one pigment stands out above all other blue pigments and chromoproteins as being by far the best known and most studied: the pigment named indigo. Indigo has, in fact, been the most important blue dye for mankind since prehistoric times. It could be extracted from various plant sources, particularly woad, *Isatis tinctoria*, in Europe, *Polygonum tinctorium* in

the Far East, and also tropical *Indigofera* species. In medieval Europe, a large industry grew around the production of indigo from woad. However, this branch of agriculture collapsed worldwide within a few years when, in the 1880s, the first commercial chemical process for synthesis of indigo was developed. Since then little scientific attention has been paid to indigo biosynthesis in plants, and the biosynthetic pathways are not yet fully known. But the advent of genetic engineering led to a surprising about-turn in the commercial biological production of indigo when indigo synthesis in *E. coli* was discovered (Ensley *et al.*, 1983). Microbial production of indigo has since become one of the best examples of commercial pathway engineering, and the biosynthetic pathways have been intensely studied (Chotani *et al.*, 2000). Using many intricate techniques of pathway engineering, the microbial production of indigo now rivals the chemically produced indigo (Weyler *et al.*, 1999).

2.2.2.1 Biosynthesis of indigo in bacteria

It has been known since the 1920s that certain microbes can produce the indigo pigment (O'Connor *et al.*, 1997). However, it was not until the early 1980s that scientists seeking a greener alternative method of indigo production looked to microorganisms. Since then several indigo-producing microorganisms have been found, the majority of which are aromatic hydrocarbon-degrading bacteria such as toluene-degrading *Pseudomonas mendocina* (Yen *et al.*, 1991), naphthalene-degrading *Pseudomonas putida* NDO and RKJ1 (Murdock *et al.*, 1993; Bhushan *et al.*, 2000), styrene-degrading *Ps. putida* S12 and CA-3 (O'Connor *et al.*, 1997), *p*-cumate-degrading *Ps. putida* F1 (Eaton and Chapman, 1995), *p*- and *m*-toluate-degrading *Ps. putida* mt-2 (Eaton and Chapman, 1995) and phenol-degrading *Bacillus stearothermophilus* (Oriel and Kim, 1998). These organisms produce indigo when grown in a medium supplemented with indole. Through their aromatic hydrocarbon-degrading ability they hydroxylate the indole to yield indoxyl. Spontaneous air oxidation and dimerization of indoxyl then leads to the formation of insoluble indigo (Figure 2.10).

2.2.2.2 Indigo genes

In principle, any microorganism that expresses an enzyme capable of introducing a hydroxyl moiety into the 3-position of indole to give indoxyl can produce indigo (Weyler *et al.*, 1999). These enzymes are typically aromatic dioxygenases, monoxygenases or hydroxylases. The

genes encoding four such enzymes have now been cloned and characterized: phenol hydroxylase from *Bacillus stearothermophilus* (Oriel and Kim, 1998), indole dioxygenase from *Rhodococcus* (Hart *et al.*, 1990), toluene-4-monooxygenase from *Pseudomonas mendocina* (Yen *et al.*, 1991) and naphthalene dioxygenase from *Pseudomonas putida* NCIB9816 (Kurkela *et al.*, 1988). Cloning of these 'indigo genes' into *E. coli*, thereby conferring on *E. coli* the ability to produce indigo, was the first step towards creation of a production strain for use in commercial microbial indigo production.

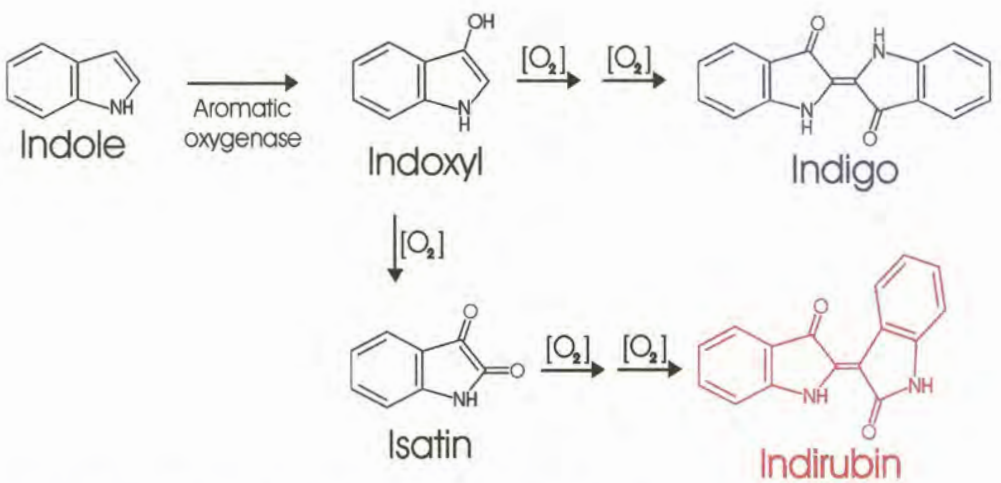

Figure 2.10: Biosynthesis of indigo from indole in bacteria. The conversion of indole to indoxyl is enzymatically catalysed by one of several aromatic oxygenases, while other steps are spontaneous. Small amounts of other pigments such as indirubin are also formed (based on Chotani *et al.*, 2000).

The *Ps. putida* naphthalene dioxygenase was the first to be characterized and is most often used in studies of commercial microbial indigo production. It is a soluble, multimeric NADPH2-dependent enzyme, composed of a reductase polypeptide, an iron-sulfur ferredoxin polypeptide and an oxygenase iron-sulfur protein. The *Ps. mendocina* toluene-4-monooxygenase is also composed of multiple different subunits, and has a relatively broad substrate specificity. In contrast, the *Rhodococcus* indole dioxygenase and *B. stearothermophilus* phenol hydroxylase are both simple, monomeric soluble proteins, approximately 43 kDa in size. Since *B. stearothermophilus* is a thermophilic organism, the phenol hydroxylase has a relatively high optimal temperature of 55°C, and is stable at elevated temperatures. Oriel and Kim (1998) have therefore suggested that the phenol hydroxylase would be a better candidate for commercial indigo production than the naphthalene dioxygenase. Unfortunately, not much more is known about the stability of these

enzymes under various conditions, nor have their relative rates of indigo production been measured.

A number of novel and unusual indigo genes have been described in the past few years. This followed the serendipitous discovery that some cytochrome P450 enzymes can also catalyze the hydroxylation of indole to form indigo. Cytochrome P450 (cyt-P450) enzymes are found throughout nature and function primarily in the oxidation of various chemicals (Gillam, 1999). Gillam *et al.* (1999; 2000) noticed that when certain human cyt-P450 enzymes are co-expressed with their natural redox partners in *E. coli*, the bacterial cultures spontaneously start to produce indigo. The most intense blue colour was seen with P450s 2A6 and 2E1. These are microsomal liver enzymes involved in the detoxification of diverse xenobiotic compounds and, accordingly, have a wide substrate specificity that includes hydroxylation of aromatic and heterocyclic compounds (Li *et al.*, 2000). The ability of such enzymes to hydroxylate indole is therefore understandable.

After this discovery, Nakamura *et al.* (2001) used random mutagenesis of the substrate recognition region of P450 2A6 to find mutants with improved ability to produce indigo. The catalytic efficiencies of these mutants have been determined, but since the same has not yet been done for the *Rhodococcus* indole dioxygenase or other bacterial aromatic oxygenases, the relative activity of the new cyt-P450 indigo genes cannot yet be directly compared with the other indigo genes. The nearest to such a comparison is the observation that the yield of indigo formed in *E. coli* cultures that express the *Rhodococcus* dioxygenase is about sevenfold higher than that obtained with P450 2A6 cultures (Hill *et al.*, 1989; Gillam *et al.*, 2000), but this is hardly a reliable measure of the relative catalytic efficiencies.

P450 2A6 and 2E1 are class II cytochromes, which means that they utilize NADPH as a cofactor and need an FAD- and FMN-containing reductase as an accessory enzyme (Gillam and Guengerich, 2001). Like all cyt-P450s, 2A6 and 2E1 bind heme as a prosthetic group. Another member of class II is the *Bacillus megaterium* P450 BM3 (Gillam, 1999). P450 BM3 is, however, unique in that the reductase and the cyt-P450 are fused together as two domains of a single large protein. The immediate proximity of the reductase increases the coupling efficiency and gives P450 BM3 the highest catalytic activity of any cyt-P450. Moreover, P450 BM3 is water-soluble, unlike most other members of the cyt-P450 family which are membrane-bound, multi-protein complexes.

P450 BM3 displays a rather narrow substrate specificity limited to long-chain fatty acids and structurally related compounds. Through directed evolution, Li *et al.* (2000) succeeded in

changing this specificity until a triple mutant was created that could efficiently hydroxylate indole and a variety of alkanes, cycloalkanes, arenes and heteroarenes (Appel *et al.*, 2001). The triple mutant that has so far been created is not yet optimal and has a little lower catalytic efficiency than P450 2A6 and its mutants (Nakamura *et al.*, 2001). Nevertheless, since P450 BM3 is far more active and stable than most other cyt-P450s, it seems ideally suited for the creation of a new, highly efficient indigo gene (Li *et al.*, 2000).

2.2.2.3 Tryptophan synthase as a source of free indole

It might at first seem rather strange that the media of these indigo-producing bacteria had to be supplemented with indole, since indole is produced by the α subunit of tryptophan synthase during the penultimate step in tryptophan synthesis in both bacteria and plants. However, it has actually been found that indole does not generally occur free inside cells. The endogenously produced indole is therefore not available for production of indigo. The reason why indole is not allowed to occur free may be that it is a small, hydrophobic molecule that could easily diffuse across membranes and away from the tryptophan-synthesizing enzymes (Streger, 1995).

The mechanism by which tryptophan synthase prevents indole from occurring free can be seen in the unique structure of the enzyme (Figure 2.11). Biochemical and X-ray diffraction studies of plant and bacterial tryptophan synthases have shown that it is a soluble, tetrameric $\alpha_2\beta_2$ complex, consisting of two α subunits and two β subunits (Radwanski *et al.*, 1995; Hyde *et al.*, 1988). Each subunit is encoded by a separate gene and catalyzes different half-reactions. The α subunit (TSA) catalyzes the cleavage of indole 3-glycerol phosphate to indole and glyceraldehyde 3-phosphate, while the β subunit (TSB) catalyzes the condensation of indole with serine in a pyridoxal phosphate (PLP)-mediated reaction. The solution of the three-dimensional structure of the enzyme from *Salmonella typhimurium* provides physical evidence for a 25 Å hydrophobic tunnel connecting the α and β active sites (Hyde *et al.*, 1988). Indole would be transferred from the active site of the α subunit to the active site of the β subunit through the connecting tunnel without diffusing into the environment. Tryptophan synthase is the first known example of such 'substrate channeling', defined as the process in which a metabolic intermediate is directly transferred from one active site to another without free diffusion.

Association of the two subunits is required for full activity of the complex, as each subunit on its own has only trace catalytic activities – about 100-fold less than when associated. But the subunits do not merely activate each other. Allosteric interactions are proposed to switch the α and β subunits between catalytically active and inactive conformations, thereby coordinating the activities at the two sites. To prevent accumulation of indole at the α site, the reaction of serine at the β site modulates the formation of indole at the α site such that indole is not produced until serine has reacted with PLP (Anderson *et al.*, 1995).

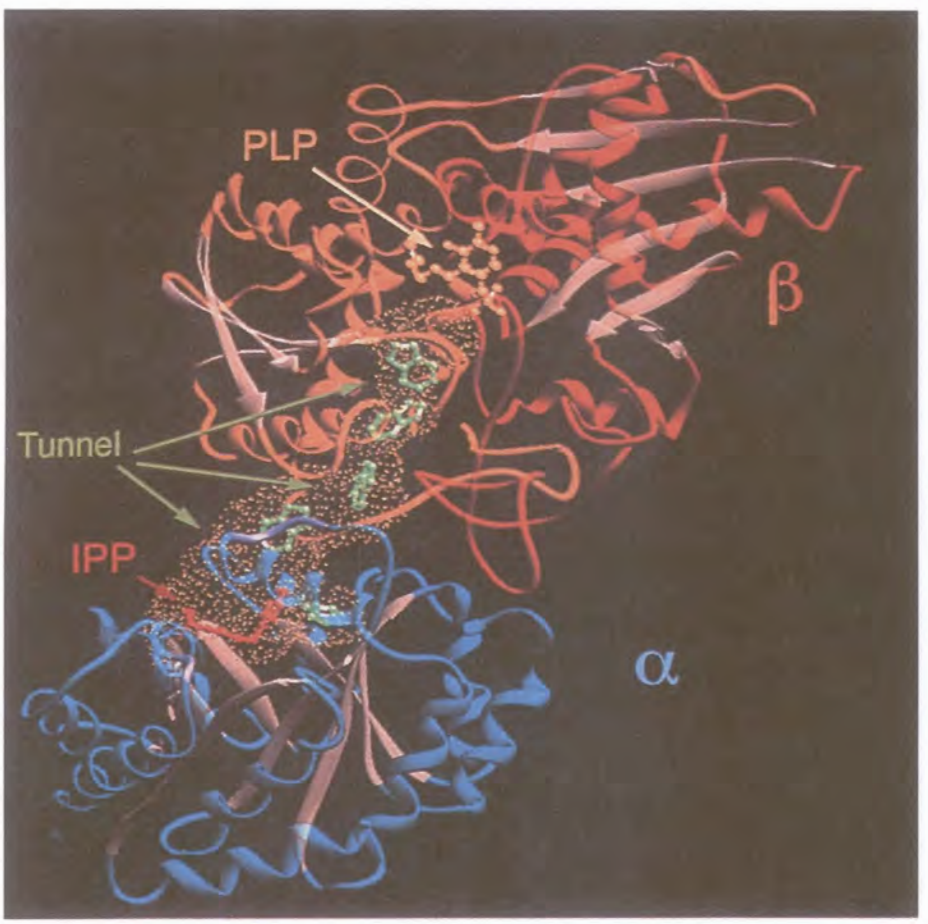

Figure 2.11: Three-dimensional structure of the *Salmonella typhimurium* tryptophan synthase. Only one αβ unit of the $\alpha_2\beta_2$ tetramer is shown. The α subunit is shown in blue and the β subunit in orange and red. The inhibitor indolpropanol phosphate is located at the α active site, while the PLP coenzyme is at the β active site. The ~25 Å-long channel connecting the α and β active sites is outlined by yellow dots. Several indole molecules have been modeled into the channel as green ball-and-stick models, thereby demonstrating the movement of the indole product of the α subunit through the channel to the β active site (Voet and Voet, 1995).

In order for microbial indigo production to be commercially viable, the production strain has to produce free indole endogenously from glucose or other cheap, renewable carbon sources. Since each subunit on its own has only trace catalytic activities, indole can not be produced by simply overexpressing the α subunit. But in a fortuitous discovery, Yanofsky and Crawford (1959) observed that *E. coli* carrying a point mutation near the C-terminus of the α subunit accumulated intracellular free indole. This mutation replaced a lysine residue that lined the indole channel with an asparagine. Presumably this disrupted the indole channel, preventing the movement of indole from the α subunit to the β subunit. Because this mutation had a deleterious effect on the cell's ability to synthesize tryptophan and was only a point mutation, it was very unstable and subject to reversion to the wild type genotype.

However, this observation gave impetus to later studies that attempted to find stable mutations that would allow indole to escape from the channel into the environment (Murdock, 1996). Through trial and error several such mutations were found, some yielding intracellular indole far in excess of that observed by Yanofsky and Crawford (1959). The mutations involved the disruption of the indole channel in either the α or β subunit, or disruption of the β subunit active site so that the β subunit is still capable of assembling into the holoenzyme but is incapable of catalyzing the conversion of serine and indole to tryptophan. To avoid spontaneous reversion to the wild type, each of the engineered mutations preferably involved more than one nucleotide base pair change. When these mutant tryptophan synthase α and β subunits were expressed in *E. coli* together with naphthalene dioxygenase, indigo was indeed produced from medium without any indole.

2.2.2.4 Isatin hydroxylase and indigo purity

Pathway engineering was also important to improve the quality of the indigo dye. At neutral pH, the indigo precursor indoxyl can, to a small extent, spontaneously oxidize to form isatin as a by-product. Dimerization with another molecule of indoxyl leads to the formation of indirubin, an isomer of indigo with a deep burgundy colour. In order to make textile grade indigo, the level of indirubin formation had to be reduced. Also, isatin was found to inhibit naphthalene dioxygenase activity and, consequently, reduces overall indigo production.

A search was therefore initiated for an enzyme capable of eliminating or removing isatin build-up (Weyler *et al.*, 1999). Numerous soil samples were screened using a nutritional selection scheme, resulting in the identification of an organism exhibiting the ability to

degrade isatin. Taxonomic studies suggested the organism is a *Pseudomonas putida*, and the enzyme was identified as an isatin hydrolase, which hydrolyzes isatin to isatic acid. After cloning and incorporating the new enzyme in an *E. coli* production strain, the indigo product performed equal to indigo produced chemically. It is interesting to note that the styrene-degrading *Pseudomonas putida* S12 and CA-3 are also known to stoichiometrically produce high-purity indigo (O'Connor *et al.*, 1997) from indole-containing medium. This suggests that these strains also have the ability to degrade isatin.

2.2.2.5 Tryptophanase as a source of free indole

Although past attempts to create production strains capable of accumulating indole have concentrated on mutant tryptophan synthases, there actually already exists an enzyme that catalyzes the formation of free indole, namely tryptophanase. Tryptophanase is a catabolic enzyme that catalyzes a pyrodixal 5'-phosphate-dependent conversion of tryptophan to indole, pyruvate and ammonia. When tryptophanase-expressing *E. coli* strains were transformed with the *Rhodococcus* indole dioxygenase, indigo was produced without the need for exogenous indole (Hart *et al.*, 1992). But although Weyler *et al.* (1999) did mention the possibility of using a tryptophanase to produce indole during commercial biosynthesis of indigo, this has not yet been attempted.

