August 2012
Ph.D. Dissertation

# Production of indigo and indirubin using recombinant *Escherichia coli* harboring a novel flavin-containing monooxygenase gene

Graduate School of Chosun University
Department of Environmental Engineering
Gui Hwan Han

# Production of indigo and indirubin using recombinant *Escherichia coli* harboring a novel flavin-containing monooxygenase gene

***Methylophaga aminisulfidivorans* MP$^{T}$ 유래 flavin-containing monooxygenase 유전자를 내포시킨 재조합 대장균을 이용한 인디고 및 인디루빈의 생산**

August 24 2012

Graduate School of Chosun University
Department of Environmental Engineering
Gui Hwan Han

# Production of indigo and indirubin using recombinant *Escherichia coli* harboring a novel flavin-containing monooxygenase gene

Supervisor : Prof. Si Wouk Kim

A dissertation submitted to the Department of Environmental Engineering and the Graduate School of Chosun University in fulfillment of the requirements for the degree of Doctor of Philosophy

April 2012

Graduate School of Chosun University
Department of Environmental Engineering
Gui Hwan Han

This certifies that the degree of Doctor of Philosophy of Gui Hwan Han is approved by

Heon Man Lim
Chair of committee

Il Soo Bang
Committee member

Hyung Joon Cha
Committee member

Jintae Lee
Committee member

Si Wouk Kim
Committee member

June 2012

Graduate School of Chosun University

# Contents

# List of tables

# List of figures

# ABSTRACT

Indigo and indirubin are considered to be the oldest natural textile dyes and drugs. They are traditionally extracted from several plant species. Recent studies have shown that indigo and indirubin can be produced via several biological processes through the use of recombinant microorganisms expressing mono- or di-oxygenase enzymes. This study aimed to optimize the production of indigo and indirubin through the use of a bacterial flavin-containing monooxygenase (FMO) gene.

A bacterial FMO gene was cloned from *Methylophaga aminisulfidivorans* $MP^T$, and a plasmid pBlue 2.0 was constructed to express the bacterial *fmo* gene in *E. coli*. The recombinant *E. coli* produced 160 mg/L of indigo from 2 g/L tryptophan medium. Transcription of the *fmo* gene was initiated from a thymine residue (+1) 279 bp (TSS1) upstream from the translation initiation codon of the *fmo* gene. On the basis of the sequenced data from the *fmo* transcripts, we identified a -35 region (5'-CTGGAA-3') and -10 region (5'-TATCCT-3'). In addition, we also identified the TG motif situated -14, -15 bp upstream from the transcription start site. To increase the production of indigo, upstream length of *fmo* gene was optimized and the response surface methodology was used. The pBlue 1.7 plasmid was prepared by deleting the sequence upstream of pBlue 2.0. The recombinant *E. coli* harboring the pBlue 1.7 plasmid produced 662 mg/L of indigo in tryptophan medium after 24 h of flask culture. Production of indigo was optimized using response surface methodology with a $2^n$ central composite design. The optimal contents of the media that produced the maximum amount of indigo were determined as follows: 2.4 g/L tryptophan, 4.5 g/L

yeast extract and 11.4 g/L sodium chloride. In addition, the optimum culture temperature and pH were determined to be 30℃ and pH 7.0, respectively. Under the optimized conditions mentioned above, the recombinant *E. coli* harboring the pBlue 1.7 plasmid produced 920 mg/L in optimum tryptophan medium after 24 h of cultivation. The combination of truncated insert sizes and culture optimization resulted in a 575% increase in the production of indigo. Indigo productions were performed by 3,000 L repeated batch fermentation and 5 L continuous fermentation. Repeated batch fermentation in a 3,000 L fermenter produced 911 ± 22 mg/L of indigo with 2 g/L tryptophan as a substrate (yield, 46.9%) under the following culture conditions; culture temperature 30℃; pH 7.0; agitation speed 200 rpm; and aeration 3 vvm. We found that sufficient dissolved oxygen (aeration rate and agitation speed) was critical for indigo production. For continuous fermentation in a 5 L fermenter, the volumetric productivity was found to be 11.3 mg/L/h up to 110 h (final accumulated indigo, 23 g) with a constant dilution rate of 0.084/h (constant feeding rate of 0.167 L/h with medium containing 3 g/L tryptophan). Recombinant *E. coli* pBlue 1.7 can withstand the toxicity of high concentrations of accumulated indigo during batch fermentation. For continuous fermentation, the recombinant cells exhibited high plasmid stability up to 110 h, after which the plasmid was lost.

In a previous study, the recombinant *E. coli* pBlue 1.7 produced indigo (920 mg/L) and indirubin (≤5 mg/L) in a 5 L fermenter containing tryptophan medium. In this study, we found that indirubin production was greatly increased when 3 mM cysteine was added to the tryptophan medium, although cysteine inhibited the growth of the recombinant *E. coli* harboring the *fmo* gene. However, the addition of cysteine did

not influence the expression level and activity of FMO in the cell. Cysteine inhibited the synthesis of 3-hydroxyindole from indole but enhanced the synthesis of 2-hydroxyindole, which might increase indirubin production. The optimal culture conditions for indirubin production in tryptophan medium were determined from analysis of the results of the response surface methodology. Under optimal conditions, the recombinant *E. coli* cells could produce 223.6 mg/L of indirubin from 2 g/L of tryptophan. Thus, indigo and indirubin production systems have potential for use in industrial applications for overcoming the environmental problems associated with synthetic indigo production and for meeting the demand for natural indigo and indirubin in the dye and drug industries.

*M. aminisulfidivorans* $MP^T$ is a restricted facultative marine methylotrophic bacterium that grows on methanol, methylated amines, dimethyl sulfide, and dimethyl sulfoxide. Here we present the high-quality draft genome sequence of *M. aminisulfidivorans* $MP^T$, consisting of a chromosome (3,092,085 bp) and a plasmid (16,875 bp). The *M. aminisulfidivorans* $MP^T$ chromosome was found to harbor 2,984 protein-coding genes, 3 rRNA operons, and 31 tRNAs, whereas the plasmid had 190 coding genes and 9 tRNAs. A gene encoding a novel bacterial FMO, which catalyzes nitrogen-containing compounds or indole by the use of oxygen through an NADPH-dependent pathway, was identified. The information obtained in the current study regarding the *M. aminisulfidivorans* $MP^T$ draft genome sequence should facilitate future studies on the metabolic diversity of the genus *Methylophaga*.

# Chapter I

# General Introduction

## 1. Dyes and pigments

Dyes and pigments are substances that impart color to a material. In ancient times, humans used natural dyes and pigments for cave paintings and for coloring their skin and clothes. The first evidence of the use of colorant materials by men goes as far back as 15,000-9,000 BC, in paintings the Altamira cave in Spain and in some parts of southern France. Archaeologically, these colored dyes and pigments were obtained from stones containing minerals such as iron, manganese, and aluminium, by burning them with plant or animal materials. Dyes may be classified in several ways. They may be classified on the basis of the nature of their chromophores (e.g., azo dyes, acridine dyes, anthroquinone dyes, nitroso dyes, quinone dyes, indophenol dyes, and thiazine dyes), by their nuclear structure (anionic dyes and cationic dyes), by their fiber type (e.g., dyes for nylon, cotton, and wool)or, by their mode of application (acid, azoic, reactive, sulphur, and basic dyes). Furthermore, a very common classification of the dyestuff is based on the source from which it is prepared. Accordingly, dyes can be classified into synthetic dyes and natural dyes.

## 2. Synthetic dyes

Man-made and synthetic dyes are derived from organic or inorganic compounds, especially from benzene and its derivatives. Historically, the development of synthetic dyes started in 1856, when an 18-year-old chemistry student, William Henry Perkin, discovered a brilliant purple solution (mauveine) while attempting to synthesize

quinine. This discovery initiated the great history of synthetic dye production and subsequently, the synthesis of numerous dyestuffs by researchers. Earlier discoveries worth mentioning include the synthesis of magenta, a basic dye (Verguin, 1859); methyl violet, a basic dye (Bardy, 1866); induline, a basic dye (Dale and Caro, 1863); malachite rreen, a basic dye (O'Fischer, 1877); indigo, an indigoid vat dye (Bayer, 1880); crystal violet, a basic dye (Kern, 1883); congo red, an azo dye (Boettiger, 1884); indigo, an indigoid vat dye prepared newer synthesis methods (Heumann, 1890 and Sandmeyer, 1899); alizarin direct blue A, an anthraquinone dye (Herzberg 1913); brilliant take colors from basic dyes (BASF, 1913); and brilliant lake colors from basic dyes (BASF, 1913).

## 3. Natural dyes

Natural dyes are obtained from plants, animals and molluscs. These dyes are often neither safer nor more ecologically sound than synthetic dyes. However, some natural dyes work on cotton, especially when set with mordants such as tannins. Among the better natural dyes for cotton are annato, tyrian purple, cutch, logwood, madder, and indigo; except for indigo, all of there require mordants. Indigo is very popular and largely used for cellulosic fibers such as cotton. Indigo dyes are widely used for dyeing and printing protein and cellulose fibers.

# 4. Indigo and indirubin

## 4-1. Plant indigo

Indigo is considered to be the oldest textile dye and is used for dyeing cotton and wool fabrics. It is traditionally extracted from several plant species (Table 1-1). Several plants that can produce indigo have come to acquire greater significance than others; one such plant is *Indigofera tinctoria*, so named and described by Linnaeus. Indigo was known throughout the ancient world for its ability to color fabrics a deep blue. Egyptian artifacts suggest that indigo was used as early as 16th century BC. and it has been found in Africa, India, Indonesia, and China. *Indigofera tinctoria* originally came from India, where it is mentioned in manuscripts dating from 4th century BC. It was recognized as a valuable blue dye by most early explorers of that region. The Venetian explorer Marco Polo described in detail the Indian indigo industry in detail; by the 11th century, Arab traders had introduced indigo to the Mediterranean region, where it replaced their native blue dye plant, woad (*Isatis tinctoria*).

The plants were grown in large open fields; the seeds were sown at the end of April or the beginning of May. They germinated after a week and in the middle of August, which was when the actual production of indigo really began if conditions were favourable. Too much or too little rain could compromise the entire harvest. Generally, to produced 1 kg of indigo powder, 100 kg of the indigo plant is required. Nineteen thousand tons of indigo were produced from plant sources in 1897. Plant indigo includes several different qualities of indigo, each of which has a different name: Bengal indigo, Java indigo, Chinese indigo, Guatemala indigo, Manilla indigo, African (Senegal) indigo and Egyptian indigo.

Table 1-1. List of plants used available for indigo and indirubin production

| Continents | Plants |
|---|---|
| South and Southeast Asia | *Indigofera tinctoria*<br>*Indigofera* sp.<br>*Polygonum tinctorium* |
| Europe, Africa | *Isatis tinctoria* (woad plant) |
| America | *Indigofera suffruticosa* |

## 4-2. Chemical indigo

Since the late 19th century, plant-derived indigo has been replaced in the textile industry by synthetic indigo. In 1882, Bayer and Drewsen began working on the synthesis of indigo [Baeyer, 1882]. The first practical method was developed by Pfleger in 1901. In this process, *N*-phenylglycine is treated with a molten mixture of sodium hydroxide, potassium hydroxide, and sodamide. This highly sensitive melt produces indoxyl, which is subsequently oxidized by air to form indigo. Variations of this method are still in use today. An alternative and viable route to indigo production was developed by Heumann in 1897. In the synthetic reaction of indigo, naphthalene is first oxidized to phthalic anhydride. Classically, potassium permanganate is used as the oxidant, but other cheaper industrial oxidants can be used. The phthalic anhydride is then converted to ammonium 2-carbamoyl benzoate. The primary amide on this molecule is converted to an amine by using the standard Hoffmann rearrangements of a base and sodium hypochlorite. Subsequent acidification results in anthranilic acid. Anthranilic acid is then converted to *N*-phenylglycine-*o*-carboxylic acid by using chloroacetic acid. Closure for forming the indole ring occurs upon fusion with lye. This results in an intermediate, which is a carboxylic acid derivative of indoxyl. This molecule then undergoes decarboxylation to form indoxyl, which is oxidized by air to form indigo (Fig. 1-1). The BASF chemical company developed a commercially feasible manufacturing process that was in use by 1897. The annual production of synthetic indigo is estimated to be 22,000 tons [Schhrott, 2000]. The current consumption of this dye is enormous because of the popularity of blue jeans, which are dyed with indigo. The consumption of indigo and other vat dyes reaches approximately 33 million kilograms annually [Roessler, 2003].

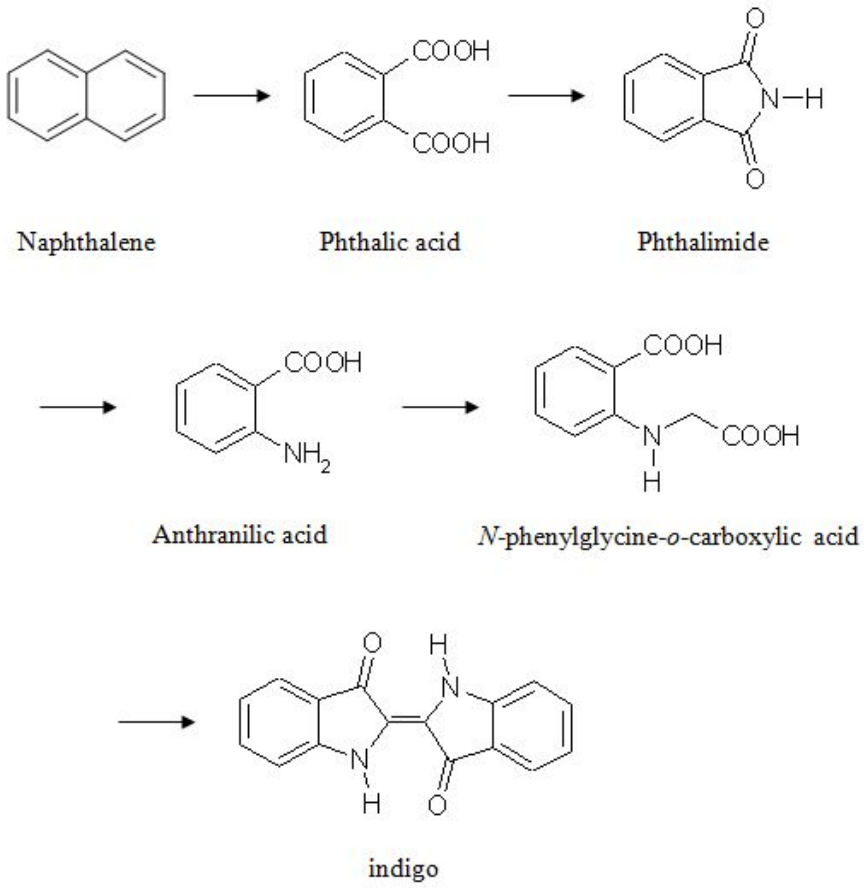



Fig. 1-1. Baeyer-Drewson indigo synthetic scheme. Naphthalene is first oxidized to phthalic anhydride. The phthalic anhydride is then converted to ammonium 2-carbamoyl benzoate. Anthranilic acid is then converted to *N*-phenylglycine-*o*-carboxylic acid using chloroacetic acid. This molecule is then decarboxylatesd to form indoxyl, which is oxidized by air to form indigo.

Various waste products are formed during indigo production and must be handled carefully. In addition to the reactants described above, other reaction side products are also produced along with indigo. This process also requires several chemical manipulations, some involving toxic materials, with the potential to injure humans and the environment. Some of these materials are considered to be hazardous and must be disposed of in accordance with local and federal chemical waste disposal guidelines. These waste chemicals can enter the environment in at least 3 different ways. The first is during the actual manufacture of the molecule, the second is when the dye is applied to the yarn, and the third is when the dye is eluted into the wash water during the initial stone washing or wet processing of the fabric. This last process typically occurs during the production of denim fabric. Nevertheless, the popularity of indigo continues to increase worldwide.

## 4-3. Biological indigo

Despite several hazards to human and the environment, indigo is still very popular dye. Many researchers have therefore been concentrating on new methods of indigo manufacturing methods that are more environmentally friendly. One promising method involves using biocatalysts in the dye reaction process. The indigo dye may be one of the first high-volume chemicals produced via a biological route. Several studies have been recently preformed on the biological production of indigo from indole by using recombinant microorganisms expressing mono- or dioxygenase [Bae, 2006; Doukyu, 2002; Madsen, 1988; Murdock, 1993; Yan, 2007]. Indigo and indirubin are presumed to be produced from 2 main reactions in recombinant microorganisms. In the

first reaction, indole is produced from tryptophan by tryptophanase (TnaA; EC 4.1.99.1), and it is then further oxidized into indole derivatives such as 2-hydroxyindole, 3-hydroxyindole and isatin by catalysis of the flavin-containing monooxygenase (FMO). Indirubin and indigo compounds are synthesized by the random condensation of the aforementioned compounds in the presence of oxygen [Eaton, 1995; Maugard, 2002]. Thus far, the substrates used for the production of indigo by using recombinant bacteria have been glucose and indole. Glucose is inexpensive and readily available, but its metabolic pathway is complicated and the resultant indigo yield is low. The toxic nature of indole also causes a decrease in the production of indigo by microorganisms [Doukyu, 1997]. Therefore, the concentration of indolr in the culture medium must be kept low to avoid this toxic effect. The aforementioned substrates are unsatisfactory for the production of indigo; moreover, the productivity of indigo is so very low that this process cannot be considered as an alternative to chemically produced indigo. Additionally, the technology and production costs involved are currently expensive.

## 4-4. Indirubin

Indirubin has been used as a natural drug and colorant since ancient times [Kohda, 1990; Xia, 1992]. The traditional Chinese prescription Dang Gui Long Hui Wan consists of 11 herbal medicines. Danggui Longhui Wan that contains indirubin has been used for the treatment of numerous chronic diseases, including chronic granulocytic leukemia [Han, 1994; Nam, 2005; Tang, 1992]. The antileukemic activity of the complex mixture was found to be due to indirubin, a minor constituent of the indigo plant. Clinical studies published in 1980 reported that indirubin induced complete

remission in 26% and partial remission in 33% of individuals with leukemia. Indirubin has been reported to inhibit DNA and protein synthesis in various cell lines [Wu, 1980]. In recent years, it has found that indirubin is a potent inhibitor of cyclin-dependent kinases (CDKs) and glycogen synthase kinase-3β (GSK-3β); therefore it may play an important role in the treatment of leukemia and Alzheimer's disease [Hoessel, 1999; Leclerc, 2001]. The mechanisms by which indirubin or its derivatives exert their effects against the cancer process are now quite well known. Current research efforts into understanding cancer are focused mainly on signaling processes or the means by which molecules communicate and are transformed through well-defined pathways leading to cell growth. CDKs are enzymes that catalyze the transfer of phosphate groups from one molecule to another. Kinases are essential for normal processes in the body, but their activity is carefully regulated. They are also essential for tumor growth and progression, but their activity is unregulated. The most successful targeted therapies include inhibitors of the kinase signaling pathways. The cell cycle plays an essential role in cell division, and the cyclin protein appears at various stages in the cycle. CDKs must bind to cyclin in order to function. Indirubin-3-monoxime causes the inhibition of CDK1 in human mammary carcinoma cell cultures. This activity prevents cells from completing the cell cycle, which is necessary for cell division [Marko, 2001].