Tryptophanases have been found in various Gram-negative bacterial species indigenous to the intestinal tracts of animals, but not in any other species (DeMoss and Moser, 1969). The genes encoding tryptophanase from several species have been cloned and characterized, such as *E. coli* (Deelay and Yanofsky, 1981), *Symbiobacterium thermophilum* (Hirahara *et al.*, 1992), *Proteus vulgaris* (Kamath and Yanofsky, 1992) and *Enterobacter aerogenes* (Kawasaki *et al.*, 1993). Tryptophanases isolated from these sources are soluble, cytoplasmic proteins with a similar multimeric structure of four identical 52 kDa subunits (Isopov *et al.*, 1998). Each monomer binds a PLP molecule as a coenzyme, and requires monovalent cations such as K^+, NH_4^+ or Tl^+ for its activity. The different tryptophanases do, however, show differences in activity, amino acid composition and antigenic properties (Kamath and Yanofsky, 1992). The *E. aerogenes* tryptophanase is a bit more active than the *E. coli* tryptophanase over a range of temperatures and pH (Kawasaki *et al.*, 1995). The *S. thermophilum* tryptophanase is thermostable and has a high optimal temperature, which is understandable as *S. thermophilum* is a thermophilic organism (Suzuki *et al.*, 1991).

Tryptophanase is actually a multifunctional enzyme and can, under higher substrate concentrations, also catalyze several other reactions (Newton *et al.*, 1965; Morino and Snell, 1967) such as:

L-serine → pyruvate + NH_3

L-cysteine → pyruvate + NH_3 + H_2S

S-Methyl-L-cysteine → pyruvate + NH_3 + CH_3SH

L-serine + indole → L-tryptophan + H_2O

L-cysteine + indole → L-tryptophan + H_2S

S-Methyl-L-cysteine + indole → L-tryptophan + CH_3SH

pyruvate + NH_3 + indole → L-tryptophan

As can be seen from this list, tryptophanase can even catalyze the synthesis of tryptophan. It has therefore been suggested that tryptophan might be economically produced from synthetic materials such as pyruvic acid, indole and ammonia by tryptophanase-expressing microorganisms (Suzuki *et al.*, 1991; Kawasaki *et al.*, 1995). In fact, most of the studies that have been done on possible commercial uses of tryptophanase have focused on this synthetic reaction. Fortunately for the use of tryptophanase in indigo production, the synthetic reaction only occurs under high substrate conditions. The very low substrate affinities that tryptophanase has for pyruvate and ammonia (Watanabe and Snell, 1972) suffice to explain why, under normal cellular conditions, only the degradation of tryptophan to indole, ammonia and pyruvate happens to a measurable extent (Isopov *et al.*, 1998).

2.2.2.6 Biosynthesis of indigo in plants

Indigo-producing plants actually have quite normal-looking, green leaves. Only when the leaves are damaged, do the wounded tissues produce indigo and become blue. The indigo is produced from two indoxyl derivatives (Figure 2.3): isatan B (indoxyl-5-ketogluconate) and indican (indoxyl 3-β-D-glucoside) (Kokubun *et al.*, 1998). By action of β-glucosidases on these molecules, the indoxyl moiety is liberated and subsequent spontaneous air oxidation of indoxyl yields indigo (Xia and Zenk 1992). Indican, and probably also isatan B, is sequestered in the vacuole (Minami *et al.*, 2000). The β-glucosidase, on the other hand, is found exclusively in the plastids (Minami *et al.*, 1997). When the cell is damaged, the physical barriers between indican, isatan B and β-glucosidase are broken down, and indigo is produced. At the same time a small amount of other pigments such as indirubin, indigo brown, indigo gluten and indoxyl brown are also formed as by-products (Kim *et al.*, 1996).

The most significant of these by-products is, as in bacteria, indirubin. Interestingly, it has been claimed that the aesthetically pleasing hue of the natural indigo, which contains small amounts of other pigments, surpasses that of the synthetic dye (Kim *et al.*, 1996).

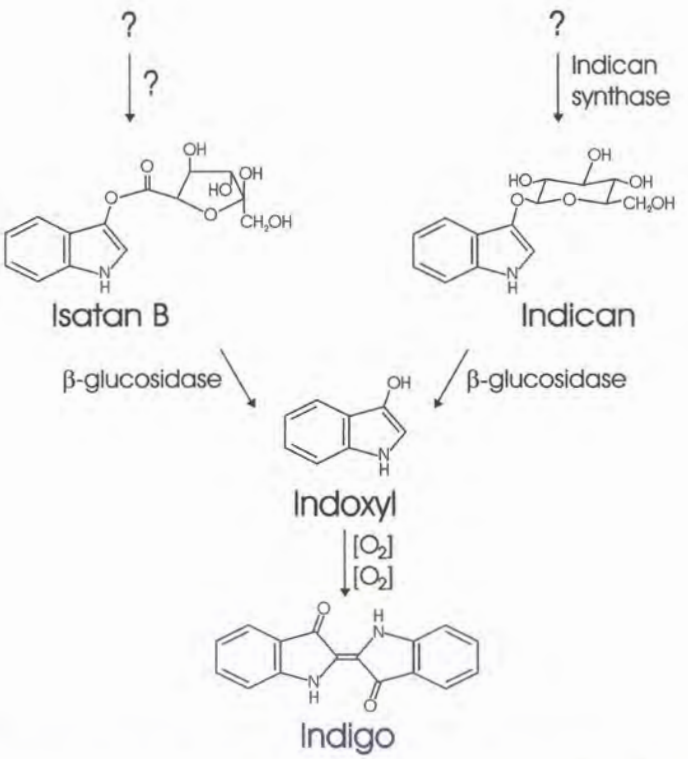

Figure 2.12: Biosynthesis of indigo in plants. Indigo is formed through spontaneous oxidation of indoxyl. Indoxyl is formed through the action of β-glucosidases on either indican or isatan B. The precursors of indican and isatan B are not yet known. As in bacterial indigo synthesis, small amounts of other pigments such as indirubin are also formed, but these pathways are not shown. (See discussion for details.)

The production of indigo and indirubin could be induced in non-wounded tissue cultures of *Polygonum tinctorium* by feeding with indole (Kim *et al.*, 1996; Shim *et al.*, 1998). This suggests that the production of indigo is not necessarily toxic to undamaged plant cells. However, it is not yet known in which subcellular compartments the indigo was produced or sequestered.

Little else is known about the biosynthesis of indigo in plants. An enzyme has recently been isolated that might be responsible for the formation of indican, but the origins and precursors of the enzyme's substrates are not known (Minami *et al.*, 2000). Only after some contradictory results in the literature concerning the role of tryptophan in the biosynthesis of

indigo, has it finally been shown that indole, and not tryptophan, is the biosynthetic precursor of both isatan B and indican (Xia and Zenk, 1992). This finding, however, actually deepens the mystery, since the origin of the free indole is not yet known. In plants, as in bacteria, indole stays trapped inside the tryptophan synthase holoenzyme and does not generally occur free inside cells. Nor has any tryptophanase activity ever been found in plants.

2.2.2.7 Indole synthase as yet another source of free indole

The answer to the mysterious origin of free indole might have been found in studies of the biosynthesis of DIMBOA (2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one), a pesticidal hydroxamic acid found in maize and other grasses. In higher plants, the tryptophan pathway provides precursors for the plant hormone auxin and many other indolic secondary products, such as hydroxamic acids. Indole is the last intermediate common to both the tryptophan and hydroxamic acid pathways, and is converted to DIMBOA in a series of cytochrome P450-mediated oxygenation steps (Frey *et al.*, 1995). In this pathway, indole is produced by an indole synthase (BX1), which catalyzes the conversion of indole-3-glycerol phosphate to indole (Frey *et al.*, 1997; Melanson *et al.*, 1997). The enzyme is homologous to TSA and can complement a tryptophan synthase α mutant in *E. coli* (Kramer and Koziel, 1995). But in contrast to TSA, whose activity depends on association with TSB, BX1 activity is independent of TSB. BX1 does not need to form complexes with TSB and simply releases free indole into the environment. In fact, the monomeric BX1 enzyme is about 29 times more efficient in the formation of indole than the active form of TSA in the $\alpha_2\beta_2$ complex (Frey *et al.*, 2000).

Indole is a volatile compound that contributes to the fragrance of many common garden flowers (Kite and Leon, 1995; Dobson *et al.*, 1997; Spiteller and Steglich, 2001; Picone *et al.*, 2002). It is also released by a variety of plant species in response to herbivore damage, together with several other volatile compounds. Herbivore predators and parasitic wasps exploit these chemical cues to locate their prey or hosts. Frey *et al.* (2000) identified a BX1-homologous enzyme in maize, named indole-3-glycerol phosphate lyase (IGL), that catalyzes the formation of free indole and is selectively activated by the oral secretions of herbivores such as beet armyworm. As is the case with BX1, IGL activity is independent of TSB, and IGL is about 3 times more efficient in the formation of indole than TSA (Frey *et al.*, 2000). For both IGL and BX1, changes relative to the TSA consensus sequence are specifically located in two domains (Frey *et al.*, 2000). These modified domains are part of the

interaction domains of TSA and TSB, based on the crystal structure of the *S. typhimurium* tryptophan synthase complex.

Both BX1 and IGL contain chloroplast-targeting peptides at their NH-termini (Kramer and Koziel, 1995; Frey *et al.*, 2000). This is similar to other enzymes involved in the aromatic amino acid biosynthetic pathways, all of which are localized to the chloroplast stroma, though encoded by nuclear genes (Zhao and Last, 1995). Protein sequences deduced from the cDNA sequences of all cloned tryptophan biosynthetic enzymes from *Arabidopsis* contain chloroplast-targeting peptides at their NH-termini. Therefore an efficient transport system exists to move tryptophan across the chloroplast membrane, and approximately 90% of tryptophan is found outside the chloroplasts. Interestingly, the cytochrome P450 enzymes responsible for conversion of indole into DIMBOA do not contain such chloroplast-targeting peptides (Frey *et al.*, 1995). Cytochrome P450 enzymes are usually bound to the endoplasmic reticulum of eukaryotic cells. This implies that indole produced by BX1 has to move out of the chloroplast into the cytosol before being converted into DIMBOA. Since indole is a small, hydrophobic molecule it readily traverses membranes and, presumably, simply diffuses out of the chloroplast (Streger, 1995).

The synthesis of indole-3-acetic acid and several plant secondary metabolites, such as indole glucosinolates, terpenoid indole alkaloids, indolic phytoalexins and some floral volatiles, depends on the tryptophan pathway (Zook, 1998; Frey *et al.*, 2000; Ouyang *et al.*, 2000). After the discovery of BX1 and IGL it has been suggested that homologous enzymes might also catalyze the first specific steps in the synthesis of these secondary metabolites (Zook, 1998; Frey *et al.*, 2000). Perhaps the same is true of the synthesis of indigo.

2.2.2.8 Genetic engineering of indigo biosynthesis in flowers

Based on the above discussion, the most promising pathways for indigo biosynthesis in flowers can now be selected. One's first thought may be to transfer the indigo biosynthetic pathway of indigo-producing plants into the ornamental species, perhaps with a few tweaks to direct the enzymes to appropriate subcellular compartments. The use of plant genes would indeed have avoided many of the potential problems during transcription and translation. However, the pathway is not yet known in nearly enough detail. The new indigo pathway would therefore have to be based on the bacterial indigo biosynthetic pathway, which consists of two steps: first, the production of indole, and then the hydroxylation of indole to indoxyl,

which spontaneously converts to indigo. It should not be necessary to include isatin-degrading steps in the pathway, since the aim is not to produce textile-grade pigments, but rather blue flowers with an 'aesthetically pleasing hue' (Kim *et al.*, 1996).

The production of indole can potentially be performed by any one of three enzymes: tryptophanase, indole synthase, or a mutant tryptophan synthase. The use of a mutant tryptophan synthase is, however, not a very attractive option. Although bacterial tryptophan synthases and plant tryptophan synthases have a similar three-dimensional structure, their subunits are not necessarily able to interact productively and activate each other (Radwanski *et al.*, 1995). This potential problem may be circumvented by expressing both subunits of the bacterial tryptophan synthase, or by trying to find indole-accumulating mutants of the plant tryptophan synthase. But a much simpler option would be to just use the indole synthases, which perform exactly the same function as the mutant tryptophan synthases and obviously function quite well in at least some plants. Also, both of the indole synthases – and especially BX1 – are much more active than the tryptophan synthases, so they would not have to be expressed at high levels that may cause dosage-dependant post-transcriptional silencing or co-suppression of homologous TSA genes.

The use of a tryptophanase to produce indole has only one potential complication: the multimeric, multi-domain structure of tryptophanase, which might not fold correctly in the cytosol. The folding of the *E. coli* tryptophanase is known to be facilitated by the bacterial chaperonins (Mizobata *et al.*, 1992). Otherwise the use of a tryptophanase seems quite promising. The cofactors needed by tryptophanase are as ubiquitous in plant cells and plastids as in bacteria (Bérczi and Møller, 2000), and plants contain no known tryptophanase homologues that might cause silencing. Among the different tryptophanases, the *E. aerogenes* tryptophanase would be the first choice. Unlike the *S. thermophilum* tryptophanase, it does not originate from a thermophilic organism and it is more active than the *E. coli* tryptophanase, while the relative activity of the *P. vulgaris* tryptophanase is not known. Also, the *E. aerogenes* tryptophanase gene is GC-rich and contains no obvious sequences that might cause undesired post-transcriptional or post-translational modifications (Kawasaki *et al.*, 1993).

The second step in the pathway – the hydroxylation of indole – can also be performed by many different indigo genes, but some are so clearly impractical for use in flowers that they can be rejected out of hand. The *B. stearothermophilus* phenol hydroxylase has such a high optimal temperature that little enzyme activity remains at room temperature (Oriel and Kim, 1998). The multimeric *Ps. putida* naphthalene dioxygenase and *Ps. mendocina* toluene-4-

monooxygenase are vastly too complex to even consider (Kurkela *et al.*, 1988; Yen *et al.*, 1991). This leaves the *Rhodococcus* indole dioxygenase and the cytochrome P450 indigo genes.

None of the cyt-P450 indigo genes can be rejected out of hand, but neither are any of them without problems. Of the various cyt-P450 indigo genes, the triple mutant P450 BM3 has several advantages. It is soluble, more stable than other cyt-P450s, and its accessory reductase need not be expressed separately (Li *et al.*, 2000). It is not known whether the endogenous plant reductases would be able to interact with exogeneous cyt-P450s, so cyt-P450 indigo genes other than P450 BM3 might have to be expressed together with their respective accessory reductases. P450 BM3 has a wide substrate range that may cause unwanted side effects in plant cells, but this is even more true for the other cyt-P450 indigo genes. Moreover, since the wild-type P450 BM3 is inherently more active than other cyt-P450s, future studies may be able to create mutants far more active than the current triple mutant, and with narrower substrate specificities.

There are, unfortunately, some drawbacks to using P450 BM3. Its gene is AT-rich, while those of P450s 2A6 and 2E1 are GC-rich (Ruettinger *et al.*, 1989; Hayashi *et al.*, 1996). Also, while P450 2A6 and 2E1 are small, eukaryotic, microsomal enzymes, P450 BM3 is a very large, multidomain, bacterial gene that may not fold efficiently in the cytosol.

The use of cyt-P450 indigo genes entails two potential problems that have little to do with their biological characteristics. Firstly, the P450 BM3 and 2A6 mutants were deliberately created instead of merely 'found', and are therefore protected by much stricter intellectual property rights. This limits the use of these genes in commercial ventures (or rather, the profits that can be made from these genes). The second complication concerns the human origins of P450s 2A6 and 2E1. No genetically modified plant with transgenes from human origin has ever been commercially released, so it not known whether such plants would gain public acceptance. Ideally, the story of how the fabled blue rose could only be created with the help of a gene isolated from humans themselves would pique the public's interest. Sadly, however, there is also a chance that it will instead lead to rejection of the flowers as 'Frankenflowers'.

In comparison to the other indigo genes, the *Rhodococcus* indole dioxygenase seems wonderfully problem-free. It is a simple, monomeric soluble enzyme with a typical optimal temperature of 37°C. The *ido* gene is GC-rich with a preference for G or C in the third codon position, shows no significant homology to any known plant gene, and contains no obvious

problematic sequences (Hart *et al.*, 1990). Although it is not yet known whether the indole dioxygenase needs a coenzyme, it probably utilizes either a metal or a flavin coenzyme like most other oxygenases (Mathews and van Holde, 1990). These cofactors are ubiquitous and easily available in plant cells, like the cofactors of tryptophanase. The indole dioxygenase is, therefore, the first choice for an enzyme to catalyze the hydroxylation of indole.

After selecting the enzymes that will be used, it still remains to decide where the enzymes should be targeted to. The indole dioxygenase should preferably remain in the cytosol, since targeting of the enzyme to the plastids may cause toxic build-up of insoluble indigo inside the plastids. From the cytosol the indigo can more easily be deposited into the cell wall or vacuole. Wherever the indole dioxygenase is targeted to, its activity should be higher than that of the indole synthase or tryptophanase to prevent the toxic effects of excess indole.

If an indole synthase is used, it can only be targeted to the plastid, since its substrates are not found anywhere else. Tryptophanase, on the other hand, can be targeted to either the cytosol or the plastid – but not to the vacuole, since tryptophanases have very little if any activity at such low pH (Kawasaki *et al.*, 1995). The relative concentrations of tryptophan in the plastids, cytosol and vacuole are not certain, so do not help in deciding between the two. The folding of tryptophanase might be more efficient in the plastid, but then the indole has to diffuse out of the plastid before the indole dioxygenase can reach it. This should not be a problem, as the same probably happens during DIMBOA synthesis. Since neither the cytosol nor the plastids can be eliminated out of hand, both should be tried to see which one gives the best results.