Indirubin has been extracted from plant cell cultures of *Indigofera tinctoria*, *Isatis tinctoria*, *Polygonum tinctorium*, and *Lonchocarpus cyanescens* [Ensley, 1983; Fitzhugh, 1997; Seldes, 1999]. Recently, several studies on the biological production of indirubin and indigo from recombinant microorganisms expressing mono- or

di-oxygenase have been published [Doukyu, 2002; Ensley, 1983; Madsen, 1988; McClay, 2005; Murdock, 19932]. However, the productivity of indirubin is very low when using chemical and biological processes.

## 5. A novel bacterial flavin-containing monooxygenase gene from *Methylophaga aminisulfidivorans* MP$^T$

Methylotrophic bacteria utilize $C_1$-compounds such as methane, methanol, and methylated amines as the sole carbon and energy sources [Anthony, 1982]. A restricted facultative marine methylotrophic bacterium, *M. aminisulfidivorans* MP$^T$, was isolated from seawater in Mokpo, South Korea. This bacterium grows on methanol, methylated amines and dimethylsulfoxide, but not in methane. A novel bacterial *fmo* gene was previously cloned from *M. aminisulfidivorans* MP$^T$ [Choi, 2003]. The amino acid sequence of the gene was highly homologous with that of mammalian FMOs. Flavoprotein monooxygenases are grouped into 6 subclasses on the basis of their structural and enzymatic properties [*van* Berke, 2006]. FMOs, microbial *N*-hydroxylating monooxygenases, and Baeyer-Villiger monooxygenases are grouped as subclass B [Fraaije, 2002]. FMOs are involved in a variety of oxidative biological processes, including drug detoxification and biodegradation of aromatic compounds [Krueger, 2005]. To date, 6 FMO genes have been identified in the human genome. Of these, FMO3 has been shown to be the most important isomer in terms of detoxification of drugs and other xenobiotics. Together with cytochrome P450, and FMO3 plays an important role in the metabolism of non-nutritional xenobiotics [Zhou, 2006]. The

addition of an oxygen atom to lipophilic xenobiotics renders them more soluble and facilitates their excretion. Among non-mammalian eukaryotic FMOs, yeast FMO (yFMO) has been shown to be involved in the oxidation of biological thiols such as glutathione [Suh, 1999], and a plant FMO from *Arabidopsis* catalyzes the hydroxylation of tryptamine during auxin biosynthesis [Zhao, 2001].

The FMO-catalyzed oxygenation reaction can be separated into a reductive reaction followed by an oxidative reaction [Cho, 2011]. In the reductive reaction, reduction of FAD consumes 1 molecule of NADPH, producing NADP+ [Beaty, 1981]. An oxygen molecule then binds to the reduced FAD generating peroxyFAD. During the oxidative reaction, in the absence of a substrate, hydrogen peroxide is produced, whereas in the presence of substrate, an oxygen atom is transferred to the substrate to generate an oxygenated species as the product (Fig. 1-2).

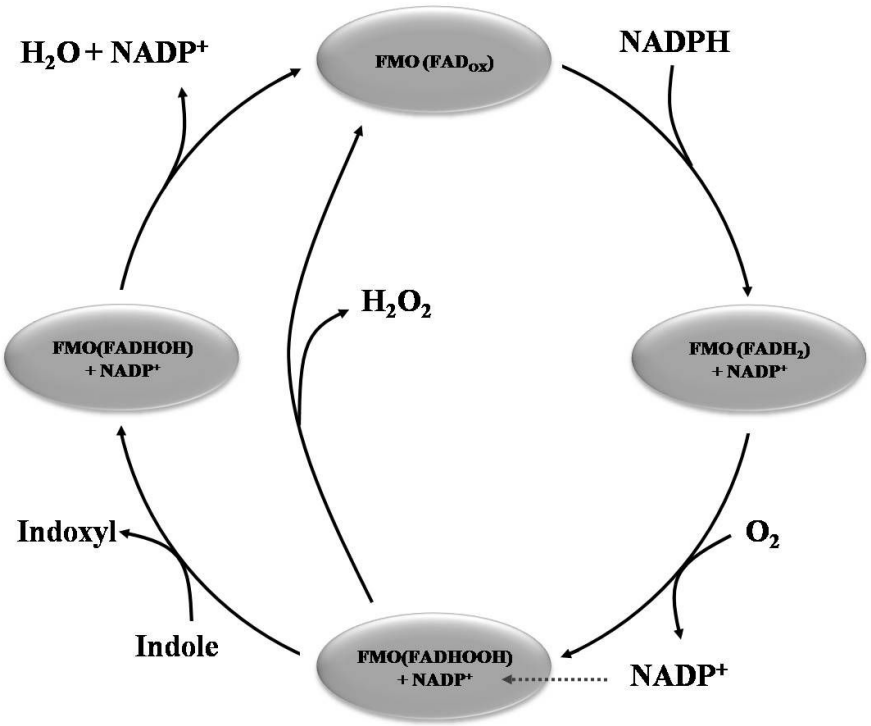



Fig. 1-2. Schematic representation of the catalytic cycle of FMO. In the reductive half of the reaction, NADPH reduces FAD bound to FMO, and the interaction of an oxygen molecule with FAD produces peroxyFAD and $NADP^+$. In the oxidative half of the reaction, in the absence of its substrate, FMO returns to the oxidized form, releasing $H_2O_2$. Indoxyl is produced in the presence of indole, releasing a water molecule [Cho, 2011; Oae, 1985].

## 6. Overall research objective

Indigo and indirubin are very popular dyes; however, their chemical productions are associated with several problems that affect humans and the environment. Various waste products are found during indigo and indirubin production and must be handled carefully. In addition to the reactants described above, other reaction by-products are also produced along with indigo and indirubin. In light of this finding, many researchers have focused on developing new methods of indigo production that are more environmentally friendly. One promising method involves using biocatalysts in the dye reaction process. A novel *fmo* gene responsible for producing indigo and indirubin has been successfully cloned earlier in this lab. In this study, the putative promoter and regulator regions of *fmo* in the plasmid were studies in order to obtain enhanced indigo and indirubin production yields. Additionally, the optimum culture conditions and optimum culture media components for increasing indigo and indirubin compound yields have been investigated (Fig. 1-3). This thesis is divided into 5 chapters and 1 appendix. Chapter II describes the promoter analysis of the *fmo* gene isolated from *M. aminisulfidivorans* $MP^T$. Chapter III and chapter IV describe the optimization of culture condition and mass production of indigo, while chapter V describes the production of indirubin by using an *fmo* gene in *E. coli*. The appendix shows the draft genome sequence of *M. aminisulfidivorans* $MP^T$. The objective of each part of the study has been described in the introduction section of each of the corresponding chapters.

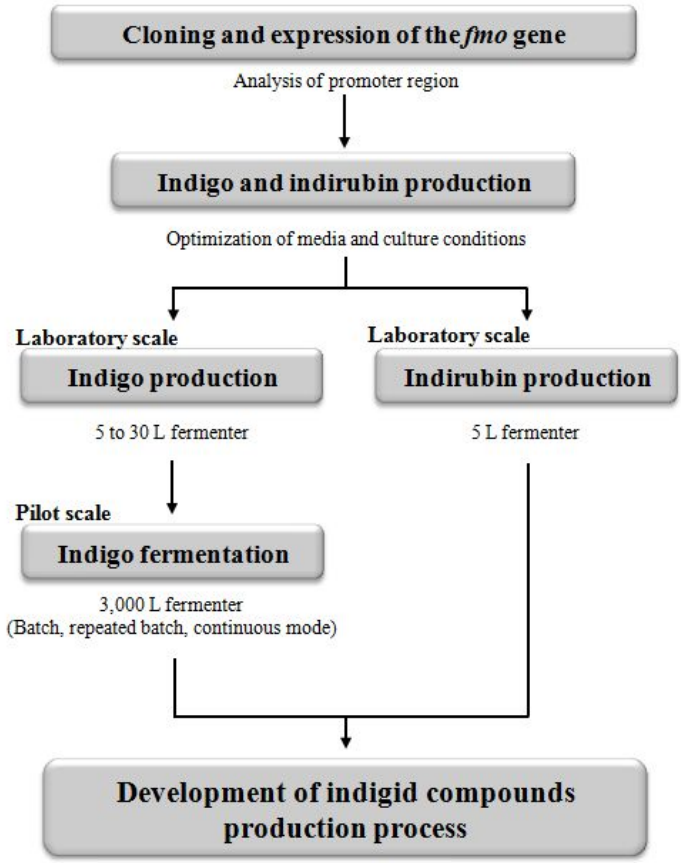



Fig. 1-3. Schematic diagram of the research objective. In this study, the putative promoter and regulator regions of *fmo* in the plasmid were investigated in order to achieve enhanced indigo and indirubin yields. Additionally, the optimum culture conditions and culture media for increased production of indigo and indirubin compounds production were investigated.

# Chapter II

## Analysis of promoter region of flavin-containing monooxygenase gene (*fmo*) from *M. aminisulfidivorans* MP$^{T}$

# 1. Introduction

Flavoprotein monooxygenases are grouped into six subclasses based on their structural and enzymatic properties [*van* Berkel, 2006]. Flavin-containing monooxygenases (FMOs), microbial N-hydroxylating monooxygenases, and Baeyer-Villiger monooxygenases are grouped as subclass B [Fraaije, 2002]. FMOs are involved in a variety of oxidative biological processes, including drug detoxification and biodegradation of aromatic compounds [Krueger, 2005]. FMOs are a polymorphic family of enzymes found in the smooth endoplasmic reticulum of multiple tissues that catalyze the oxidation of soft nucleophilic heteroatom substances to their respective oxides [Cashman, 2004; Krueger, 2005; Ziegler, 1988]. FMO substrates include alkaloids, pesticides, and pharmaceutical compounds [Cashman, 1995; Rettie, 1999]. In mammals, five families of distinct FMO genes have been identified and classified based on the amino acid sequence [Lawton, 1994]. These isoforms differ in tissue distribution, regulation, and substrate specificity [Cashman, 1995; Krueger, 2005; Luo, 2001]. Of these, FMO3 has been shown to be the most important isomer in terms of detoxification of drugs and other xenobiotics. Together with cytochrome P450 and FMO3 plays an important role in the metabolism of non-nutritional xenobiotics [Zhou, 2006]. The addition of an oxygen atom to lipophilic xenobiotics renders them more soluble, and facilitates excretion. Among non-mammalian eukaryotic FMOs, yeast FMO (yFMO) has been shown to be involved in the oxidation of biological thiols such as glutathione [Suh, 1999], and a plant FMO from *Arabidopsis* catalyzes the hydroxylation of tryptamine during auxin biosynthesis [Zhao, 2001].

*M. aminisulfidivorans* $MP^T$ contained a novel bacterial flavin-containig monooxygenase gene (*fmo*) [Choi, 2003]. It contained the three distinct sequence motifs characteristic of the FMO family: FAD binding domain (GXGXXG), NADPH binding domain (GXGXX(A/G)), and FMO-identifying sequence motif (FXGXXXHXXX(Y/F). A novel *fmo* gene reported to be responsible for producing indigo was successfully cloned.

Recombianat *E. coli* pBlue 2.0 can effectively metabolize tryptophan for the production of indigo. Recombianat *E. coli* pBlue 2.0 was constructed in a pBluescript SKII(+) expression system which contained *fmo* structural gene (1361bp) and its upstream region (663bp). Although pBluescript SKII(+) expression system required IPTG as a inducer, pBlue 2.0 did not require IPTG for FMO expression. Promoters are key players in gene regulation. They receive signals from various sources and control the level of transcription initiation, which largely determines gene expression. One rationale for why pBlue 2.0 gave the results was probably attributed to the presence of a putative promoter region. In this study, putative promoter and regulator region of *fmo* in the pBlue 2.0 were investigated to produce indigo.

## 2. Materials and methods

### 2-1. Strains, media and chemicals

*E. coli* DH5α, which harbors a recombinant pBlues construct with a novel bacterial *fmo* was used in this study. The recombinant *E. coli* pBlues were grown on LB containing 100 μg/mL ampicillin at 37℃ overnight. The tryptophan medium (2 g/L tryptophan, 5 g/L yeast extract, and 10 g/L sodium chloride) used for promoter analysis and indigo produciton. All chemicals used in this study were analytical grade and commercially available.

### 2-2. Cloning of different upstream length of *fmo* gene

A plasmid pBlue 2.0 was constructed to express the bacterial FMO in *E. coli* [Walter, 1989]. The plasmid pBlue 2.0 consisted of a 2,024 bp *M. aminisulfidivorans* $MP^T$ DNA fragment, containing 651bp upstream sequences of the *fmo* gene. Recombinant *E. coli* DH5α cells harboring pBlue2.0 produced a small amount of bacterial FMO [Choi, 2003]. To increase the production of indigo, PCR amplification of DNA fragment containing various lengths of upstream regions and *fmo* ORF from *M. aminisulfidivorans* $MP^T$ chromosomal DNA was performed [Merkel, 2003; Walter, 1989]. The PCR primer sets used in this study are summarized in Table 2-1. The PCR products were digested with *Sac*I and *Xba*I. Digested PCR products were ligated with *Sac*I - *Xba* I digested pBluescript SKII (+). Each of these recombinant plasmids were transformed into *E. coli* DH5α.

Table 2-1. Primers used in this study

| Name | Sequence (5' → 3') | Usage |
|---|---|---|
| pBlue 2.2 F | ATA*GAGCTC*TTATTAATCAATAC | construction of pBlues |
| pBlue 2.0 F | ATA*GAGCTC*TCAGTCTGTGAATA | 〃 |
| pBlue 1.9 F | ATA*GAGCTC*TTGCCTGCCATGTTGTATCG | 〃 |
| pBlue 1.8 F | ATA*GAGCTC*CATGTTAAATCAGTCGCAAG | 〃 |
| pBlue 1.7 F | ATA*GAGCTC*ACATTTTCACCGTTTAATAT | 〃 |
| pBlue 1.6 F | ATA*GAGCTC*GGCCGGGGGCTGTACTCGAA | 〃 |
| pBlue 1.5 F | ATA*GAGCTC*TACCCCCTTTTTTACCCTT | 〃 |
| pBlue 1.4 F | ATA*GAGCTC*ATGGCAACTCGTATTGCGA | 〃 |
| pBlue R | TAT*TCTAGA*TTAAGCTTCTTTAGCCACAG | 〃 |
| *fmo* GSP1 | AAATTCAGGCACGTAAGGT | 5' RACE and primer extension |
| *fmo* GSP2 | GTTGAGAAGTGACCGGTAC | 〃 |
| *fmo* GSP3 | ATGGCTACTCGTATTGCG | 〃 |
| 5SF | GAGAGTAGGGAACTGCCA | 〃 |

## 2-3. 5'-rapid amplification of cDNA ends (5'-RACE)

Rapid amplification of cDNA ends (5´-RACE) was initiated by treating total RNA with tobacco acid pyrophosphatase (Epicentre, USA). To ligate the 3´-hydroxyl end of 5S rRNA to the 5´ phosphate end of *fmo* mRNA, a 25 μl of reaction volume containing 10 μl of total RNA, 2.5 μl of 10× reaction buffer (Ambion), 10 units of T4 RNA ligase (Ambion), and 20 units of RNasin (Promega) was incubated at 37℃ for 4 h. Using this RNA preparation, realtime reverse transcription-polymerase chain reaction (RT-PCR) was performed. Total RNA (20 μg) was treated with 15 units of DNase I (Takara) and 1.5 units of alkaline phosphatase (Takara) at 37℃ for 20 min in 100 μl of reaction volume. This alkaline phosphatase treatment removes the 5'-phosphate of the RNA and is necessary to prevent RNA self-ligation during ligation reaction. RNA ligation was performed at 37℃ for 3 h in 50 μl of reaction volume containing 5 μg of total RNA, 2 nM of a synthetic RNA oligomer which has 5'-phosphate and 3'-inverted deoxythymidine (idT, Tharmacon), 10 units of T4 RNA ligase and 20 units of RNasin. The ligated RNA was cleaned using a G-50 column and 1 μg of RNA was reverse transcribed at 37℃ for 2 h in 60 μl of reaction volume containing 4 units of Omniscript$^{TM}$ reverse transcriptase, 0.5 mM of dNTPs, 0.4 μM of *fmo* GSP1 primer complementary to the RNA oligomer and 10 units of RNasin. Two μl of this reaction mixture was used as template for the following PCR amplification. The PCR amplification of *fmo* cDNA was performed using the *fmo* GSP2 primer and the 5SF primer using HotStarTaq$^{TM}$ DNA polymerase. Cloning and sequencing of the amplified 5'-end for the *fmo* mRNA were carried out as described above. To assay the 3'-ends of the *fmo* mRNAs, the amplified cDNA was purified, and used as a template. The primer

extension reaction was performed in 20 μl of γ-$^{32}$P labeled *fmo* GSP3 primer and 1 unit of Taq polymerase (Qiagen). The reaction products were resolved on 8% polyacrylamide-urea sequencing gel, and quantified using Phosphor-Imager (Amersham Biosciences).

## 2-4. Primer extension analysis

When progressive 5'RACE reactions could no longer amplify the additional upstream region from the *fmo* GSP1 primer site, we performed the primer extension from the putative 5'-end. The *fmo* GSP2 primer was end-labeled with [γ-$^{32}$P] dATP and it was annealed to 30 μg of total RNA prepared from *E. coli* pBlue 1.7 at 42℃ for 30 minutes, followed by 10 min. incubation at room temperature. Extension was performed as described in the Primer extension kit (Promega) using 50 U of Avian Myeloblastosis Virus-Reverse Transcriptase at 42℃ for 50 minutes. The reaction was halted by the addition of loading dye (98% formamide, 10 mM EDTA, 0.1% xylene xyanol, 0.1% bromophenol blue). ΦX174 DNA/*Hinf*I dephosphorylated marker was labeled with [γ-$^{32}$P]dATP. Manual sequencing of the fragment was performed using a T7 Sequenase Version 2.0 kit (Amersham Pharmacia Biotech). The sample was denatured for 10 minutes at 90℃ before loading onto 6% sequencing gel. The gel was then vacuum-dried and auto-radiographed for 4 days at −70℃.