Lastly one should attempt to predict the eventual fate of the pigment. Since indigo is a foreign, insoluble compound, it will likely be excreted from the cytosol into either the cell wall or vacuole (Harvey *et al.*, 2002). Before excretion, the pigment may be conjugated to endogenous hydrophilic molecules such as glutathione, sugars or malonyl derivatives. Even after entering the vacuole, the conjugates may be further metabolized. There is, unfortunately, only scarce knowledge about the factors that determine how such foreign, insoluble compounds are metabolized and where they will be deposited. These factors also differ from species to species. The eventual fate of indigo pigments in plant cells can therefore not be predicted with certainty.

At least the cellular metabolism of indigo should not alter the colour of the pigment, since normal, blue indigo was produced in live cells from *Polygonum tinctorium* after feeding the cells with indole (Kim *et al.*, 1996; Shim *et al.*, 1998) or electroporating β-glucosidase into

the cells (Minami *et al.*, 2000). *P. tinctorium* and the other indigo-producing plants are not specialized for the production of indigo in undamaged cells, so it is unlikely that the metabolism of indigo in ornamental plants will be very different. Furthermore, indigo is pH stable (Hart *et al.*, 1992) – in contrast to delphinidin-3-glucoside – so the relatively acidic pH of rose petal vacuoles should no longer present any obstacles to the creation of blue roses.

2.2.3 Melanin as a novel black pigment

Flowers can be given a black appearance by extremely high expression of one or more of a variety of different pigments. However, it would probably be just that much more easy if the pigment is black to begin with. The choice of such a black pigment also seems rather easy, since there is only one well-known black pigment, namely melanin. (The only other known black pigment being an obscure heme pigment of *Porphyromonas gingivalis* and some other anaerobes (Smalley *et al.*, 2000).) Unfortunately, the choice of pigment is not as simple as it might appear at first, as the term 'melanin' has historically been used rather indiscriminately to mean any dark pigment, and now refers to a large and varied class of pigments (Riley, 1997).

Even today there is still contention about the precise definition of a melanin. This might reflect the resistance of many black or dark natural pigments to degradation, and the resulting intractability toward analysis (Butler and Day, 1998). Accordingly, the isolation and purification of melanins require rather drastic procedures. The resulting pigment is insoluble in water and may contain protein, carbohydrate and lipid moieties surrounding an indolic nucleus. The nucleus consists of an irregular polymer of indoles but also contains variable amounts of other pre-indolic products of the synthetic pathway such as quinones, hydroquinones, and semiquinones (Figure 2.13; Bell and Wheeler, 1986; Riley, 1997). The nucleus is apparently the chromophore, but the extent to which it is bonded to the other moieties in the natural melanin is uncertain. Even foreign phenols, such as pesticides or plant phenols can be incorporated into melanins by secreted melanin-forming enzymes.

Besides the obvious light-absorption property, melanins also have powerful cation chelating properties, can take part in redox reactions, and are highly resistant against enzymatic lysis (although biological degradation of melanins by two fungal ligninases, lignin peroxidase and manganese peroxidase, has recently been reported) (Riley, 1997; Butler and Day, 1998). These properties allow melanin to serve a wide range of functions that, surprisingly, all have to do with protection. These include protection against radiation, protection of certain fungi

against enzymatic lysis, binding of toxic metals to prevent entry into the cell, resistance against certain fungicides by binding and inactivating them, protection against extremes of temperature, protection against oxygen radicals, and protection against predation through camouflage. Even the protective ink released by squids and octopuses consists of a fine suspension of melanin granules.

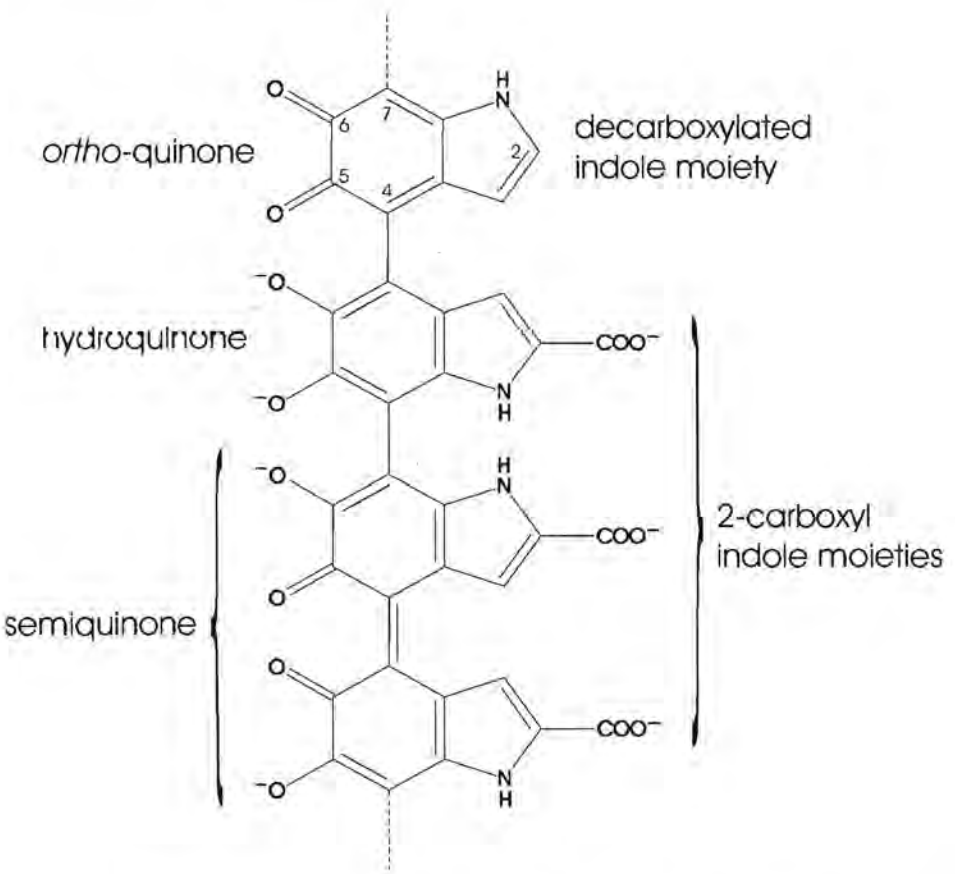

Figure 2.13: Structure of the chromophore nucleus of indolic melanin as proposed by Riley (1997). Carboxylic acid groups can be attached to C2. The C5 and C6 functionalities can assume structures such as carbonyls in the *ortho*-quinone form, deprotonated hydroxyls in the catecholic form, or an equilibrium form of linked semiquinones.

This description is wide enough to even include melanin-like pigments that are not black, such as the red-brown pyomelanin formed by *Vibrio cholerae* (Kotob *et al.*, 1995), the red phaeomelanin formed by animals (Ito, 1993), and the blue-green melanin-like pigment produced by *Aspergillus fumigatus* (Tsai *et al.*, 1999). In spite of all this complexity, most intracellular melanins can be grouped into a small number of classes, named according to the major indolic precursor of the chromophore nucleus. The chromophore nucleus of an

intracellular melanin pigment almost always appears to come primarily from a single biochemical pathway and is formed by the condensation of one or a few similar indolic precursors. Melanins that are more complex, being derived from diverse indoles, amino acids, lipids and carbohydrates, are usually just referred to as heterogeneous melanins, and include most extracellular melanins. There are several enzymes, such as laccase, tyrosinase, and perhaps peroxidases and catalases, that will form heterogeneous melanins from indolic compounds in the environment (Bell and Wheeler, 1986).

2.2.3.1 Catechol, GHB and DHN melanins

Catechol melanin and glutaminyl-4-hydroxybenzene (GHB) melanin are two minor, little-known classes found in certain microorganisms. Catechol melanin is produced by *Azotobacter chroococcum* and certain Basidiomycetous mushrooms, and seems to be derived from tyrosinase oxidation of catechol (Bell and Wheeler, 1986; Shivprasad and Page, 1989). The melanins in some other Basidiomycetes are derived from GHB via the shikimate pathway. GHB is apparently oxidized by a tyrosinase to yield glutaminyl-3,4-hydroxybenzene as the immediate precursor of the melanin polymer (Bell and Wheeler, 1986).

The most widespread fungal melanin is 1,8-dihydroxynaphthalene (DHN) melanin. It is also sometimes found in plants and is then usually referred to as 'allomelanin'. The pathway is encoded by at least five genes, only four of which have been cloned and characterized (Kimura and Tsuge, 1993; Edens *et al.*, 1999; Tsai *et al.*, 1999). The starting molecule of the DHN pathway, 1,3,6,8-tetrahydroxynaphthalene, is formed by the joining and cyclization of acetate molecules by polyketide synthase. Through a series of reduction and hydroxylation reactions DHN is formed. In the final step the DHN molecules are polymerized to form the melanin chromophore. There are a number of candidate enzymes for this step including phenoloxidases such as tyrosinase, peroxidases, laccases and maybe even catalases (Butler and Day, 1998; Edens *et al.*, 1999).

2.2.3.2 DOPA melanin

Dihydroxyphenylalanine (DOPA) melanin is by far the best-known melanin class. It is the major pigment found in animals – referred to as 'eumelanin' by animal melanin researchers –

and is also produced by many plants, fungi and bacteria. In comparison to DHN melanin, the biosynthetic pathway of DOPA melanin is relatively simple, as shown in figure 2.14 (Ito, 1993). The first two steps in eumelanogenesis are catalyzed by the enzyme tyrosinase: the hydroxylation of tyrosine to L-dihydroxyphenylalanine (L-DOPA) and the oxidation of L-DOPA to dopaquinone. The subsequent steps occur spontaneously and involve the cyclization of dopaquinone to dopachrome, the oxidation of dopachrome to 5,6-dihydroxyindole (DHI), and the polymerization of DHI to form the black melanin pigment.

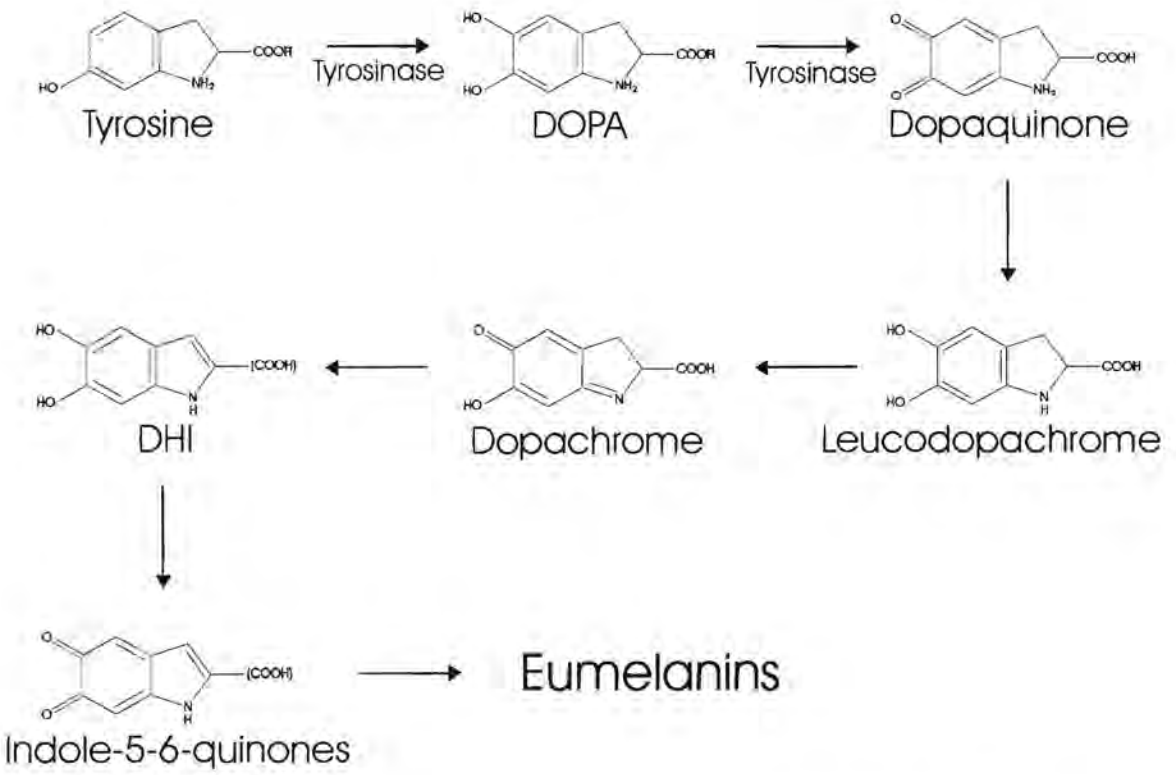

Figure 2.14: Raper-Mason scheme for eumelanogenesis (based on Ito, 1993). Tyrosinase catalyzes the first two reactions in eumelanogenesis, and the other steps are spontaneous.

2.2.3.3 Tyrosinases

Tyrosinase is a copper-containing monooxygenase widely distributed in nature and is mainly involved in the formation of melanin pigments. It is a bifunctional enzyme that catalyzes the orthohydroxylation of monophenols (cresolase activity) and also the oxidation of diphenols to quinones (catecholase activity). More than 20 tyrosinase sequences are known, ranging from

humans to bacteria (Ikeda *et al.*, 1996; van Gelder *et al.*, 1997; Goldman *et al.*, 1998). Within major taxa such as prokaryotes, fungi, plants and mammals the sequence homology of tyrosinases is high, and conserved domains can be identified for each taxon. Between the taxa, however, homology is quite low and the only conserved regions seem to be the two Cu-binding domains (van Gelder *et al.*, 1997).

2.2.3.3.1 Animal tyrosinase

The dark melanins produced by animals are usually referred to as eumelanins, but not all of the eumelanin is necessarily DOPA melanin. In some species up to half of the dopachrome is tautomerized and diverted into the formation of 5,6-dihydroxyindole-2-carboxylic acid (DHICA) melanin (Hearing and King, 1993). This alternative pathway is, however, not very relevant as it is more complex and DHICA melanin is not as dark as DOPA melanin.

In vertebrates, melanin is only synthesized in specialized cells called 'melanocytes' in warm-blooded vertebrates and 'melanophores' in cold-blooded vertebrates. These cells contain hundreds of melanosomes, which are melanin-synthesizing organelles derived from the Golgi apparatus and endoplasmic reticulum (Bell and Wheeler, 1986). Tyrosinases are found exclusively in the melanosomes, with a single membrane-spanning helix located in the C-terminal part of the protein and the bulk of the protein inside the melanosome. The transmembrane region is not necessary for enzyme activity and can be cleaved off by trypsin digestion to release the soluble part of the protein as an active monomeric enzyme (van Gelder *et al.*, 1997). In contrast, it has been established that N-glycosylation is essential for correct folding and subsequent activity of the enzyme (Petrescu *et al.*, 2000). The activity of tyrosinase fills the melanosome with melanin until it becomes a dense granule. Some melanocytes retain their melanin granules while others transfer their melanin granules to receptor cells such as keratinocytes through cytocrine transfer. The remaining granules generally remain dispersed in the cytoplasm and retain the melanosome membrane (Bell and Wheeler, 1986).

2.2.3.3.2 Plant tyrosinase

In most fruits and vegetables, tyrosinase is responsible for enzymatic browning following bruising, cutting or pathogen-induced cell damage. This might indicate a defensive role for

plant tyrosinases, perhaps involving oxidative detoxification of pathogen-produced phytotoxins. Plant tyrosinases are nuclear-encoded but are targeted to the plastids where they are stored in a latent, membrane-bound form (van Gelder *et al.*, 1997). It has been suggested that plant tyrosinases are glycosylated, but whether this is essential for correct folding or activity is not known (Flurkey, 1985; Hunt *et al.*, 1993). In damaged or senescent cells, the physical barriers between the enzyme in the plastid and the substrates in the cytoplasm and vacuole are broken down, allowing the enzyme to come into contact with its substrates. The enzyme is then activated by proteolytic removal of the C-terminal peptide, leaving the soluble N-terminal domains as an active monomeric enzyme (Dry and Robinson, 1994). The function of the C-terminal peptide remains unresolved, but several possibilities have been proposed (Dry and Robinson, 1994; van Gelder *et al.*, 1997): targeting of the protein to the chloroplast thylakoid, binding of the enzyme to the membrane by forming a transmembrane α-helix, interacting with the active site to render the enzyme inactive until cleavage, assisting in correct assembly of the protein within the chloroplast, or assisting in correct insertion of copper atoms.

2.2.3.3.3 Fungal tyrosinase

Fungal melanins are found in the cell wall, where they often form large electron-dense melanin granules scattered randomly in the matrix (Butler and Day, 1998). Fungal tyrosinases are soluble cytoplasmic enzymes (van Gelder *et al.*, 1997), but it has been suggested that the melanin precursors may actually be formed in cytoplasmic vacuoles before being secreted to the cell wall (Butler and Day, 1998). As in plants, fungal tyrosinases are produced as latent precursor proteins that are activated by proteolytic cleavage of the C-terminal peptide (van Gelder *et al.*, 1997). Physiologically, the function of tyrosinase in fungi and the stimuli that cause tyrosinase activation are not yet understood, but it may be defense-related.

2.2.3.3.4 Bacterial tyrosinase

Bacterial tyrosinase genes of four *Streptomyces* species have been cloned and characterized (Bernan *et al.*, 1985; Huber *et al.*, 1985; Kawamoto *et al.*, 1993; Ikeda *et al.*, 1996). In *Streptomyces*, melanin is produced by a single operon designated *melC*. The *melC* operon contains only two genes, coding for a tyrosinase and a chaperone protein. The operon has

been studied in great depth, and has even been used to construct insertional-inactivation cloning vectors for various Gram-negative bacteria (Altenbuchner, 1988; Tseng *et al.*, 1990) and as a reporter gene to probe promotor sequences (Sugiyama *et al.*, 1990).