## 2-5. Western blot analysis

For western blot analysis, six sheets of 3 MM paper (Whatman), a sheet of nitrocellulose (NC) membrane (Osmonics), and unstained gel were pretreated in the transfer buffer (1× transfer buffer; 2.9 g glycine, 5.8 g Tris-base, 0.37 g SDS, and 200 ml methanol per liter) and assembled. The cell free protein extracted from respective pBlues were separated on SDS-polyacrylamide gel and transferred onto NC membrane by electro-blotting using a transfer electroblot unit (Bio-Rad). Membrane was washed three times with TBS-T buffer (20 mM Tris (pH 7.5), 150 mM NaCl, and 0.1 % Tween 20) and blocked for 4 hours at 25℃ in the same solution with 5 % skim milk. Membrane was probed with antibody (1:50,000 dilutions; *fmo* antibody) and washed three times with TBS-T buffer and incubated in alkaline phosphatase (AP)-conjugated secondary antibody (1:50,000 dilution; Promega Corp., USA). Then, the membrane was washed thrice with TBS-T buffer and rinsed once in AP-substrate buffer (0.1 M Tris (pH 9.5), 100 mM NaCl, 50 mM $MgCl_2$). To visualize antibody-specific proteins, the membrane was reacted with APsubstrate solution (Western Blue Stabilized Substrate Solution, Promega) and the color reaction was terminated with 0.05 M EDTA solution.

# 3. Results and discussion

## 3-1. Identification of transcription start site of *fmo* ORF

Recombinant *E. coli* cell containing pBlue 1.7 showed higher indigo production than that containing pBlue 2.0 and pBlue 1.4, respectively. This is due to the presence of a strong promoter in pBlue 1.7.

Firstly, the transcription start site was identified through the 5'-RACE and primer extension using plasmid pBlue 1.7. The relative number of transcripts initiated from the promoter of *fmo* gene was determined during different growth phases. The *E. coli* pBlue 1.7 was grown under tryptophan medium. The transcription level of *fmo* gene was compared in the early-exponential phase (4h, $OD_{600}$ = 0.6) and mid-exponential phase (12h, $OD_{600}$ = 2.0) during *E. coli* pBlue 1.7 growth. Total RNA was prepared from a fixed number of cells, $1 \times 10^8$ cells, as determined by optical density and colony forming unit. After RNA ligation, the cDNA was generated and the amplified cDNA sequences include the 5S RNA and 5'-ends of the *fmo* transcripts. The primer extension assay using this products and a $^{32}$P-labeled primer was performed using standard procedure. The result was visualized by electrophoresis on DNA sequencing gel. Six transcription start sites (TSS) of the *fmo* gene were identified (Fig. 2-1). Six signals were located at -279 (TSS1), -129 (TSS2), -65 (TSS3), -59 (TSS4), -48 (TSS5) and -30 (TSS6), respectively. TSS1 was present in the pBlue 1.7 region and TSS2-TSS6 were located in the pBlue 1.5 region (Fig. 2-2). The TSS1 and TSS2 signal were increased on mid-exponential phase than early-exponential phase.

Secondly, the expression levels of FMO in cells cultured between

early-exponential phase (4h, $OD_{600}$ = 0.6) and mid-exponential phase (12h, $OD_{600}$ = 2.0) were compared by western blot analysis (Fig. 2-3). Early-exponential phase, all *E. coli* cells containing pBlues were not produce indigo, the pBlue 1.7 showed the highest expression level. Mid-exponential phase, initial stage of indigo production, the recombinant *E. coli* harboring pBlue 1.7 observed the highest expression level of FMO and indigo production. However, FMO expression level was not increasing at any appreciable rate.

5'RACE and western blot results showed, The TSS1 is an important part of FMO expression and indigo production in the *E. coli* harboring pBlue 1.7. Based on the sequenced data of the *fmo* transcripts, the transcription of the *fmo* gene was T(+1) which is in 279bp (TSS1) upstream from the translation initiation codon of the *fmo* gene. The -35 region and –10 region were identified 5'-CTGGAA-3' and 5'-TATCCT-3', respectively. In addition, the sequence consist of the TG at -14, -15bp in the upstream region from the TSS1 that is considered as the extended -10 promoter.

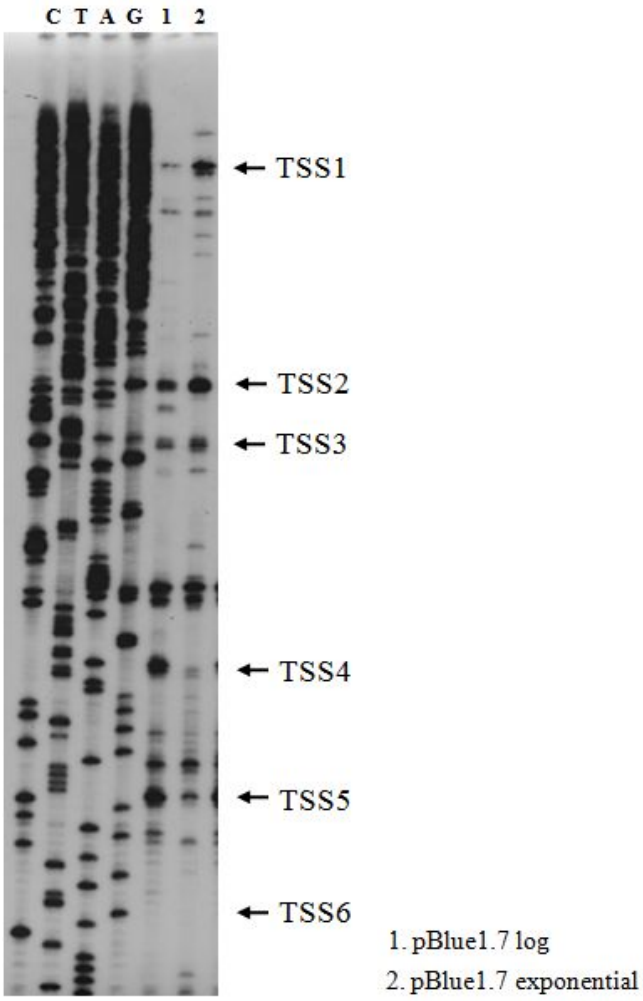



Fig. 2-1. Identification of the transcription start site of *fmo* gene. The transcription start site of *fmo* gene was identified by 5'-RACE and primer extension, mapping used $^{32}$P-labbled oligonucleotide primer which is complementary to the nucleotide position 224 to 244 downstream of the *fmo* gene. 1, early-exponential phase (4h, $OD_{600}$ = 0.6); 2, mid-exponential phase (12h, $OD_{600}$ = 2.0).

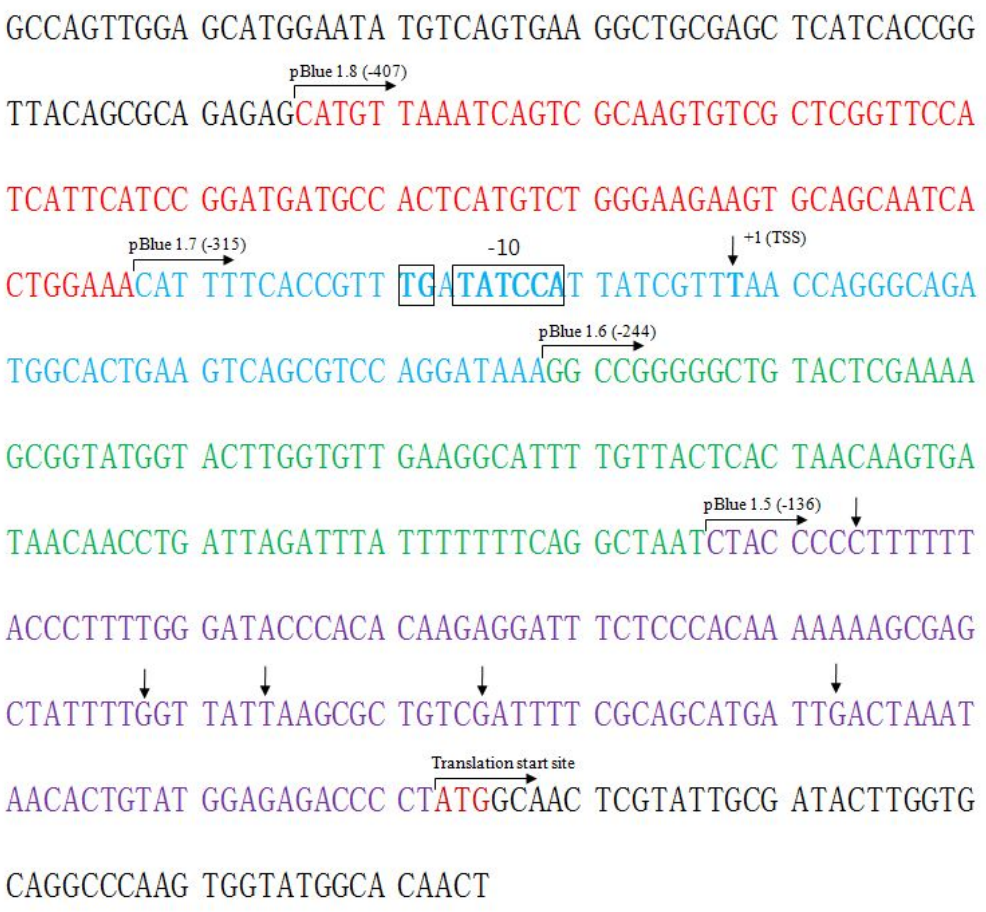



Fig. 2-2. Transcription start site mapping. Six signals were located at -279 (TSS1), -129 (TSS2), -65 (TSS3), -59 (TSS4), -48 (TSS5) and -30 (TSS6), respectively. TSS1 was present in the pBlue 1.7 region and TSS2-TSS6 were located in the pBlue 1.5 region

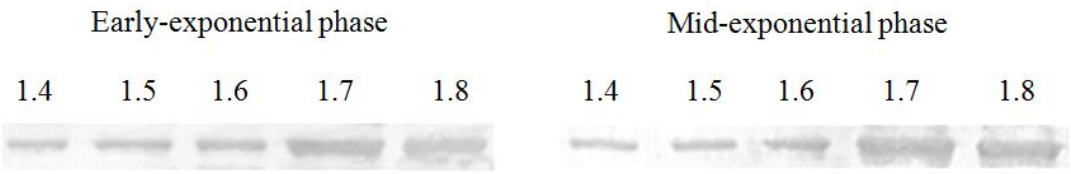



Fig. 2-3. Expression level of FMO in recombinant *E. coli* by western blot analysis.

### 3-2. Relationship between *fmo* transcripts and indigo production

The relative number of *fmo* transcripts determined the pBlue 1.5, 1.7 and 1.8 by real time PCR. These results showed the differences in indigo productivity. The 5'-RACE assay requires the RNA ligation reaction that might lower the numbers cDNA templates (*fmo* transcripts) for the following DNA amplification procedures. A designed forward primer named as *fmo* GSP1 used for 5'-RACE and primer extension for all pBlues. The highest amount of transcripts of *fmo* gene was obtained in pBlue 1.7, showed the highest indigo production. On the other hand the pBlue 1.5 showed the lowest amount of transcripts. The number of trasncripts of *fmo* gene is proportional to the indigo production. From the results, it was confirmed to accurately concur the relationship between the amount of *fmo* gene transcript and indigo production (Table 2-2).

Table 2-2. Relationship between *fmo* transcripts and indigo production

| | Real time PCR (cycles) | Indigo production (mg/L) |
|---|---|---|
| pBlue 1.5 | 19.36 | 2.6 |
| pBlue 1.7 | 13.43 | 190 |
| pBlue 1.8 | 17.56 | 166 |

# Chapter III

# Optimization of indigo production using recombinant *E. coli* harboring *fmo* gene

This chapter is published in Enzyme and Microbial Techenology 42 (7) 617-623 (2008)

# 1. Introduction

Indigo is considered the oldest textile dye, which is used for the dyeing of cotton and wool fabrics. It was traditionally It was traditionally extracted from several plant species, such as *Indigofera* sp. and *Polygonum tinctorium*. Since the late 19th century, plant-derived indigo has been replaced in the textile industry by the use of synthetic indigo. The annual production of synthetic indigo is estimated to be 22,000 metric tons [Schhrott, 2000]. Recently, several investigations on the biological production of indigo from indole using recombinant microorganisms expressing mono- or dioxygenase have been reported [Bae, 2006; Doukyu, 2002; Madsen, 1988; Murdock, 1993]. Thus far, the substrates used for production of bio-indigo using recombinant bacteria have been glucose and indole. Glucose is inexpensive and readily available, but its metabolic pathway is so complicated that the resultant bio-indigo yield is low. Indole is toxic, causing a reduction in the production of indigo by the microorganisms [Doukyu, 1997]. Therefore, the indole concentration in the culture medium must be kept low to avoid this toxic effect. The above two substrates are unsatisfactory for the production of indigo and; moreover, the productivity of indigo is so low that this process could not be considered as an alternative to chemical indigo.

Methylotrophic bacteria utilize $C_1$-compounds, such as methane, methanol and methylated amines, as sole carbon and energy sources [Biasiolo, 1995]. A restricted facultative marine methylotrophic bacterium, *Methylophaga aminisulfidivorans* $MP^T$, was isolated from seawater at Mokpo, South Korea [Choi, 2003; Kim, 2007]. This bacterium grows on methanol, methylated amines and dimethyl sulfoxide, but not on methane.

Previously, a novel flavin-containing monooxygenase (FMO) gene reported to be responsible for producing indigo was successfully cloned [Choi, 2003]. The deduced amino acid sequence of the gene showed rather high similarity with mammalian FMOs that convert nitrogen- and sulfur- containing xenobiotics to more hydrophilic metabolites using oxygen through an NADPH-dependent pathway [Ziegler, 1988]. This seems to be the first member of the FMO family characterized in prokaryotes.

Recombinant *E. coli* cells harboring a single *fmo* gene can effectively metabolize tryptophan for the production of indigo. Tryptophan has been converted to indole by tryptophanase in the cell, and indole oxidized to 2-hydroxyindole and 3-hydroxyindole by FMO [McClay, 2005; Rui, 2005]. Thereafter, two molecules of 3-hydroxyindole was combined to produce a indigo molecule and one molecule of 2-hydroxyindole and one molecule of 3-hydroxyindole combined to give a indirubin in the presence of oxygen (Fig. 3-1) [Eaton, 1995; Maugard, 2002].

In this study, attempts were made to optimize the production of indigo using a gene deletion approach with the *fmo* gene, and the expression of FMO in various bacteria hosts such as *E. coli* DH5α, *Bacillus subtilis* and *Corynebacterium glutamicom*. Moreover, the culture conditions and culture media compositions were optimized using response surface methodology (RSM) for maximum indigo production.

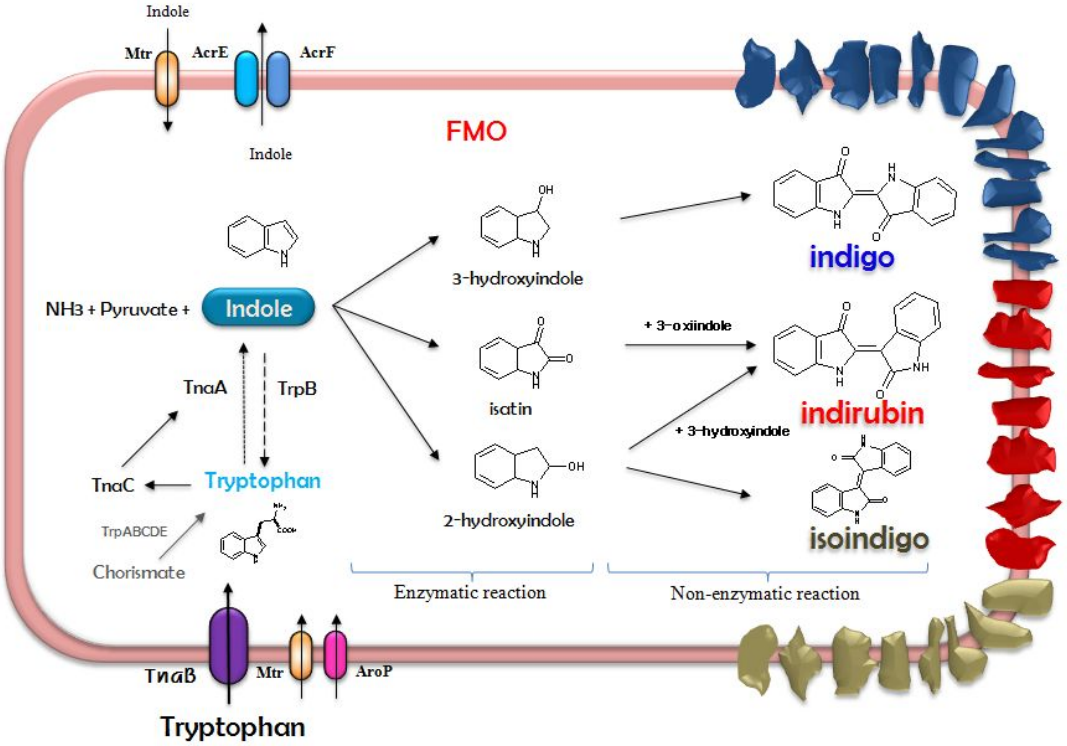



Fig. 3-1. Proposed biosynthesis pathway of indigoid compounds production in *E. coli*. The *Tna*A (dotted line) and *Trp*B (dashed line) indicate feedback inhibition and repression, respectively [Lee, 2007; Lee, 2011]. The indole converted indole derivatives by FMO activity, thereafter indigoid compounds were synthesized by the combination of the above compounds in the oxygen dependent condition.

# 2. Materials and methods

## 2-1. Materials

Dimethyl sulfoxide (DMSO) was purchased from Sigma-Aldrich (St. Louis, USA). Restriction enzyme, T4 ligase and Taq polymerase were purchased from Takara (Shiga, Japan), and the indigo standard from Fluka (St. Gallen, Switzerland). All other chemicals used were of analytical grade.

## 2-2. Strains, plasmids, media and fermenters

For FMO gene expression the bacteria strains such as *E. coli* DH5α, *Bacillus subtilis* and *Corynebacterium glutamicom* and the vectors like PHCEIIB, pGEX4T-1, PBSK, PHY300PLK and PEKEx2 were used. The recombinant bacteria cells were grown on LB agar plates at 37℃ for overnight. A single colony was inoculated in 100 mL of tryptophan medium (seed culture) and incubated at 30℃ for 16h. The seed culture was inoculated in a 10 L fermenter (Bioflow 3000 model, New Brunswick) with 2 L working volume.