Approximately half of the total cellular tyrosinase is secreted, but strangely, the tyrosinase does not contain a signal peptide. The secretion was found to be mediated by the chaperone, which plays a dual role in the expression of tyrosinase: a copper chaperone (Chen *et al.*, 1992) and a secretion accessory (Leu *et al.*, 1992). The newly synthesized chaperone and apotyrosinase form a transient association, perhaps even while still being translated. The resulting complex – which is folded in a three-dimensional structure – is then secreted without first being unfolded. While associated, the chaperone transfers two copper ions to the apotyrosinase. This induces a conformational change in both the chaperone and the newly activated tyrosinase, and the complex dissociates.

The copper-transfer and secretion functions of the chaperone can be decoupled, as seen when the hydrophobic core of the chaperone signal peptide is mutated (Leu *et al.*, 1992). The apotyrosinase is then still activated, but accumulates intracellularly. It is interesting to note that when the positively charged amino acids at the beginning of the signal peptide was mutated, the chaperone did lose its ability to activate the apotyrosinase, indicating that the positively charged region is also involved in copper transfer or perhaps the folding of the chaperone itself.

It has been found that the *Streptomyces* strain *S. castaneoglobisporus* HUT6202 has a much higher ability to synthesize melanin pigments than other *Streptomyces* species (Ikeda *et al.*, 1996). This is not due to higher expression levels, but rather due to a very active chaperone, which efficiently facilitate the incorporation of copper into apotyrosinase. The *S. castaneoglobisporus melC* operon has therefore been used to explore the possibilities of commercial microbial phenol waste removal (Han *et al.*, 1994) and pigment production (Solaiman and Somkuti, 1995).

2.2.3.4 Genetic engineering of melanin biosynthesis in flowers

The fact that all tyrosinases – and actually also all other melanin-forming laccases, catalases and peroxidases – are copper enzymes, is the deciding factor when selecting a melanin biosynthetic pathway to express in flowers. Copper does not simply occur free inside cells,

easily available to any copper enzyme. The same reactivity of copper that makes it so useful in redox reactions also contributes to its inherent toxicity, should it ever occur free inside the cell. Cells have therefore developed an elaborate system of intracellular copper-trafficking pathways through which copper is accumulated and distributed while preventing its toxic effects (Harrison *et al.*, 2000; Himelblau and Amasino, 2000; Rosenzweig and O'Halloran, 2000). An important component of these pathways are a variety of cytoplasmic copper chaperones. These high-affinity copper proteins quickly scavenge any free copper, sequester it in a non-reactive form and deliver it to copper enzymes and copper transporters. During the transfer of copper from a copper chaperone to a copper enzyme, the two proteins interact closely through a key-and-lock mechanism (Harrison *et al.*, 2000). The copper transfer process is therefore highly specific, and each copper enzyme has its very own copper chaperone. The copper enzyme can receive copper only from its own chaperone or homologous chaperones from other species – not from any other copper chaperone.

This implies that whenever a tyrosinase or any other copper enzyme is expressed, its copper chaperone must also be present. For the expression of melanin in flowers, then, only one possibility remains: the *Streptomyces* tyrosinases, which are the only tyrosinases whose copper chaperones have been identified (Chen *et al.*, 1992).

The expression of the *Streptomyces antibioticus* tyrosinase and chaperone in tobacco plants has, in fact, been attempted by McBride and Stalker (2001). They tried two strategies: expressing the *melC* genes without any modification, or expressing the *melC* genes with plastid-targeting peptides added to each gene. When the enzymes were targeted to the plastids, melanin was indeed produced and dark regions appeared in various parts of the plants. However, the effects on the plants were highly deleterious. The growth of the plants was severely stunted, leaves were narrow and meristems were often aborted. The causes of these symptoms were not investigated, but might involve a depletion of the tyrosine amino acid pools in the plastids, or perhaps a toxic build-up of the insoluble, reactive melanin inside the plastid, turning the plastid into a kind of melanosome.

When the *melC* genes were expressed without modification, no effect on the phenotype could be observed (McBride and Stalker, 2001). This led the authors to assume that the concentration of free tyrosine in the cytosol is not high enough for melanin synthesis. However, they did not take into account that the chaperone has a strong signal peptide. The plant cell would most likely recognize the signal peptide and translocate the nascent polypeptide long before the tyrosinase can be activated. And even in the unlikely scenario

that the signal peptide is not recognized, such a prominent hydrophobic domain can easily cause aggregation and instability of the chaperone.

A much better strategy would be to express the chaperone with mutations in the hydrophobic core of the signal peptide. This should allow the chaperone to remain in the cytoplasm without affecting its ability to activate the tyrosinase (Leu *et al.*, 1992). Because any remaining large, hydrophobic regions may still destabilize the chaperone, the mutations should involve not one, but several of the amino acids. Furthermore, the tyrosinase and chaperone genes should rather come from the *S. castaneoglobisporus melC* operon than from *S. antibioticus*, since the *S. castaneoglobisporus* chaperone is far more active (Ikeda *et al.*, 1996). Such expression of the melanin-biosynthetic pathway in the cytosol is not expected to be deleterious – unlike expression in the plastids – as the large, hydrophobic melanin molecules would safely be deposited in the cell wall (Harvey *et al.*, 2002; Karam, 1998).

hapter 3

Materials & Methods

3.1 Materials

Plant tissue culture media and hormones were obtained from Sigma Chemical Company. Enzymes used during molecular cloning were obtained from Roche, Promega or Amersham. All other chemicals used were of analytical grade. *Zea mays* cv. PNR473 (Advanta) was obtained from Stephanus de Meillon (University of Pretoria, South Africa). *Nicotiana tabacum* cv. Samsun and *Agrobacterium tumefaciens* C58 were obtained from Karl Kunert (University of Pretoria, South Africa). Custom-made primers were ordered from either Roche or Gibco BRL. SP6 en T7 primers were obtained from Promega.

Plasmids used. Several different plasmids were used as sources for the promoter, terminator and genes during construction of the expression vectors. Plasmid pUCM-2 (Ikeda *et al.*, 1996), containing *the S. castaneoglobisporus melC* operon, was obtained from Masanori Sugiyama (Hiroshima University, Japan). Plasmid pSLH8 (Hart and Woods, 1992), containing the *Rhodococcus* indole dioxygenase gene, was obtained from Sharon Reid (University of Cape Town, South Africa). Plasmid pKT421 (Kawasaki *et al.*, 1993), containing the *E. aerogenes* tryptophanase gene, was obtained from Kosei Kawasaki (Hokkaido University, Japan). Plasmid pAHC27 (Christensen and Quail, 1996), containing the nopaline synthase terminator, was obtained from Peter Quail (University of California, California). Plasmid pCAMBIA2300 (Roberts *et al.*, 1997), containing the Cauliflower Mosaic Virus 35S promoter, was obtained from the Center for the Application of Molecular Biology to International Agriculture (CAMBIA, Australia), and was also used as a binary vector during plant transformation. Cloning vectors pGEM®-T Easy and pGEM®-3Zf+ were obtained from Promega.

3.2 General molecular techniques

Standard cloning techniques. Competent *E. coli* HB101 cells with competences of approximately 1x 10^9 cfu were prepared and transformed as described by Inoue *et al.* (1990). Plasmid DNA was isolated from *E. coli* cultures using Nucleospin DNA extraction kits (Macherey-Nagel) according to the manufacturer's instructions. Purification of DNA from agarose gels was performed with QIAquick gel extraction kits (QIAGEN) according to the manufacturer's instructions. All other standard molecular cloning techniques were performed as described in Sambrook *et al.* (1989).

Polymerase chain reactions and DNA sequencing. Polymerase chain reactions (PCRs) were performed using SuperTherm *Taq* DNA polymerase and buffers (Southern Cross Biotechnology), and a GeneAmp PCR System 9700 (Applied Biosystems). DNA sequencing was performed through the dideoxy-DNA chain-termination method (Sanger *et al.*, 1977), using an ABI PRISM™ automated sequencer (Perkin Elmer) and BigDye Terminator Cycle Sequencing Reaction Kits (Perkin Elmer). Whenever GC-rich regions had to be sequenced, 5% DMSO was added to the cycle sequencing reaction, and the denaturation temperature was lowered to 95°C.

3.3 Isolation and characterization of maize Bx1 *cDNA*

Unlike the other genes used during this study, the maize *Bx1* gene could not be obtained as a ready-made insert in a plasmid, and therefore had to be isolated anew from the cultivar PNR473 using a reverse transcription PCR (RT-PCR). To test whether the PNR473 cultivar does indeed produce DIMBOA, one-week-old shoots were dipped in a ferric chloride reagent (50g $FeCl_3 6H_2O$ in 500 ml methanol and 5 ml of 10M HCL). Development of a blue colour in the shoots was taken as evidence of the presence of DIMBOA (Wilkes *et al.*, 1999).

Reverse transcription. Total RNA was extracted from one-week-old shoots using a guanidinium isothiocyanate-based extraction buffer containing 2% (w/v) high molecular weight polyethylene glycol (HMW-PEG) as described in Gehrig *et al.* (2000). The concentrations were determined spectrometrically. Approximately 10 µg RNA was added to 400 pmol oligo dT_{15} primer in a final volume of 21 µl. The mixture was incubated at 70°C for 10 min and then snap-cooled on ice. The remaining components of the reverse transcription reaction were then added so that the reaction contained 1x reverse transcription buffer, 10 mM dithiothreitol (DTT), 50 U avian myeloblastosis virus reverse transcriptase, and 1 mM each dNTPs in a final volume of 40 µl. The reaction was incubated at 42°C for 60 min.

Polymerase chain reaction. 3 µl of the reverse transcription reaction was used as template in a 50 µl PCR containing 1x PCR buffer with 1.5 mM $MgCl_2$, 2 U *Taq* DNA polymerase and 200 µM each dNTPs. At first only the forward primer (5'-CAG GAT CCG CCA CCA TGG CTT TCG CGC CCA-3') was added to the reaction (to a final concentration of 1 µM). The reaction was incubated at 95°C for 30 sec, 55°C for 30 sec and 72°C for 1 min. The reverse primer (5'-CTT CTA GAT CAT GGC AGC GCG TTC TTC-3') was then also added to a final

concentration of 1 μM, and the PCR was continued for 35 cycles of 95°C for 30 sec, 65°C for 10 sec, 72°C for 50 sec.

Sequencing of the cDNA. The RT-PCR product was sequenced using the forward and reverse primers to determine the sequence of the gene. The RT-PCR product was then cloned into pGEM®-T and the individual clones sequenced using the T7 and SP6 primers until a clone with the expected sequence was found. The resulting sequence was searched against the GenBank and EMBL nucleotide databases using the BLASTN algorithm (Altschul *et al.*, 1997).

3.4 Sequencing of melC genes

Since all the PCR products of the *S. castaneoglobisporus melC* genes had sequences differing from the published sequences, these genes were sequenced to verify the observed sequences. To test whether the tyrosinase and chaperone expressed by pUCM-2 are functional, the *E. coli* culture was grown in LB-medium supplemented with 2 mM tyrosine and 0.1 mM $CuCl_2$. A functional tyrosinase and chaperone should cause strong black pigmentation after overnight growth in such a supplemented medium, due to melanin production (Tseng *et al.*, 1990). pUCM-2 was digested with *Eco* RI, and the two fragments that contained the *melC* operon (both approximately 720bp in length) were subcloned into pGEM®-3Zf+. Five clones of each of the fragments were sequenced using SP6 and T7 primers. The resulting sequences were searched against the GenBank and EMBL nucleotide databases using the BLASTN algorithm (Altschul *et al.*, 1997).

3.5 Construction of plant expression vectors

Three binary expression vectors were constructed to test expression of the novel pigmentation pathways in plants. The first vector, pC-ido-tnaK (Figure 3.1b), was used to test expression of indigo and consists of the pCAMBIA2300 binary vector carrying the *Rhodococcus* indole dioxygenase (*ido*) and the *Enterobacter aerogenes* tryptophanase (*tnaK*) genes. Each gene is under control of a Cauliflower mosaic virus 35S (CaMV35S) promoter, and each has a nopaline synthase (NOS) terminator at its 3' end. The second vector, pC-tyrC-melC1 (Figure 3.1c), contains the two genes of the *Streptomyces castaneoglobisporus* melanin operon, *tyrC*

and *melC1*. Again, each of these genes is provided with a CaMV35S promoter and a NOS terminator. The third vector, pC-tyrC-melC1Δ (Figure 3.1d), is the same as pC-tyrC-melC1 except that part of the signal sequence of *melC1* has been excised.

Addition of restriction sites and Kozak sequences. To construct these vectors, appropriate restriction sites first had to be introduced at the ends of each of the genes, promoter and terminator. Also, a Kozak sequence (Kozak, 1997) was added to each of the genes to ensure accurate and efficient translation of the prokaryotic genes in plants. This was accomplished through PCR amplification of each of these regions, using primers that contain these restriction sites and Kozak sequences. The primers used to amplify these regions are shown in table 3.1. All PCRs contained standard concentrations of reagents (1x PCR buffer with 1.5 mM $MgCl_2$, 200 µM each dNTPs, 4 U *Taq* DNA polymerase per 100 µl reaction, 1 µM each forward and reverse primers, and approximately 1 ng plasmid DNA template). Most of these regions were routinely amplified using standard temperature cycling consisting of 30 cycles of 95°C for 30 sec, 55°C for 30 sec, 72°C for 1 min. The *melC* genes, however, could only be dependably amplified using 40 cycles of 97,5°C for 10 sec, 60°C for 30 sec, 72°C for 30 sec. Since the high denaturing temperature inactivates the *Taq* polymerase, the cycling was interrupted after 20 cycles to add an additional 3 U Taq per 100 µl reaction, and the PCR was then continued for another 20 cycles.

Construction of expression cassettes. After PCR amplification of the CaMV35S promoter from pCAMBIA2300, the product was cloned into pGEM®-T Easy and sequenced to find a plasmid that contained the promoter in a forward direction and with no mutations. The NOS terminator was amplified from pAHC27, the PCR product digested with *Xba* I and *Kpn* I, and then cloned into pGEM®-3Zf+. After sequencing the insert to ensure that its sequence was correct, the terminator was excised with with *Pst* I and *Nde* I, and cloned into the promoter-containing pGEM®-T Easy. This created an expression cassette inside pGEM®-T Easy and the resulting plasmid was named pG-3-N. The four genes (*ido*, *tnaK*, *tyrC* and *melC1*) were each amplified by PCR, digested with *Bam* HI and *Xba* I, and cloned into the expression cassette of pG-3-N. The resulting plasmids were sequenced, using the same primers as during the PCR amplification, to find clones with the expected sequence. These steps resulted in the four plasmids pGT-3-ido-N, pGT-3-tnaK-N, pGT-3-tyrC-N and pGT-3-melC1-N, which contain the *ido*, *tnaK*, *tyrC* and *melC1* genes, respectively.

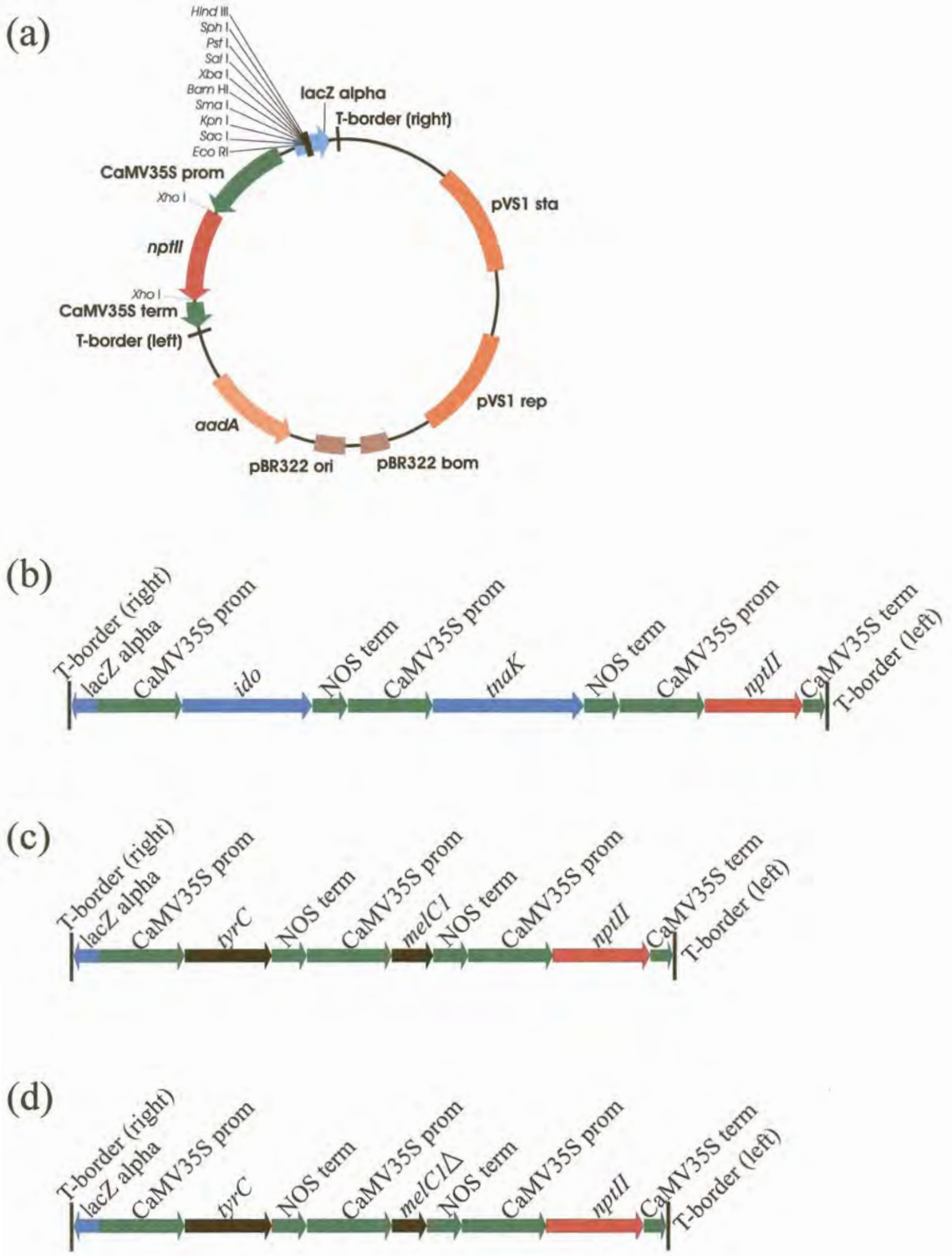

Figure 3.1: Binary expression vectors used for plant transformation. (a) Vector map of the original pCAMBIA2300 binary vector (Roberts *et al.*, 1997). (b) T-DNA region of pC-ido-tnaK. (c) T-DNA region of pC-tyrC-melC1. (d) T-DNA region of pC-tyrC-melC1Δ. Abbreviations: CaMV35S prom, CaMV35S promoter with double enhancer; NOS term, nopaline synthase terminator; *nptII*, neomycin phosphotransferase; *aaadA*, aminoglucoside phosphotransferase.