## 2-3. General DNA manipulation

Total DNA from *M. aminisulfidivorans* MP$^{T}$ was prepared using an iNtRONBio G-spin$^{TM}$ Genomic DNA extraction kit (iNtRON Biotechnology, Seongnam, Gyeonggi-Do, Korea). Plasmid DNA was obtained using a NucleoGen plasmid purification mini kit (NucleoGen, Sihueng, Gyeonggi-Do, Korea). DNA ligation, transformation and other recombinant DNA manipulation methods followed the standard

methods described by Sambrook et al. [Sambrook, 2001]. DNA sequencing and primer synthesis were performed at Bionics Inc. (Seoul,Korea).

## 2-4. Deletion of plasmid pBlue 2.0

A plasmid pBlue 2.0 was constructed to express the bacterial FMO in *E. coli* [Kim, 2007]. The plasmid pBlue 2.0 consisted of a 2,024 bp *M. aminisulfidivorans* $MP^T$ DNA fragment, containing 651bp upstream sequences of the *fmo* gene. Recombinant *E. coli* DH5α cells harboring pBlue2.0 produced a small amount of bacterial FMO [Choi, 2003]. To increase the production of indigo, PCR amplification of DNA fragment containing various lengths of upstream regions and *fmo* ORF from *M. aminisulfidivorans* $MP^T$ chromosomal DNA was performed [Merkel, 2003; Walter, 1989]. The PCR primer sets used in this study are summarized in Table 3-1. The PCR amplification was carried out under the following conditions: 95℃ for 5min of initial denaturation; 30 cyles of denaturation at 95℃ for 30 s, annealing at 58℃ for 30s, extension at 72℃ for 1min, and a final extension at 72℃ for 5min. The amplified products were visualized by agarose gel electrophoresis and the correctly sized product was purified and concentrated by PCR purification kit (Invitrogen). The PCR products were digested with *Sac*I and *Xba*I (Fig. 3-2). Digested PCR products were ligated with *Sac*I - *Xba* I digested pBluescript SKII (+). Each of these recombinant plasmids were transformed into *E. coli* DH5α.

Table 3-1. Primers used in this study

| Name | Sequence (5' → 3') | Usage |
|---|---|---|
| pBlue 2.2 F | ATA*GAGCTC*TTATTAATCAATAC | construction of pBlues |
| pBlue 2.0 F | ATA*GAGCTC*TCAGTCTGTGAATA | 〃 |
| pBlue 1.9 F | ATA*GAGCTC*TTGCCTGCCATGTTGTATCG | 〃 |
| pBlue 1.8 F | ATA*GAGCTC*CATGTTAAATCAGTCGCAAG | 〃 |
| pBlue 1.7 F | ATA*GAGCTC*ACATTTTCACCGTTTAATAT | 〃 |
| pBlue 1.6 F | ATA*GAGCTC*GGCCGGGGGCTGTACTCGAA | 〃 |
| pBlue 1.5 F | ATA*GAGCTC*TACCCCCTTTTTTACCCTT | 〃 |
| pBlue 1.4 F | ATA*GAGCTC*ATGGCAACTCGTATTGCGA | 〃 |
| pBlue R | TAT*TCTAGA*TTAAGCTTCTTTAGCCACAG | 〃 |
| fmo-F-SacI | TAT*GAGCTC*CATTTTCACCGTTTGATAT | pHCE II B |
| fmo-R-XbaI | TAT*TCTAGA*TTAAGCTTCTTTAGCCACAG | pHCE II B |
| fmo-F-PstI | TAT*CTGCAG*CATTTTCACCGTTTGATAT | pEKE x2 |
| fmo-R-KpnI | TAT*GGTACC*TTAAGCTTCTTTAGCCACA | pEKE x2 |
| fmo-F-XbaI | TAT*TCTAGA*CATTTTCACCGTTTGATATCC | pHY300PLK |
| fmo-R-PstI | TAT*CTGCAG*TTAAGCTTCTTTAGCCACA | pHY300PLK |
| fmo-F-BamHI | TAT*GGATTC*CATTTTCACCGTTTGATAT | pGEX-4T-1 |
| fmo-R-XhoI | TAT*CTCGAG*TTAAGCTTCTTTAGCCA | pGEX-4T-1 |

## 2-5. Selection of hosts for indigo production

Amplification of gene encoding flavin-containing monooxygenase from pBlue1.7 by PCR was carried out with Genepro (Bioer). The PCR products were cloned into the vector like PHCEIIB, pGEX4T-1, PBSK, PHY300PLK and PEKEx2. The recombinant plasmid PHCEIIB 1.7, pGEX4T-1 1.7 and PBSK 1.7 (pBlue 1.7) were transformed into *E. coli* DH5α, PHY300PLK 1.7 was transformed into *Bacillus subtilis*, PEKEx2 1.7 was transformed into *Corynebacterium glutamicom*, respectively. The nucleotide sequencing and DNA oligonucleotide syntesis were performed commercially at the DNA sequencing facility of Bionics (Seoul).

## 2-6. Statistical procedure

The production of indigo was optimized using a Response Surface Methodology (RSM) based, two type $2^n$ central composite design (CCD) [Gasparetti 2006; Kim, 2004; Myers, 2002]. The first experiment was based on a $2^3$ factorial central composite experimental plan with three medium components, i.e, tryptophan, yeast extract and sodium chloride. The range and levels of the experimental variables investigated in this study are given in Table 3-2. A set of 20 experiments was carried out with three variables, with each variable studied at five levels ($\alpha$=1.5). The second experimental design was a $2^2$ factorial central composite experimental plan with two culture conditions; pH and temperature. A set of 13 experiments was carried out with two variables, with each variable studied at five levels ($\alpha$=1.5). For statistical calculations including the solutions of the second-order polynomial model equation [Myers, 2002]; The Design Expert (version 7.0, Stat-Ease Inc., Minneapolis, USA) software package was used.

Table 3-2. Experimental design (range, levels and response) to assess the effects of medium composition (A), pH and temperature (B)

| (A) Effect of medium composition | | | | | | | | (B) Effect of pH and temperature | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RUN[1] | Tryptophan[2] $A_1$ (%) | | Yeast extract[3] $A_2$ (%) | | NaCl[4] $A_3$ (%) | | Indigo concentration[5] $Y_1$ (mg/l) | RUN[1] | pH[6] $B_1$ | | Temperature[7] $B_2$ (℃) | | Indigo concentration[5] $Y_1$ (mg/l) |
| 1 | 0.1 | -1.0 | 0.4 | -1.0 | 1.0 | -1.0 | 409.1 | 1 | 6.6 | -1.0 | 25 | -1.0 | 2.04 |
| 2 | 0.3 | +1.0 | 0.4 | -1.0 | 1.0 | -1.0 | 746.2 | 2 | 7.4 | +1.0 | 25 | -1.0 | 114.2 |
| 3 | 0.1 | -1.0 | 0.6 | +1.0 | 1.0 | -1.0 | 418.1 | 3 | 6.6 | -1.0 | 35 | +1.0 | 81.9 |
| 4 | 0.3 | +1.0 | 0.6 | +1.0 | 1.0 | -1.0 | 757.9 | 4 | 7.4 | +1.0 | 35 | +1.0 | 136.0 |
| 5 | 0.1 | -1.0 | 0.4 | -1.0 | 2.0 | +1.0 | 421.7 | 5 | 6.4 | -1.0 | 30 | 0.0 | 53.6 |
| 6 | 0.3 | +1.0 | 0.4 | -1.0 | 2.0 | +1.0 | 755.4 | 6 | 7.6 | +1.0 | 30 | 0.0 | 20.6 |
| 7 | 0.1 | -1.0 | 0.6 | +1.0 | 2.0 | +1.0 | 466.5 | 7 | 7.0 | 0.0 | 22.5 | -1.5 | 101.9 |
| 8 | 0.3 | +1.0 | 0.6 | +1.0 | 2.0 | +1.0 | 781.1 | 8 | 7.0 | 0.0 | 37.5 | +1.5 | 70.6 |
| 9 | 0.05 | -1.5 | 0.5 | 0.0 | 1.5 | 0.0 | 224.6 | 9 | 7.0 | 0.0 | 30 | 0.0 | 846.3 |
| 10 | 0.35 | +1.5 | 0.5 | 0.0 | 1.5 | 0.0 | 752.7 | 10 | 7.0 | 0.0 | 30 | 0.0 | 873.9 |
| 11 | 0.2 | 0.0 | 0.35 | -1.5 | 1.5 | 0.0 | 741.2 | 11 | 7.0 | 0.0 | 30 | 0.0 | 880.0 |
| 12 | 0.2 | 0.0 | 0.65 | +1.5 | 1.5 | 0.0 | 757.6 | 12 | 7.0 | 0.0 | 30 | 0.0 | 829.3 |
| 13 | 0.2 | 0.0 | 0.5 | 0.0 | 0.75 | -1.5 | 722.7 | 13 | 7.0 | 0.0 | 30 | 0.0 | 850.1 |
| 14 | 0.2 | 0.0 | 0.5 | 0.0 | 2.25 | +1.5 | 739.0 | | | | | | |
| 15 | 0.2 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 752.9 | | | | | | |
| 16 | 0.2 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 762.2 | | | | | | |
| 17 | 0.2 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 748.3 | | | | | | |
| 18 | 0.2 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 750.5 | | | | | | |
| 19 | 0.2 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 762.8 | | | | | | |
| 20 | 0.2 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 740.7 | | | | | | |

[1]Run order

[2]Left col.; actual tryptophan concentration(%), right col.; level (coded unit)

[3]Left col.; actual yeast extract concentration(%), right col.; level (coded unit)

[4]Left col.; actual sodium chloride concentration(%), right col.; level (coded unit)

[5]Actual indigo concentration (mg $l^{-1}$)

[6]Left col.; actual pH, right col.; level (coded unit)

[7]Left col.; actual temperature (℃), right col.; level (coded unit)

## 2-7. Determination of indigo

After the flask and fermentor cultures, the culture broth was centrifuged at 10,000 x *g* for 1 min to yield dark blue pellets, which were collected and washed twice with water. The material was then suspended in 10 ml of DMSO, and subjected to repeated sonication (Branson, Danbury, USA) with a microprobe for 1 min. The amount of bio-indigo produced in the supernatant was determined using a standard indigo solution dissolved in DMSO. High-performance liquid chromatography, equipped with a photodiode array detector (Agilent 1200 HPLC, Santa Clara, USA), was used, with an eluent flow rate of 1 ml/min, and monitored at 620 nm. A column used was YMC-pack ODS-A ($C_{18}$, 250× 4.6 mm) with isocratic elution of methanol and water (80:20, v/v). The amount and purity of indigo was estimated as described previous study [Han, 2008; Han, 2011].