Table 3.1: Primers used during construction of binary expression vectors

Region amplified	Polarity	Sequence
ido	Forward	*Bam*H I Kozak 5' CAGGATCCGCCACCATGGACATCACCCGCACCGAACTCATG 3'
	Reverse	*Xba* I 5' AATCTAGATCCCAACAGCTCG 3'
tnaK	Forward	*Bam*H I Kozak 5' GGATCCGCCACCATGAAACGTATCCCTGAG 3'
	Reverse	*Xba* I 5' AATCTAGAGCGCCCTGATAGG 3'
tyrC	Forward	*Bam*H I Kozak 5' CAGGATCCGCCACCATGACCGTACGCAAG 3'
	Reverse	*Xba* I 5' GATCTAGATCAGGCGTCGAAC 3'
melC1	Forward	*Bam*H I Kozak 5' GGATCCGCCACCATGCCCGAGATCACC 3'
	Reverse	*Xba* I 5' GATCTAGAAGTCCTTGCTGGG 3'
CaMV35S promoter	Forward	*Kpn* I 5' AAGGTACCGTATTGGCTAGAGCAG 3'
	Reverse	*Bam*H I 5' AAGGATCCTGCGAAAGCTCGAG 3'
NOS terminator	Forward	*Xba* I 5' ACTCTAGAGCTCGAATTTCCC 3'
	Reverse	*Kpn* I 5' AAGGTACCGATCTAGTAACATAG 3'

Construction of pC-ido-tnak. The expression cassette of pGT-3-tnaK-N was excised with *Kpn* I and cloned into pGEM®-3Zf+. The resulting plasmids were screened, using restriction enzyme digestions, to find one in which the expression cassette had inserted with the NOS terminator next to the plasmid's *Eco*R I site. This plasmid was named pGZ-3-tnaK-N. The expression cassette of pGT-3-ido-N was then excised with *Eco*R I, blunt-ended and cloned into the *Sma* I site of pGZ-3-tnaK-N. A plasmid containing the second expression cassette in the same direction as the first was selected, and named pGZ-ido-tnaK. The *Hind* III-*Eco*R I fragment of pGZ-ido-tnaK was subcloned into pCAMBIA2300 to create the binary expression vector pC-ido-tnaK.

Construction of pC-tyrC-melC1. The expression cassette of pGT-3-tyrC-N was excised with *Kpn* I, blunt-ended and cloned into pGEM®-3Zf+ that was digested with *Xba* I and also blunt-ended. A plasmid was selected in which the expression cassette had inserted with the NOS terminator next to the *Eco*R I site, and named pGZ-3-tyrC-N. The expression cassette of pGT-3-melC1-N was excised with *Kpn* I, blunt-ended and cloned into the *Sma* I site of pGZ-3-tyrC-N. A plasmid that contained the second expression cassette in the same direction as the first was selected, and named pGZ-tyrC-melC1. The *Hind* III-*Kpn* I fragment of pGZ-tyrC-melC1 was cloned into pCAMBIA2300 to create pC-tyrC-melC1.

Construction of pC-tyrC-melC1Δ. To create pC-tyrC-melC1Δ, pGZ-tyrC-melC1 was simply digested with *Sac* II and recircularized to remove the small *Sac* II fragment, resulting in plasmid pGZ-tyrC-melC1Δ. Finally, The *Hind* III-*Kpn* I fragment of pGZ-tyrC-melC1Δ was cloned into pCAMBIA2300 to create pC-tyrC-melC1Δ.

Confirmation of vector construction. To show that all the genes were present in their appropriate binary expression vectors, and each with an upstream CaMV35S promotor and a downstream NOS terminator, each gene was amplified from its vector together with either its promotor or terminator. To amplify a promotor-gene region, the forward primer of the promotor and the reverse primer of the gene were used. To amplify a gene-terminator region, the forward primer of the gene and the reverse primer of the terminator were used. PCR conditions were as given above.

3.6 Transformation of tobacco

Transformation and regeneration. In three separate reactions, *A. tumefaciens* was transformed with either pC-ido-tnaK, pC-tyrC-melC1 or pC-tyrC-melC1Δ, using a freeze-thaw method (An *et al.*, 1988). Transformation and regeneration of tobacco were performed through a leaf disk method, using these transformed *A. tumefaciens* cultures, as described in Curtis *et al.* (1995). Ninety plantlets were regenerated in total – thirty for each of the three expression vectors. Regenerated plants were maintained under continual light conditions at 24°C on a growth medium containing MS basal salts and vitamins (Murashige and Skoog, 1962), 2% sucrose and 1% agar. The plants were micropropagated through nodal cuttings. For the first two months after transformation, the medium also included 50 µg/ml kanamycin and 500 µg/ml cefotaxime.

Confirmation of transformation. To confirm that the plants were indeed transformed, genomic DNA was extracted from young leaves using a cetyltriethylammonium bromide (CTAB) method (Sambrook *et al.*, 1989), and the presence of the transgenes or the CaMV35S promoter was detected by PCR amplification. The same PCR parameters as given in section 3.5 were used, except that 1 µg of genomic DNA was used as template.

3.7 Testing of transgene expression

Since none of the transformed plants turned visibly blue or black, expression of the transgenes had to be confirmed through Northern dot-blot analysis of RNA extracts from transformed plants. The plants were tested approximately six months after transformation.

Dot-blots of bacterial mRNA. To ensure that the dot-blots can, in fact, detect RNA, the experiment was first performed using total RNA extracts from *E. coli* cultures that are known to express *ido*. *E. coli* cultures carrying the pSLH8 plasmid were grown to early log phase in LB-medium containing 30 µM isopropyl-β-D-thiogalactopyranoside (IPTG). To collect the cells, 1.5 ml of the culture was centrifuged at 1000 x g for 1 minute. The pelleted cells were gently resuspended in 100 µg protoplasting solution (15 mM Tris-HCl, pH 8.0, 0.45 M sucrose, 8 mM EDTA, 40 µg/ml lysozyme) and incubated at room temperature for 10 minutes. 1 ml TriZOL reagent (GibcoBRL) was then added and the remaining steps of the extraction proceeded as described by the manufacturer. Half of the RNA extract was incubated in a DNase solution (10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 0.1 mM $CaCl_2$, 5 units RNase-free DNase) at 37°C for 30 min to remove contaminating plasmid DNA. The remaining half of the RNA extract was incubated in the same DNase solution, but with 5 µg/ml RNase A added to the reaction. This served as a negative control to show that the dot-blots actually detect RNA and not some remaining pSLH8 DNA. As another negative control to ensure that the dot-blots do not detect ribosomal RNA, total RNA was also extracted from an *E. coli* culture carrying the pGEM®-3Zf+ plasmid.

Dilution series of the RNA extracts were blotted onto positively charged nylon membranes (Roche) using a dot-blot vacuum manifold (Bio-Rad) as described in Sambrook *et al.* (1989). The RNA extract that was treated with RNase was blotted onto a separate membrane and kept separate from the other membranes, because of the presence of RNase. A dilution series of pC-ido-tnaK was also blotted onto each membrane as a positive control. A DIG-labeled DNA probe was generated from a PCR product of *ido* using a DIG High Prime DNA Labeling and

Detection Kit (Roche). Hybridization was carried out as described in the DIG System User's Guide for Filter Hybridization (Roche).

Testing expression of the four pigmentation genes. Total RNA was extracted from young leaves using a guanidinium isothiocyanate-based extraction buffer containing 2% (w/v) HMW-PEG as described in Gehrig *et al.* (2000). The concentrations were determined spectrometrically, and approximately 10 µg total RNA of each plant was used per dot. The samples were blotted onto positively charged nylon membranes as above. Dilution series of both pC-ido-tnaK and pC-tyrC-melC1Δ were also blotted onto each membrane as a reference to estimate the concentration of the target mRNA. DIG-labeled DNA probes were generated from PCR products of each of the four genes using a DIG High Prime DNA Labeling and Detection Kit (Roche). Hybridization was carried out as described in the DIG System User's Guide for Filter Hybridization (Roche). If a membrane had to be re-probed, the blue colour precipitate was first removed by washing the membrane with dimethyl formamide (DMF) at 60°C. The hybridized probes were then stripped as described in Ausubel *et al.* (1992).

Testing expression of kanamycin resistance gene. Expression of the kanamycin resistance gene (*nptII*) was tested using dot-blots as described above, except that a DIG-labeled probe was generated from the *nptII*-containing *Xho* I-fragment of pCAMBIA2300. As an additional test, the plants were placed onto medium supplemented with 50 µg/ml kanamycin.

Chapter 4

Results

4.1 Isolation and characterization of maize Bx1 *cDNA*

Unlike the other genes used during this study, the maize *Bx1* gene could not be obtained as a ready-made insert in a plasmid. It was therefore isolated from the cultivar PNR473 using a RT-PCR. The quality of the extracted total RNA is shown in Figure 4.1A, while the RT-PCR product is shown in Figure 4.1B.

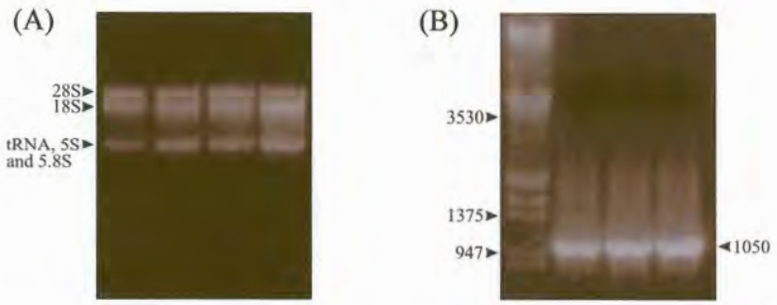

Figure 4.1: (A) Total RNA extracted from maize cv. PNR473 and used for RT-PCR. (B) *Bx1* RT-PCR product, approximately 1050bp in size, next to DNA molecular weight marker III.

The RT-PCR product was sequenced to determine the correct sequence of the gene. The resulting electropherograms consistently showed double peaks at specific bases, indicating the possibility that more than one *Bx1* copy were expressed. As expected, when the RT-PCR product was cloned into pGEM®-T and the individual clones sequenced, two distinct sequences were found.

The nucleotide sequences the PNR473 *Bx1* genes are shown in Figure 4.2, compared with the gene originally isolated from the CG00237 cultivar by Kramer and Koziel (1995). The corresponding amino acid sequences are shown in Figure 4.3. All of the differences between the nucleotide sequences of the three genes are synonymous substitutions, except for three nucleotide insertions present in the PNR473 genes (at positions 208, 250 and 251) relative to the CG00237 gene. These insertions cause a block of 16 amino acids in the first exon of the CG00237 gene to differ from the corresponding 17 amino acids of the PNR473 genes.

The PNR473 sequences were searched against the GenBank and EMBL nucleotide databases using the BLASTN algorithm (Altschul *et al.*, 1997). For both sequences the most significant alignment was not to the *Bx1* gene from the CG00237 cultivar, but to an expressed sequence tag from the PCO064449 cultivar (GenBank accession number AY104439), with E values of

zero. The PNR473 sequences were submitted to GenBank as accession numbers AY254103 and AY254104.

```
(A)  atggctttcg cgcccaaaac gtcctcctcc tcctcgctgt cctcggcgtt gcaggcagct cagtcgccgc     70
(B)  atggctttcg cgcccaaaac gtcctcctcc tcctcgctgt cctcggcgtt gcaggcagct cagtcgccgc     70
(C)  atggctttcg cgcccaaaac gtcctcctcc tcctcgctgt cctcggcgtt gcaggcagct cagtcgccgc     70

(A)  cgctgctcct gaggcggatg tcgtcgaccg caacaccgag acggaggtac gacgcggccg tcgtcgtcac    140
(B)  cgctgctcct gaggcggatg tcgtcgaccg caacaccgag acggaggtac gacgcggccg tcgtcgtcac    140
(C)  cgctgctcct gaggcggatg tcgtcgaccg caacaccgag acggaggtac gacgcggccg tcgtcgtcac    140

(A)  taccaccacc actgctagag ctgcggcggc tgctgtcacg gttcccgccg ccccgccgca ggcgttc-cgc   210
(B)  taccaccacc actgctagag ctgcggcggc tgctgtcacg gttcccgccg ccccgccgca ggcgtccggcgc   211
(C)  taccaccacc actgctagag ctgcggcggc tgctgtcacg gttcccgccg ccccgccgca ggcg---ggcgc   211

(A)  cggcgccggt gccgccaaag caagcggcgg cacccgc---ga ggaggagccg tccggtgtcg gacaccatgg   280
(B)  cggcgccggt gccgccaaag caagcggcgg cacccgccgaga ggaggagccg tccggtgtcg gacaccatgg   283
(C)  cggcgccggt gccgccaaag caagcggcgg cacccgccgaga ggaggagccg tccggtgtcg gacaccatgg   283

(A)  cggcgctcat ggccaagggc aagacggcgt tcatcccgta catcaccgcc ggcgacccgg acctagcgac    350
(B)  cggcgctcat ggccaagggc aagacggcgt tcatcccgta catcaccgcc ggcgacccgg acctagcgac    353
(C)  cggcgctcat ggccaagggc aagacggcgt tcatcccgta catcaccgcc ggcgacccgg acctagcgac    353

(A)  gacggccgag gcgctgcgtc tgctggacgg ctgtggcgcc gacgtcatcg agctgggcgt cccctgctcg    420
(B)  gacggccgag gcgctgcgtc tgctggacgg ctgtggcgcc gacgtcatcg agctgggcgt cccctgctcg    423
(C)  gacggccgag gcgctgcggc tgctggacgg ctgtggcgcc gacgtcatcg agctgggcgt cccctgctcg    423

(A)  gacccctaca tcgacgggcc catcatccag gcgtcggtgg cgcgggctct ggccagcggc accaccatgg    490
(B)  gacccctaca tcgacgggcc catcatccag gcgtcggtgg cgcgggctct ggccagcggc accaccatgg    493
(C)  gacccctaca tcgacgggcc catcatccag gcgtcggtgg cgcgggctct ggccagcggc accaccatgg    493

(A)  acgccgtgct ggagatgctg agggaggtga cgccggagct gtcgtgcccc gtggtgctcc tctcctacta    560
(B)  acgccgtgct ggagatgctg agggaggtga cgccggagct gtcgtgcccc gtggtgctcc tctcctacta    563
(C)  acgccgtgct ggagatgctg agggaggtga cgccggagct gtcgtgcccc gtggtgctcc tctcctacta    563

(A)  caagcccatc atgtctcgca gcttggccga gatgaaagag gcgggggtcc acggtcttat agtgcctgat    630
(B)  caagcccatc atgtctcgca gcttggccga gatgaaagag gcgggggtcc acggtcttat agtgcctgat    633
(C)  caagcccatc atgtctcgca gcttggccga gatgaaagag gcgggggtcc acggtcttat agtgcctgat    633

(A)  ctcccgtacg tggccgcgca ctcgctgtgg agtgaagcca agaacaacaa cctggagctg gtgctgctga    700
(B)  ctcccgtacg tggccgcgca ctcgctgtgg agtgaagcca agaacaacaa cctggagctg gtgctgctga    703
(C)  ctcccgtacg tggccgcgca ctcgctgtgg agtgaagcca agaacaacaa cctggagctg gtgctgctga    703

(A)  caacaccagc cataccagaa gacaggatga aggagatcac caaggcttca gaaggcttcg tctacctggt    770
(B)  caacaccagc cataccagaa gacaggatga aggagatcac caaggcttca gaaggcttcg tctacctggt    773
(C)  caacaccagc cataccagaa gacaggatga aggagatcac caaggcttca gaaggcttcg tctacctggt    773

(A)  gagcgtcaac ggagtgacag gtcctcgcgc aaacgtgaac ccacgagtgg agtcactcat ccaggaggtt    840
(B)  gagcgtcaac ggagtgacag gtcctcgcgc aaacgtgaac ccacgagtgg agtcactcat ccaggaggtt    843
(C)  gagcgtcaac ggagtgacag gtcctcgcgc aaacgtgaac ccacgagtgg agtcactcat ccaggaggtt    843

(A)  aagaaggtga ctaacaagcc cgttgctgtt ggcttcggca tatccaagcc cgagcacgtc aagcagattg    910
(B)  aagaaggtga ctaacaagcc cgttgctgtt ggcttcggca tatccaagcc cgagcacgtc aagcagattg    913
(C)  aagaaggtga ctaacaagcc cgttgctgtt ggcttcggca tatccaagcc cgagcacgtc aagcagattg    913

(A)  cgcagtgggg cgctgacggg gtgatcatcg gcagcgccat ggtgaggcag ctgggcgaag cggcttctcc    980
(B)  cgcagtgggg cgctgacggg gtgatcatcg gcagcgccat ggtgaggcag ctgggcgaag cggcttctcc    983
(C)  cgcagtgggg cgctgacggg gtgatcatcg gcagcgccat ggtgaggcag ctgggcgaag cggcttctcc    983

(A)  caagcaaggc ctgaggaggc tggaggagta tgccaggggc atgaagaacg cgctgccatg a           1041
(B)  caagcaaggc ctgaggaggc tggaggagta tgccaggggc atgaagaacg cgctgccatg a           1044
(C)  caagcaaggc ctgaggaggc tggaggagta tgccaggggc atgaagaacg cgctgccatg a           1044
```

Figure 4.2: Nucleotide sequences of the indole synthase gene isolated by Kramer and Koziel (1995) from maize cv. CG00237 (A) in comparison with genes isolated from cv. PNR473 in this study (B and C). Bases at which the sequences differ are indicated by black shading. Deletions needed for alignment are indicated by dashes.

```
(A)   MAFAPKTSSS SSLSSALQAA QSPPLLLRRM SSTATPRRRY DAAVVVTTTT    50
(B)   MAFAPKTSSS SSLSSALQAA QSPPLLLRRM SSTATPRRRY DAAVVVTTTT    50
(C)   MAFAPKTSSS SSLSSALQAA QSPPLLLRRM SSTATPRRRY DAAVVVTTTT    50

(A)   TARAAAAAVT VPAAPPQ--- ---------- ---RRSRPVS DTMAALMAKG  100
(B)   TARAAAAAVT VPAAPPQEAP EAEVEENQRA APAERRSRPVS DTMAALMAKG  101
(B)   TARAAAAAVT VPAAPPQEAP EAEVEENQRA APAERRSRPVS DTMAALMAKG  101

(A)   KTAFIPYITA GDPDLATTAE ALRLLDGCGA DVIELGVPCS DPYIDGPIIQ  150
(B)   KTAFIPYITA GDPDLATTAE ALRLLDGCGA DVIELGVPCS DPYIDGPIIQ  151
(C)   KTAFIPYITA GDPDLATTAE ALRLLDGCGA DVIELGVPCS DPYIDGPIIQ  151

(A)   ASVARALASG TTMDAVLEML REVTPELSCP VVLLSYYKPI MSRSLAEMKE  200
(B)   ASVARALASG TTMDAVLEML REVTPELSCP VVLLSYYKPI MSRSLAEMKE  201
(C)   ASVARALASG TTMDAVLEML REVTPELSCP VVLLSYYKPI MSRSLAEMKE  201

(A)   AGVHGLIVPD LPYVAAHSLW SEAKNNNLEL VLLTTPAIPE DRMKEITKAS  250
(B)   AGVHGLIVPD LPYVAAHSLW SEAKNNNLEL VLLTTPAIPE DRMKEITKAS  251
(C)   AGVHGLIVPD LPYVAAHSLW SEAKNNNLEL VLLTTPAIPE DRMKEITKAS  251

(A)   EGFVYLVSVN GVTGPRANVN PRVESLIQEV KKVTNKPVAV GFGISKPEHV  300
(B)   EGFVYLVSVN GVTGPRANVN PRVESLIQEV KKVTNKPVAV GFGISKPEHV  301
(C)   EGFVYLVSVN GVTGPRANVN PRVESLIQEV KKVTNKPVAV GFGISKPEHV  301

(A)   KQIAQWGADG VIIGSAMVRQ LGEAASPKQG LRRLEEYARG MKNALP      346
(B)   KQIAQWGADG VIIGSAMVRQ LGEAASPKQG LRRLEEYARG MKNALP      347
(C)   KQIAQWGADG VIIGSAMVRQ LGEAASPKQG LRRLEEYARG MKNALP      347
```

Figure 4.3: Amino acid sequences encoded by the indole synthase gene isolated by Kramer and Koziel (1995) from maize cv. CG00237 (A) in comparison with the genes isolated from cv. PNR473 in this study (B and C). Amino acids at which the sequences differ are indicated by black shading. Deletions needed for alignment are indicated by dashes.

4.2 Sequencing of melC *genes*

Since all the PCR products of the *S. castaneoglobisporus melC* genes in the pUCM-2 plasmid had sequences differing from those given by Ikeda *et al.* (1996), these genes were sequenced to verify the observed sequences. The resulting sequences of *tyrC* and *melC1* are shown in Figures 4.5 and 4.6, compared to the published sequences. The *E.coli* culture from which pUCM-2 was extracted turned black during overnight growth in tyrosine- and $CuCl_2$-supplemented LB-medium, indicating that the tyrosine and chaperone were expressed and functional (Figure 4.4).

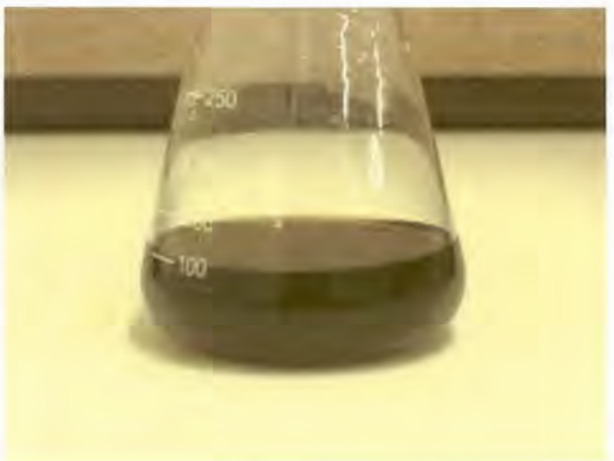

Figure 4.4: Overnight culture of *E. coli* carrying the pUCM-2 plasmid. The medium was supplemented with tyrosine and $CuCl_2$. The black colour indicates that the tyrosine and chaperone were expressed and functional.

```
(A)     atgaccgtac gcaagaacca ggccaccctg accgccgacga gaagcgccgg ttcgtcgcc      60
(B)     atgaccgtac gcaagaacca ggccaccctg accgccgacga gaagcgccgg ttcgtcgcc

(A)     gccgtcctcg aactcaagcg cagcgggcgc tacgacgagtt cgtccgcacg cacaacgaa      120
(B)     gccgtcctcg aactcaagcg cagcgggcgc tacgacgagtt cgtccgcacg cacaacgaa

(A)     ttcatcatgt cggacaccga cagcggggaa cggaccggcca ccgctccccc tcgttcctg      180
(B)     ttcatcatgt cggacaccga cagcggggaa cggaccggcca ccgctccccc tcgttcctg

(A)     ccctggcacc gcagattcct gctcgacttc gaacaggcgct gcagtccgtg gactcctcc      240
(B)     ccctggcacc gcagattcct gctcgacttc gaacaggcgct gcagtccgtg gactcctcc

(A)     gtcacgctgc cgtactggga ctggtccgcc gaccgcaccgt gcgggcctcg ctgtgggcg      300
(B)     gtcacgctgc cgtactggga ctggtccgcc gaccgcaccgt gcgggcctcg ctgtgggcg

(A)     ccggacttcc tcggcggcac cgggcgcag[c] [c]ccgacggccg ggtgatggac gggcc[g]ttc      360
(B)     ccggacttcc tcggcggcac cgggcgcag[g] [g]ccgacggccg ggtgatggac gggcc[a]ttc

(A)     gccgcgttca cgggcaactg gccgatcaac gtgcgggtcga cagccggacg tacctgcgc      420
(B)     gccgcgttca cgggcaactg gccgatcaac gtgcgggtcga cagccggacg tacctgcgc

(A)     cgctcgctcg gcggaagcgt ggccgaactg cccacgcgcgc cgaggtggag tcggtactg      480
(B)     cgctcgctcg gcggaagcgt ggccgaactg cccacgcgcgc cgaggtggag tcggtactg

(A)     gcgatatcgg cgtacgacct gccgccgtac aacagcgcctc ggagggcttc cgcaatcac      540
(B)     gcgatatcgg cgtacgacct gccgccgtac aacagcgcctc ggagggcttc cgcaatcac

(A)     ctggagggct ggcgcggggt caatctgcac aaccgcgtcca cgtgtgggtc ggcgggcag      600
(B)     ctggagggct ggcgcggggt caatctgcac aaccgcgtcca cgtgtgggtc ggcgggcag

(A)     atggccaccg gggtctcccc caacgacccg gtgttctggct gcaccacgcc tacgtcgac      660
(B)     atggccaccg gggtctcccc caacgacccg gtgttctggct gcaccacgcc tacgtcgac

(A)     aagctgtggg ccgagtggca gcggcggcac ccggactccgc gtacgtgccg acgggcggc      720
(B)     aagctgtggg ccgagtggca gcggcggcac ccggactccgc gtacgtgccg acgggcggc

(A)     acgccggacg tggtggacct gaacgagacc atgaagccctg gaacaccgtg cgcccggcg      780
(B)     acgccggacg tggtggacct gaacgagacc atgaagccctg gaacaccgtg cgcccggcg

(A)     gatctgctgg accacactgc ctactacacg ttcgacgcctg a            821
(B)     gatctgctgg accacactgc ctactacacg ttcgacgcctg a
```

Figure 4.5: Nucleotide sequence of the *S. castaneoglobisporus tyrC* gene as determined in this study (A), compared with the sequence given by Ikeda *et al.* (1996) (B). Bases at which the sequences differ are indicated by black shading.

```
(A)     atgcccgaga tcacccgccg ccgcgcgctc accgcggcag ccgccgtcgc cgcgaccgcc      60
(B)     atgcccgaga tcacccgccg ccgcgcgctc accgcggcag ccgccgtcgc cgcgaccgcc

(A)     tcggcggccg tcaccctcgc cgcccccgcc gcgtcggccg cgggccacca cgagcccgcc     120
(B)     tcggcggccg tcaccctcgc cgcccccgcc gcgtcggccg cgggccacca cgagcccgcc

(A)     gcgcccgagt ccttcgacga ggtctacaag ggccgccgga tacagggccg tccggcggcc     180
(B)     gcgcccgagt ccttcgacga ggtctacaag ggccgccgga tacagggccg tccggcggcc

(A)     ggcggccgcgc accaccacga acacggcggc ggatacgagg tgttcgtcga cggcgtgcag     240
(B)     ggcggccgcgc accaccacga acacggcggc ggatacgagg tgttcgtcga cggcgtgcag

(A)     ctgcacgtga tgcgcaacgc cgacggcagc tggatcagcg tcgtcagcca ctacgacccg     300
(B)     ctgcacgtga tgcgcaacgc cgacggcagc tggatcagcg tcgtcagcca ctacgacccg

(A)     gtgcccaccc cgcgcgccgc cgcccgtgcc gccgtggacg agctgcaggg cgccccgctg     360
(B)     gtgcccaccc cgcgcgccgc cgcccgtgcc gccgtggacg agctgcaggg cgccccgctg

(A)     ctgccgttcc ccgccaactg a      381
(B)     ctgccgttcc ccgccaactg a
```

Figure 4.6: Nucleotide sequence of the *S. castaneoglobisporus melC1* gene as determined in this study (A), compared with the sequence given by Ikeda *et al.* (1996) (B). Bases at which the sequences differ are indicated by black shading.

The *S. castaneoglobisporus tyrC* and *melC1* sequences determined in this study were submitted to GenBank as accession numbers AY254101 and AY254102, respectively. The sequences were searched against the GenBank and EMBL nucleotide databases using the BLASTN algorithm (Altschul *et al.*, 1997). For both *tyrC* and *melC1* the most significant alignment was to the corresponding *S. galbus* genes (GenBank accession number X95705), with E values of zero for the tyrosinase genes and 10^{-118} for the chaperone genes.

Figures 4.7 and 4.8 show the amino acid sequences of the *S. castaneoglobisporus* tyrosinase and chaperone as determined in this study, compared with the sequences given by Ikeda *et al.* (1996), and also the sequences of the tyrosinases and chaperones of *S. antibioticus* (Bernan *et al.*, 1985), *S. lincolnensis* (GenBank accession number X95703) and *S. galbus* (GenBank accession number X95705). From these figures it can be seen that the amino acid sequences of the *S. castaneoglobisporus* tyrosinase and chaperone as determined in this study show closer similarity to the consensus sequences of *Streptomyces* tyrosinases and chaperones than the sequences given by Ikeda *et al.* (1996) do. Of the six differences between the nucleotide sequences determined in this study and those given by Ikeda *et al.* (1996), two are synonymous substitutions (at positions 417 in *tyrC* and 18 in *melC1*), while one causes a substitution of proline with threonine at a non-conserved position in the consensus sequence (position 391 in *tyrC*). The remaining three (at positions 178 and 185 in *melC1*, and 391 in *tyrC*) cause their corresponding amino acids to be replaced by those found in the consensus sequence.

```
(A)   MTVRKNQATL TADEKRRFVA AVLELKRSGR YDEFVRTHNE FIMSDTDSGE RT--GHRSPS-FL   60
(B)   MTVRKNQATL TADEKRRFVA AVLELKRSGR YDEFVRTHNE FIMSDTDSGE RT--GHRSPS-FL   60
(C)   MTVRKNQASL TAEEKRRFVA ALLELKRTGR YDAFVTTHNA FILGDTDNGE RT--GHRSPS-FL   60
(D)   MTVRKNQATL TADEKRRFVT AVLSSSAA-R YDTFVTTHNE FIVADTDNGE RT--GHRSPS-FL   59
(E)   MTVRKNQAAL TADEKRRFVA AVLELKRNGR YDEFVRTHNE FIMSDTRTGR RGGPGHRLPLPFL   63

(A)   PWHRRFLLDF EQALQSVDSS VTLPYWDWSA DRTVRASLWA PDFLGGTGR  DGRVMDGPF   120
(B)   PWHRRFLLDF EQALQSVDSS VTLPYWDWSA DRTVRASLWA PDFLGGTGR  DGRVMDGPF   120
(C)   PWHRRFLLEF ERALQSVDAS VALPYWDWST DRSTRSSLWA PDFLGGTGR  DGQVMDGPF   120
(D)   PWHRRFLLEF ERALQSVDAS VALPYWDWST DRSARSSLWA PDFLGGTGR  NGRVTDGPF   119
(E)   PWHRRFLLDF EQALQSVDSS VALPYWDWST DRTVRASLWA PDFLGGTGR  DGRVMDGPF   123

(A)   AAFTGNWPIN VRVDSRTYLR RSLGGSVAEL PTRAEVESVL AISAYDLPPY NSASEGFRNH   180
(B)   AAFTGNWPIN VRVDSRTYLR RSLGGSVAEL PTRAEVESVL AISAYDLPPY NSASEGFRNH   180
(C)   AASAGNWPIN VRVDGRTFLR RALGAGVSEL PTRAEVDSVL AMATYDMAPW NSGSDGFRNH   180
(D)   RAATGVWPIT VRLDGRTYLR RALGGAGREL PTRAEVDSVL SIPTYDMAPW NSASDGFRNH   179
(E)   AASTGNWPVN VRVDGRTFLR RSLGTGVREL PTRAEVDSVL SMATYDMAPY NSASDGFRNH   183

(A)   LEGWRGVNLH NRVHVWVGGQ MATGVSPNDP VFWLHHAYVD KLWAEWQRRH PDSAYVPTGG   240
(B)   LEGWRGVNLH NRVHVWVGGQ MATGVSPNDP VFWLHHAYVD KLWAEWQRRH PDSAYVPTGG   240
(C)   LEGWRGVNLH NRVHVWVGGQ MATGVSPNDP VFWLHHAYID KLWAEWQRRH PSSPYLPGGG   240
(D)   LEGWRGVNLH NRVHVWVGGQ MATGVSPNDP VFWLHHAYID KLWAQWQRRH RTPAYVPAAG   239
(E)   LEGWRGVNLH NRVHVWVGGQ MATGVSPNDP VFWLHHAYNR QLWAEWQRRH PGAGYVPTGG   243

(A)   TPDVVDLNET MKPWNTVRPA DLLDHTAYYT FDA          273
(B)   TPDVVDLNET MKPWNTVRPA DLLDHTAYYT FDA          273
(C)   TPNVVDLNET MKPWNDTTPA ALLDHTRHYT FDV          273
(D)   TPDVVDLDET MKPWHDSSPA DLLDHTGHYT FDTD         273
(E)   TPDVVDLNDT MKPWNDVRPA DLLTHTAHYT FDV          276
```

Figure 4.7: Amino acid sequence of (A) the *S. castaneoglobisporus* tyrosinase as determined in this study, compared with (B) the sequence given by Ikeda *et al.* (1996) and also the tyrosinase sequences of (C) *S. antibioticus* (Bernan *et al.*, 1985), (D) *S. lincolnensis* (GenBank accession number X95703) and (E) *S. galbus* (GenBank accession number X95705). Amino acids at which sequences A and B differ are indicated by black shading. Deletions needed for alignment are indicated by dashes.

```
(A)   MPEITRRRAL TAAAAVAATA SAAVTLAAPA ASAAGHHEPA APES----------------FDEVYK   50
(B)   MPEITRRRAL TAAAAVAATA SAAVTLAAPA ASAAGHHEPA APES----------------FDEVYK   50
(C)   MPELTRRRAL GAAAVVAAGV PLVALPAARA DDR-GHHT-- -PEVPGNPAASGAPAA----FDEIYK   58
(D)   MPRLTRRRAL TAAAALASGA GAGAGAQAAA APGAAAHDHG SPDVPLPCSLLDLLPLLPLLLDEVIK   66
(E)   MPDITRRRAY TTAAAVAATA SAAAPTAAPA ATAAARHDHT APDS----------------FDEVYK   50

(A)   GRRIQGRPA----------------  GGAHHHEHGG GYEVFVDGVQ LHVMRNADGS WISVVSHYDP   100
(B)   GRRIQGRPA----------------  GGAHHHEHGG GYEVFVDGVQ LHVMRNADGS WISVVSHYDP   100
(C)   GRRIQGRTVTDGGGHHGGGHGGDGH  GG-HH---GG GYAVFVDGVE LHVMRNADGS WISVVSHYEP   120
(D)   GRGIQGGP----------------HR  GG-HHH--GT GYAVFLDGVE LHVMRNADGS WISVVSHYDP   114
(E)   GRRIQGGPAS--------------  GG-HHHEHGG GYAVFVDGVQ LHVMQNADGT WISVVSHYAP   100

(A)   VPTPRAAARA AVDELQGAPL LPFPAN          126
(B)   VPTPRAAARA AVDELQGAPL LPFPAN          126
(C)   VDTPRAAARA AVDELQGARL LPFPSN          146
(D)   VPTPRAAARA AVDELQGAKL VPFPAN          140
(E)   VATPRAAARA AVDELQGAPL LPFPTN          126
```

Figure 4.8: Amino acid sequence of (A) the *S. castaneoglobisporus* MELC1 chaperone as determined in this study, compared with (B) the sequence given by Ikeda *et al.* (1996) and also the sequences of (C) *S. antibioticus* (Bernan *et al.*, 1985), (D) *S. lincolnensis* (GenBank accession number X95703) and (E) *S. galbus* (GenBank accession number X95705). Amino acids at which sequences A and B differ are indicated by black shading. Deletions needed for alignment are indicated by dashes.

4.3 *Construction of expression vectors*

Three binary expression vectors were constructed to test expression of the novel pigmentation pathways in plants (Figure 3.1). To show that all the genes were present in their appropriate binary expression vectors, and each with an upstream CaMV35S promotor and a downstream NOS terminator, each gene was amplified by PCR from its vector together with either its promotor or terminator. Figure 4.9 shows the results of these PCRs, confirming that all the promotor-gene and gene-terminator regions could be amplified, and that all are of the expected lengths.

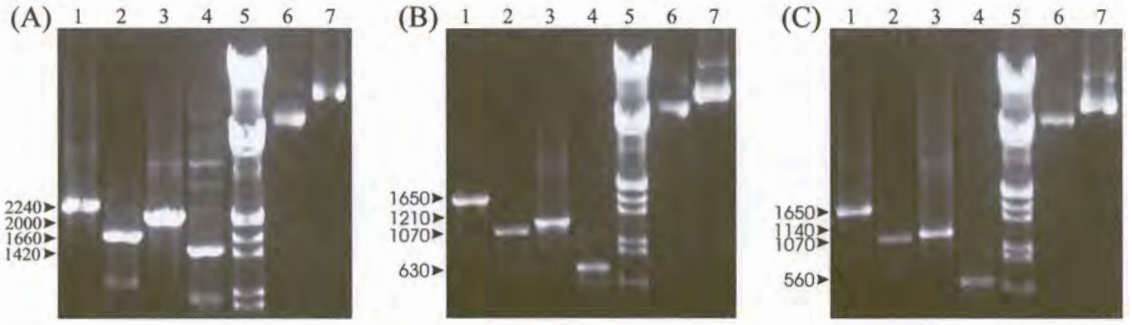

Figure 4.9: PCR products confirming vector construction.

(A) pC-ido-tnaK: PCR-amplified promotor-*tnaK* region of ~2240 bp (lane 1), *tnaK*-terminator region of ~1660 bp (lane 2), promotor-*ido* region of ~2000 bp (lane 3) and *ido*-terminator region of ~1420 bp (lane 4), next to DNA molecular weight marker III (lane 5), uncut pCAMBIA2300 (lane 6) and the larger pC-ido-tnaK (lane 7).

(B) pC-tyrC-melC1: PCR-amplified promotor-*tyrC* region of ~1650 bp (lane 1), *tyrC*-terminator region of ~1070 bp (lane 2), promotor-*melC1* region of ~1210 bp (lane 3) and *melC1*-terminator region of ~630 bp (lane 4), next to DNA molecular weight marker III (lane 5), uncut pCAMBIA2300 (lane 6) and the larger pC-tyrC-melC1 (lane 7).

(C) pC-tyrC-melC1Δ: PCR-amplified promotor-*tyrC* region of ~1650 bp (lane 1), *tyrC*-terminator region of ~1070 bp (lane 2), promotor-*melC1* region of ~1140 bp (lane 3) and *melC1*-terminator region of ~560 bp (lane 4), next to DNA molecular weight marker III (lane 5), uncut pCAMBIA2300 (lane 6) and the larger pC-tyrC-melC1Δ (lane 7).

4.4 *Transformation of tobacco*

Phenotypes of transformed plants. Tobacco leaf disks were transformed with the three vectors pC-ido-tnaK, pC-tyrC-melC1 or pC-tyrC-melC1Δ, and plantlets regenerated. All the transformed shoots that grew from the leaf disks were the normal, pale green colour, and no blue or black colouring could be seen (Figure 4.10 A). The regenerated plants also had normal phenotypes (Figure 4.10 B) – except for one unusual difference: the stigmas and styles of about a quarter of the pC-ido-tnaK-transformed plants were blue. This was, however, only observed during the first five months after transformation. All flowers produced after the first five months had normal, green stigmas and styles.

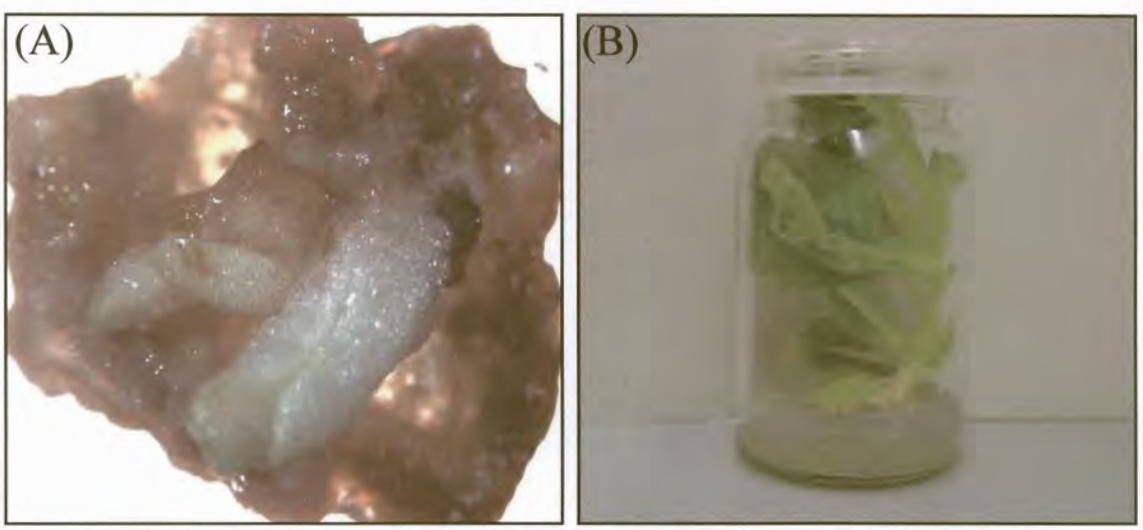

Figure 4.10: Phenotypes of transformed plants. (A) Kanamycin-resistant shoots growing from leaf disks. The shoots are a normal, pale green colour without any blue or black colouring. (B) Regenerated plant, also a normal, green colour.

Confirmation of transformation. To confirm that the plants were indeed transformed, genomic DNA was extracted from young leaves, and the presence of the transgenes detected by PCR amplification. Both *ido* and *tnaK* were easily amplified from all thirty plants transformed with pC-ido-tnaK, confirming that the plants were indeed stably transformed (Figures 4.11 A and B). The *melC* genes, on the other hand, could not be amplified reliably from any of the plants transformed with pC-tyrC-melC1 or pC-tyrC-melC1Δ, even using the special cycling parameters given in chapter 3.5. Instead, the presence of the CaMV35S promoter was used as indicator of transformation. Using the PCR parameters given in chapter 3.5, the CaMV35S promoter was detected in all plants transformed with pC-tyrC-melC1 or pC-tyrC-melC1Δ (Figure 4.11 C).

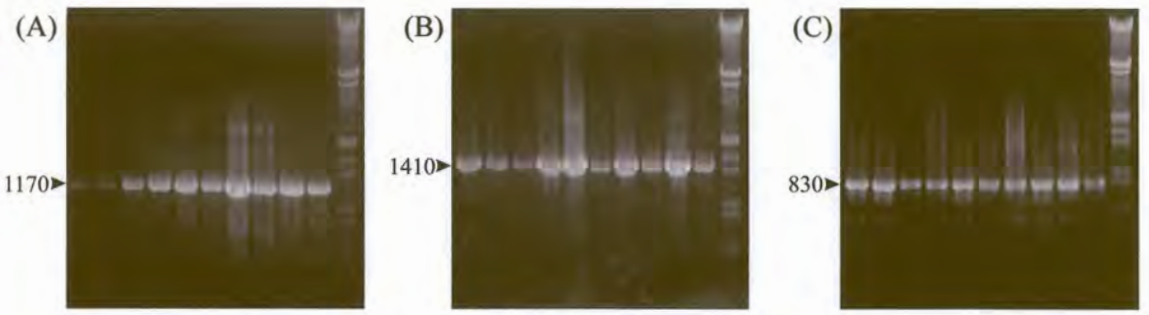

Figure 4.11: Examples of T-DNA regions amplified by PCR from tobacco plants to confirm transformation and transgene integration. In each photo the rightmost lane contains DNA molecular weight marker III. (A and B) *ido* and *tnaK* transgenes amplified from pC-ido-tnaK-transformed plants, showing the expected sizes of approximately 1170 bp and 1410 bp, respectively. (C) CaMV35S promoter amplified from pC-tyrC-melC1-transformed plants, showing the expected size of approximately 830 bp.

4.5 Testing of transgene expression

Dot-blots of bacterial RNA. The dot-blots of bacterial RNA showed that target RNA can indeed be detected (Figure 4.12). The DNase-treated RNA from pSLH8-carrying *E. coli* cultures gave strong signals thoughout the dilution series. The two negative controls (RNA from pGEM®-3Zf+ cultures and RNase-treated extracts from pSLH cultures) gave no signal. In all cases the pC-ido-tnaK dilution series could routinely be detected in amount as little as 50 pg, which corresponds to approximately 5 pg of homologous DNA.

Testing of transgene expression in plants. The dot-blots of plant RNA could not detect any expression of the transgenes. The dilution series of pC-ido-tnaK could, as before, easily be detected down to 50 pg using either the *ido*, *tnaK* or *nptII* probes. The detection of pC-tyrC-melC1Δ was just as sensitive when using the *tyrC*, *melC1* or *nptII* probes. The blots of plant RNA, on the other hand, never showed any signal, irrespective of which probe was used. Moreover, when the plants were placed back onto medium containing kanamycin, the previously kanamycin-resistant plants withered and died. Figure 4.13 shows examples of dot-blots that were performed with *ido* and *nptII* probes.

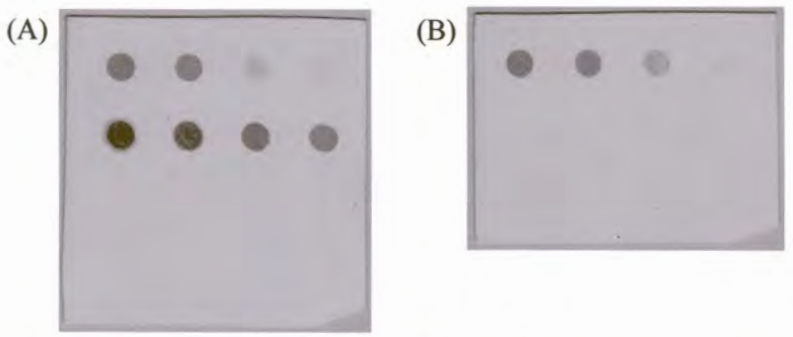

Figure 4.12: Dot-blots of bacterial RNA. Dilution series of pC-ido-tnaK are in the top row of both membranes. Dilution series of DNase-treated RNA from pSLH8- and pGEM®-3Zf+-carrying *E. coli* cultures are in the second and third rows, respectively, of membrane A. A dilution series of RNase- and DNase-treated extracts from pSLH8 cultures is in the second row of membrane B. The DNase-treated RNA from pSLH8-carrying cultures is strongly detected, but the other bacterial extracts give no signal at all.

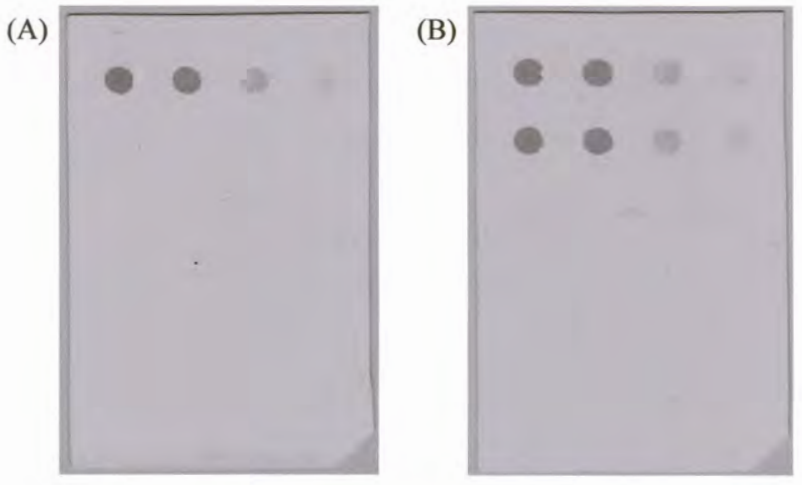

Figure 4.13: Dot-blots of plant RNA. Dilution series of pC-ido-tnaK are in the top row of both membranes. Dilution series of pC-tyrC-melC1Δ are in the second row of both membranes. All the other unseen dots on the membranes are of plant RNA. Membrane A was probed with a DIG-labeled *ido* probe. As little as 50 pg of pC-ido-tnaK could easily be detected, but none of the plant RNA could be detected. Membrane B was probed with a DIG-labeled *nptII* probe. Both the pC-ido-tnaK and pC-tyrC-melC1Δ dilution series were detected but, again, none of the plant RNA could be detected.

Chapter 5

Discussion

5.1 The novelty of using indigo and melanin biosynthetic pathways in plants

The aim of this study was to test the feasibility of using novel pigmentation pathways to create blue of black flowers. Alas, the idea of using novel pigmentation pathways is itself not so novel anymore. During the course of this study other researchers proposed very similar ideas. Gillam and Guengerich (2001), who discovered that certain human cytochrome P450s can catalyze the formation of indigo, quickly recognized the potential of their discovery for the creation of blue flowers or blue cotton. However, they did not elaborate the idea by suggesting, for instance, other enzymes with which the P450s can be combined to form an indigo biosynthetic pathway.

McBride and Stalker (2001) went much further in their proposal. Not only did they propose the use of an indigo biosynthetic pathway in plants to blue flowers and cotton, but also the use of a melanin biosynthetic pathway to create black cotton. Even though they did not give reasons for their choices, the pathways that they proposed are remarkably similar to those tested in this study: they proposed the use of a melanin biosynthetic pathway based on the *S. antibioticus melC* operon, and an indigo biosynthetic pathway that uses the *Rhodococcus* indole dioxygenase and a tryptophanase. In their proposal they also did some preliminary testing of the expression of the *S. antibioticus* genes in tobacco, as mentioned in chapter 2.

This study had to go at least one step further than these proposals. Firstly, the pathways used in this study were not just proposed out of the blue. All the various possible pigmentation pathways were discussed and evaluated so that the most promising ones could be chosen. Secondly, the study did not stop after merely suggesting new pathways. Instead, an attempt was made to actually express these pathways in a model plant, tobacco.

5.2 Sequencing of melC *genes and transformation of tobacco*

The *Rhodococcus* indole dioxygenase, the *E. aerogenes* tryptophanase and the *S. castaneoglobisporus melC* genes were placed in plant expression vectors and successfully introduced into tobacco. In the course of constructing these vectors it was found that all the PCR products of the *melC* genes had sequences slightly different from the published sequences. Since the same discrepancies occurred in all PCR products, it was unlikely to be

due to PCR-induced mutations. To determine whether the sequences of the *melC* genes in the pUCM-2 plasmid are actually the same as the published sequences, the genes were sequenced without intervening PCR steps. The resulting sequences were indeed the same as those of the PCR products, slightly different from the published sequences. The *E. coli* cultures from which the pUCM-2 plasmid was isolated turned black during overnight growth in tyrosine- and copper-enriched medium, indicating that the enzymes encoded by the pUCM-2 *melC* genes were functional (Figure 4.4). Furthermore, the amino acid sequences of the *S. castaneoglobisporus* tyrosinase and chaperone as determined in this study show closer similarity to the consensus sequences of *Streptomyces* tyrosinases and chaperones than the sequences given by Ikeda *et al.* (1996) do (Figures 4.7 and 4.8). It is therefore contended that the new sequences are, in fact, the correct ones.

Genes with such high GC-content as the *melC* genes are notoriously difficult to sequence or amplify by PCR – Han *et al.* (1994) even admitted that their attempts to amplify the *S. antibioticus melC* genes by PCR were unsuccessful and that they had to clone the genes by other means. In this study, the *melC* genes could easily be amplified from a plasmid template by simply increasing the denaturing and annealing temperatures. Because of the high denaturing temperature, extra *Taq* polymerase had to be added after 20 cycles. However, even these measures were not enough to prevent amplification from being unreliable and unspecific when using plant genomic DNA as template. The presence of the *melC* transgenes in kanamycin-resistant, transformed plants therefore had to be confirmed indirectly by amplification of the CaMV35S promoter.

Fortunately, the pCAMBIA2300 vector contains the kanamycin resistance gene next to the T-DNA left border (Figure 3.1), and since T-DNA integration starts at the right border, other parts of the T-DNA region are integrated into the plant genome before the kanamycin resistance gene (Hellens *et al.*, 2000). Any plant with kanamycin resistance, and especially if the CaMV35S promoter could be amplified from it, is almost certain to contain the entire T-DNA region. Amplification of the promoter region is therefore a reliable test of whether the plant is transformed with the *melC* genes.

5.3 Characterization of maize Bx1 *alleles*

The *Bx1* gene copies of the PNR473 cultivar were only characterized long after the *ido*, *tnaK* and *melC* genes have already been used to transform tobacco, so the results of using *Bx1*

could not included in this study. The protocol that was used during the RT-PCR contains only one noteworthy deviation from general protocols such as found in Sambrook *et al.* (1989). During the first cycle of the PCR, only the forward primer, which binds to the first strand cDNA that was synthesized during reverse transcription, is added. This was found to greatly reduce the amount of non-specific binding and background amplification products. Since the reverse primer does not yet have a complementary strand to bind to during the first temperature cycle, inclusion of it during the first cycle can only decrease the specificity of the PCR.

The RT-PCR product contained two distinct sequences. Since the PNR473 cultivar is not an inbred line, but rather a hybrid (Stephanus de Meillon, personal communication), it allows for the possibility that the two sequences represent two alleles at the same locus. Both of these PNR473 sequences contain three insertions relative to the CG00237 *Bx1* gene (Figure 4.2). However, the changes occur within the first exon, which has no counterpart in the bacterial TSA gene and which encodes the chloroplast targeting signal (Radwanski *et al.*, 1995; Zhao and Last, 1995). The exact amino acid sequence of such a targeting peptide is not important. As long as it is rich in serine, threonine and small hydrophobic amino acids, and poor in glutamic acid and aspartic acid, it should still function as a chloroplast targeting signal (Bruce, 2000). The amino acid sequence between the three insertions is Pro-Ala-Pro-Pro-Ala-Pro-Val-Pro-Pro-Lys-Gln-Ala-Ala-Ala-Pro-Ala-Glu (Figure 4.3), which fulfills these criteria. It is interesting to note that such proline-rich regions are also present in the chloroplast targeting signals of several other proteins (Bruce, 2000).

5.4 Silencing of the transgenes

Even though the tobacco plants were successfully transformed (Figure 4.11), their phenotypes showed disappointingly little change of colour (Figure 4.10). The only unusual aspect of their appearance was the blue stigmas and styles that were observed in some flowers. This absence of blue or black colouring in most plants could either have been due to some of the genes not being expressed strongly enough, or to some of the gene products not functioning as intended. The first possibility seems to have been the case in this study – or more precisely, it was the case five months after transformation – since no expression could be detected. This inability to detect expression is unlikely to have been caused by experimental failure, since the DNA controls and the bacterial *ido* mRNA could easily be detected. The amount of total RNA tested per plant, 10 µg, typically includes between 50 and 300 ng mRNA (SuperScript™

Lambda System instruction manual, Life Technologies). Even a rare transcript that makes up only 0.01% of the mRNA fraction should, therefore, have been present in amounts greater than 5 pg. This amount of complimentary DNA could be detected (Figures 4.12 and 13). The plant RNA extraction buffer included HMW-PEG, which removes polysaccharides and polyphenols that might have hindered the binding of the RNA to the membrane or to the probes (Gehrig *et al.*, 2000). Furthermore, the plants that were previously kanamycin-resistant withered when placed back on a medium supplemented with kanamycin, indicating that the *nptII* transgene was also silenced.

If it is assumed that the expression vectors were constructed correctly, then the absence of strong expression is surely due to gene silencing. However, the fact that not even one single plant showed expression is quite perplexing. Even though the expression vectors were not optimised to minimize silencing, it is highly unusual for such strong silencing to occur in all plants. Many studies have reported expression when the transgenes were present in multiple, perhaps even rearranged copies, and contained repeat sequences (Neuhuber *et al.*, 1994; Kohli *et al.*, 1999; Iyer *et al.*, 2000; Domínguez *et al.*, 2002). Silencing was found to be random – occurring in some plants but not all.

The expression vectors used in this study contained three copies of the CaMV35S promoter, which are substrates for repeat-induced transcriptional gene silencing; and two of the transgenes had the same terminator, which may have induced homology-dependent post-transcriptional gene silencing (de Wilde *et al.*, 2000). The vectors also contained no matrix-associated regions, which would have minimized position effects (Allen *et al.*, 2000). This is very similar to the vectors used by Kohli *et al.* (1999) and Domínguez *et al.* (2002) – and theirs even contained the promoters or terminators as inverted repeats. It was, therefore, expected that some gene silencing would occur, but not that all transgenes would be silenced in all plants. Since the aim was simply to test the activity of the novel pathways in tobacco – not to create new, commercial varieties that express the genes stably and strongly over many years – some silencing would have been acceptable.

Silencing can also be influenced by environmental conditions. Abiotic stress such as the growing of plants under high light intensity or in closed culture vessels prior to transfer to a greenhouse has been found to increase the percentage of plants affected by silencing (de Wilde *et al.*, 2000; Palauqui *et al.*, 1996). In this study the plants were grown under continual light conditions and in closed culture vessels, which might therefore have contributed to the excessive silencing.

It is not known whether the pigment genes were expressed during the first few months immediately after transformation, as the dot-blots were only performed five months after transformation. The regenerated shoots were the normal, pale green colour, so if the genes were expressed it would imply that the pigmentation pathways did not work as intended. However, the shoots only appeared after a few weeks, and this is sufficient time for silencing effects to develop (Iyer *et al.*, 2000).

The kanamycin resistance gene was certainly expressed during the first two months, since the plants grew on kanamycin-supplemented medium (Figure 4.10). But this does not necessarily mean that the other genes were also expressed. The pigment genes may have been more susceptible to silencing than the kanamycin resistance gene. For instance, each pigment gene had a NOS terminator, which could have led to post-transcriptional co-suppression. Silencing and the accompanying methylation do not necessarily spread to other transgenes in the vicinity (Kohli *et al.*, 1999). A silenced transgene can even be found in between two flanking, functional transgenes (Iyer *et al.*, 2000). It is therefore conceivable that the pigment genes could have been silenced before the kanamycin resistance gene.

The study failed to tell anything about the feasibility of the use of the *melC* genes to create a melanin biosynthetic pathway in flowers, since all the *melC*-transformed plants were phenotypically normal and no expression of the genes could be detected. The study will have to be repeated – but in the next attempt an expression vector that is less likely to induce gene silencing should be used, and the plants should be tested for expression as early after transformation as possible. Any abiotic stress that may contribute to silencing should also be carefully avoided.

5.5 The strangely blue stigmas and styles

Some of the pC-ido-tnaK-transformed plants showed an unusual phenotype that deserves further discussion. For the first five months after transformation, the stigmas and styles of the flowers produced by these plants were blue. Wildtype tobacco stigmas and styles are green, and tobacco does not even produce blue, delphinidin-derived pigments (Shimada *et al.*, 1999). Since the blue stigmas and styles only occurred in the pC-ido-tnaK-transformed plants, it must somehow be the result of the *ido* and *tnaK* transgenes. Unfortunately, the plants stopped producing flowers with blue stigmas and styles after five months, before they could be studied

further. Therefore nothing definitive can be said about the reasons for the blue stigmas. Nevertheless, one can still speculate.

One possibility is that the genes were expressed in the stigmas and styles of certain plants while silenced in all other organs. However, this is unlikely. There have never before been reports of transgenes that are expressed only in stigmas and styles even though they are under control of constitutive promoters. Stigmas and styles are also not known to be 'silencing-priviliged' organs, meaning that silencing does not occur in those organs. At least, homology-dependent post-transcriptional silencing has been found to occur in stigmas (Takasaki *et al.*, 2001). Strong CaMV35S promoter activity is found in all parts of inflorescences, including the stigmas and styles (Holtorf *et al.*, 1995), which rules out the possibility that dosage-dependent post-transcriptional silencing did not occur due to a lower level of transcription.

If silencing does not occur systemically throughout the whole plant, it usually occurs in a random manner in different parts of the plant. This can be seen in the variegated patterns in flowers with antisense *chsA* suppression (Jorgensen *et al.*, 1996), or when silencing is limited to intraveinal spots in leaves due to a diffused silencing signal (Fagard and Vaucheret, 2000). Sometimes there is a gradient of silencing throughout the plant, so that silencing is more pronounced in either the younger, topmost leaves (Palauqui *et al.*, 1996) or the older, bottom leaves (van Houdt *et al.*, 1997). These random patterns of silencing lead to the possibility that the *ido* or *tnaK* transgenes were randomly expressed in several different parts of some plants, but that conditions were favourable for the production of indigo in only the stigmas and styles. If, by chance, the genes were expressed in the stigma and style of a certain plant, these parts turned blue.

Following this trail of thought, one cannot simply presume that both enzymes were active in the stigma, style and other parts. Although both genes may have been expressed, probably only one of the two enzymes was functional. Otherwise not only the stigma and style, but also all those other parts of the plants should have turned blue. It is hard to imagine that the stigma and style are any different from most other tissues with regard to their processing of nascent peptides, their amount of free tryptophan, their metabolism of indigo, or anything else that could have cause the pathway to be functional in the stigma and style but not in other tissues, if both enzymes were active and functional in these tissues. So, if both enzymes were functional in the stigma and style, both should have been functional in all other tissues in which their genes were expressed.

Even if only one of the two enzymes was functional, either of them could conceivably have combined with endogenous enzymes to produce indigo in the stigma and style. A functional tryptophanase, for instance, could have combined with an endogenous cyt-P450 enzyme to form an indigo biosynthetic pathway. Most plants, including tobacco, express many more different cyt-P450 enzymes than humans do (Gillam, 1999). So, if several human cyt-P450 could catalyze the oxygenation of indole to form indigo, it is likely that there are tobacco cyt-P450s that can do the same. It is also more common for the expression of an enzyme, such as a cyt-P450, to be limited to the same tissues in different plants, than it is for silencing to be limited in such a way.

If only the indole dioxygenase was functional and not the tryptophanase, there had to have been an endogenous source of free indole. Frey *et al.* (2000) mention three classes of secondary metabolites that depend on the tryptophan biosynthetic pathway and whose biosynthesis may therefore involve BX1-homologous enzymes, namely indole glucosinolates, indole phytoalexins and terpenoid indole alkaloids. However, contrary to what they proposed, only the biosynthesis of indole phytoalexins such as camalexin is known to involve the formation of free indole (Zook, 1998). The known biosynthetic pathways of indole glucosinolates and indole alkaloids actually start from tryptophan and do not involve BX1-homologous enzymes or indole (Chen and Andreasson, 2001; Facchini, 2001). And even if other undiscovered biosynthetic pathways do exist, they could not be the source of free indole, since tobacco does not even produce any of these three classes of secondary metabolites (Rodman, 1991; Pedras *et al.*, 2000; Facchini, 2001). Tobacco also does not produce free indole as part of its floral scent, since neither indole nor any of its known derivatives could be detected in the headspace of *Nicotiana tabacum* flowers (Loughrin *et al.*, 1990; Spiteller and Steglich, 2001).

The only metabolite that is known to be produced in tobacco and whose biosynthetic pathway does involve indole, is the important auxin hormone, indole-3-acetic acid (IAA). IAA can be produced by either a tryptophan-dependent pathway or a tryptophan-independent pathway. The tryptophan-dependent pathway takes tryptophan as its starting point and does not involve indole. The tryptophan-independent pathway, on the other hand, branches from the tryptophan biosynthetic pathway at indole-3-glycerol phosphate, and produces free indole as an intermediate (Ouyang *et al.*, 2000). This tryptophan-independent pathway is the dominant source of IAA in tobacco (Sitbon *et al.*, 2000).

IAA is known to be present in the stigma and style of the flowers of certain citrus cultivars (Kojima, 1996), so is likely to be present in the stigma and style of tobacco too. However, the

concentration of IAA in the stigma and style of the citrus cultivars is about threefold less than in the stamens and petals. IAA is actually present in most parts of the plant, with rapidly growing tissues having the greatest concentration (Ljung *et al.*, 2001). Therefore, even though the tryptophan-independent IAA pathway is a potential source of free indole in the stigmas and styles, it would also have been a source of indole in other tissues. If the indole dioxygenase was active in other tissues than the stigma and style, these tissues should also have turned blue.

In the end, however, all this speculation about the possible causes of the blue stigmas and styles is just that – mere speculation. Before anything more definite can be said about the use of the indigo biosynthetic pathway in plants, the study will have to be repeated. The same precautions against excessive silencing should be taken as mentioned above for the melanin biosynthetic pathway. An additional experiment that can be performed is to regenerate and grow *ido*-transformed plants in an indole-supplemented medium. This way the activity of the indole dioxygenase can be evaluated as soon as the regenerated shoots start to appear. The *tmaK* gene does not need to be included in this experiment, which would reduce the chances of silencing or other complications.

At least the blue stigmas and styles showed that some part of the indigo biosynthetic pathway was functional, even if it is not known how or why. So, in the true age-old fashion of fairytale quests, the first leg of the journey did not yet lead to success, but it did give a glimmer of hope that the quest will one day succeed…

hapter 6

Summary / Opsomming

Summary

Blue or black flowering varieties are still missing from important ornamental flowers such as rose, tulip, chrysanthemum and carnation. All the past, failed attempts to create blue or black flowers have focused on the manipulation of the flavonoid pigmentation pathway. The aim of this study was, therefore, to investigate the use of alternative, novel pigmentation pathways that may lead to the expression of novel blue or black pigments not normally found in flowers. Two promising pathways were chosen: an indigo biosynthetic pathway that uses a *Rhodococcus* indole dioxygenase (IDO) together with either an *Enterobacter aerogenes* tryptophanase (TNAK) or a *Zea mays* indole synthase (BX1), and a melanin biosynthetic pathway that uses the tyrosinase (TYRC) and chaperone (MELC1) of the *Streptomyces castaneoglobisporus melC* operon.

The sequences of the *melC* genes that were observed in this study differed slightly from the published sequences. It is proposed that the observed sequences are correct, and therefore the new sequences were submitted to GenBank as accession numbers AY254101 and AY254102. The maize *Bx1* gene was isolated from the PNR473 cultivar using an RT-PCR. Two different *Bx1* copies were found, and the sequences were also deposited in GenBank as accession numbers AY254103 and AY254104. Although the *Bx1* genes have been characterized, they have not been used further in this study.

Tobacco was transformed with either the *ido* and *tnaK* genes or the *tyrC* and *melC1* genes. All the regenerated plants had normal phenotypes – except for one unusual difference: for the first five months after transformation, about a quarter of the plants transformed with *ido* and *tnaK* produced flowers with blue stigmas and styles. However, after five months, all new flowers that were produced had the normal, green stigmas and styles.

Expression of transgenes was tested through Northern dot-blot analysis of total RNA extracts, approximately six months after transformation. Although complimentary DNA and bacterially-expressed *ido* mRNA could be detected at low concentrations, no transgene expression could be detected in any of the plants. It is proposed that the transgenes were silenced within the first few months after transformation.

The study failed to tell anything about the feasibility of the use of the melanin biosynthetic pathway to create black flowers, since no transgene expression could be detected and the transformed plants showed no change in appearance. At least the blue stigmas and style

showed that some part of the indigo biosynthetic pathway was functional. The study will have to be repeated, but with additional precautions to avoid gene silencing.

Opsomming

Niemand kon nog ooit 'n ware blou of swart roos, tulp, krisant of angelier kweek nie. Al die vorige mislukte pogings om blou of swart blomme te kweek het gefokus op manipulasie van die biosintese van flavonoïedpigmente. Die doel van hierdie studie was dus om die gebruik van alternatiewe, nuwe pigmentasie-biosinteseweë te ondersoek. Sulke biosinteseweë kan lei tot die uitdrukking van pigmente wat nie normaalweg in blomme gevind word nie. Twee belowende biosinteseweë was gekies: 'n indigo-biosinteseweg wat gebruik maak van 'n *Rhodococcus* indool-dioksigenase (IDO) in kombinasie met óf 'n *Enterobacter aerogenes* triptofanase (TNAK) óf 'n *Zea mays* indool sintase (BX1), en 'n melanien-biosinteseweg wat gebruik maak van die tirosinase (TYRC) en chaperone (MELC1) van die *Streptomyces castaneoglobisporus melC* operon.

Die nukleotiedvolgordes van die *melC*-gene wat in hierdie studie verkry is, verskil van die gepubliseerde volgordes. Dit word voorgestel dat die waargenome volgordes korrek is, en die nuwe volgordes is voorgelê aan GenBank met aanwinsnommers AY254101 en AY254102. Die *Zea mays Bx1*-geen was geïsoleer vanuit die PNR473 kultivar. Twee verskillende *Bx1*-kopieë is gevind, en die volgordes is by GenBank ingedien met aanwinsnommers AY254103 en AY254104. Alhoewel die *Bx1*-gene gekarakteriseer is, was hulle nie verder in hierdie studie gebruik nie.

Tabak is onderskeidelik met óf die *ido*- en *tnaK*-gene, óf die *tyrC*- en *melC1*-gene getransformeer. Al die geregenereerde plante het normale fenotipes vertoon – behalwe vir een ongewone verskil: vir die eerste vyf maande na transformasie het ongeveer 'n kwart van die plante wat met *ido* en *tnaK* getransformeer was blomme met blou stigmas en style geproduseer. Na die eerste vyf maande het al die plante egter slegs blomme met normale groen stigmas en style geproduseer.

Uitdrukking van die transgene is getoets deur middel van Noordelike klad-analise van totale RNA-ekstrakte, ongeveer ses maande na transformasie. Alhoewel komplimentêre DNA en bakteriële *ido* mRNA teen lae konsentrasies bespeur kon word, kon geen ekspressie van die

transgene in die plante bespeur word nie. Daar word voorgestel dat ekspressie van die transgene onderdruk was binne die eerste paar maande na transformasie.

Die studie kon dus geen aanduiding gee van die bruikbaarheid van die melanien-biosinteseweg om swart blomme te kweek nie, aangesien geen transgeen-ekspressie bespeur kon word nie en die getransformeerde plante geen verandering in voorkoms getoon het nie. Ten minste het die blou stigmas en style getoon dat sekere dele van die indigo-biosinteseweg wel funksioneel was. Die studie sal herhaal moet word, maar met beter voorsorgmaatreëls teen geen-onderdrukking.