# 3. Results

## 3-1. Construction of recombinant plasmid pBlues

Previously, the maximum indigo production (160 mg/L) was obtained by recombinant *E. coli* pBlue 2.0, which containing a 2,024bp DNA fragment of *M. aminisulfidivorans* $MP^T$ was reported [Choi, 2003]. In this study, each primer sets were designed to give various PCR products containing the *fmo* gene and its upstream regions (Fig. 3-2). Transformants corresponding to each pBlues containing various truncated inserts were cultured and their indigo production was measured. After culturing *E. coli* pBlues on LB medium at 37℃ for 16h, the blue colonies were observed. Each colony was picked and transferred into the tryptophan medium. Fig. 3-2 shows the results of the indigo production from the pBlues containing various truncated inserts. Among the variants, pBlue 1.7 showed the most abundant indigo production (822±22.2 mg/L). The pBlue 1.4 and 1.5 produced no bio-indigo at all.

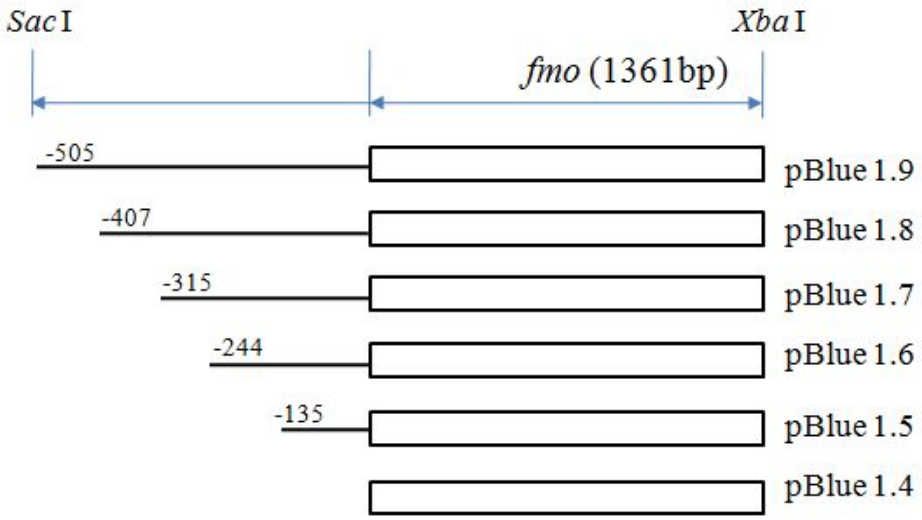



Fig. 3-2. Construction of recombinant plasmid pBlues. The pBlues containing various lengths of upstream regions and *fmo* ORF.

## 3-2. Indigo production using different bacterial hosts harboring *fmo* gene

Each recombinant bacteria cells were grown on LB agar plates at 37℃ overnight. A single colony was inoculated in 100 mL of tryptophan medium and incubated at 30℃ for 16h. Fig. 8 shows the results of the time course of indigo production from the *B. subtilis* pHY300PLK 1.7 and *C. glutamicum* pEKEx2 1.7 for 24h cultivation. The culture conditions were as follows: temperature 30℃, pH 7.0 and agitation speed 200 rpm. Under these conditions, *B. subtilis* pHY300PLK 1.7 was produced a maximum of 182 mg/L indigo for 24h cultivation, *C. glutamicum* pEKEx2 1.7 was produced 377 mg/L indigo, respectively (Fig. 3-3).

The recombinant *E. coli* DH5α applied three different expression vector for indigo production. Culture conditions was described above. Among the variants, *E. coli* pBSK 1.7 (pBlue 1.7) showed the most abundant production of indigo (673.2 mg/L). The *E. coli* pGEX4T 1.7 and *E. coli* pHCEIIB 1.7 produced 638.5 mg/L and 400.2 mg/L of indigo, respectively (Fig. 3-4). The *E. coli* pBSK 1.7 (pBlue 1.7) and *E. coli* pGEX4T 1.7 showed similar indigo productivity, however the *E. coli* pBSK 1.7 (pBlue 1.7) was required IPTG induction for FMO expression. Thus, *E. coli* pBSK 1.7 (pBlue 1.7) was for further experiments.

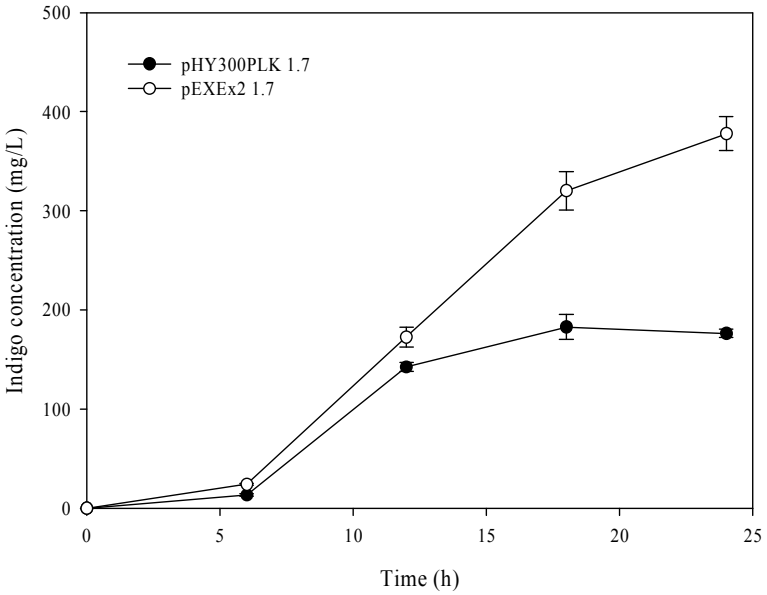



Fig. 3-3. The indigo production by recombinant *B. subtilis* pHY300PLK 1.7 and *C. glutamicum* pEKEx2 1.7. The culture conditions were as follows: temperature 30℃, pH 7.0 and agitation speed 200 rpm.

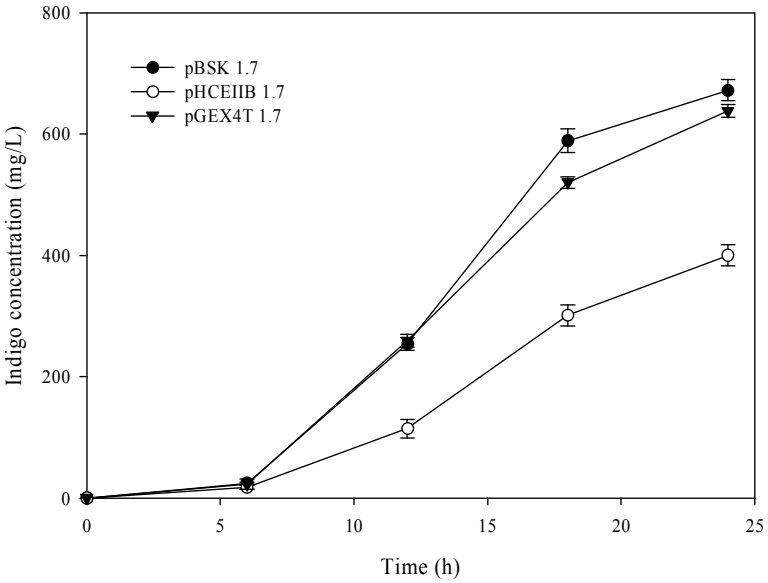



Fig. 3-4. The indigo production by recombinant *E. coli* pBSK 1.7 (pBlue 1.7), *E. coli* pGEX4T 1.7 and *E. coli* pHCEIIB 1.7. The culture conditions were as follows: temperature 30℃, pH 7.0 and agitation speed 200 rpm.

## 3-3. Setup of model equations

To search for the pin point optimization of medium compositions (tryptophan, yeast extract, and sodium chloride) and culture conditions (pH and temperature) for the best production of indigo, the experiments were performed according to a central composite design experimental plan (Table 3-2). The initial ranges were selected based on our preliminary work. The mathematical model, which represents a second order polynomial, is given by Eqns (1) and (2), where the variables take their coded values. $Y_1$ and $Y_2$ represent the responses, which are dependent variables and the indigo concentration (mg/L). The two experiments were performed independently. $A_1$, $A_2$ and $A_3$ are the coded values of the independent variables; tryptophan, yeast extract and sodium chloride, in the medium, respectively. $B_1$ and $B_2$ are the coded values of the independent variables, pH and temperature, respectively.

$$Y_1 = 756.98 + 170.98A_1 + 7.66A_2 + 7.83A_3 + 0.44A_1A_2 - 1.05A_1A_3 + 3.71A_2A_3 - 126.47A_1^2 - 10.57A_2^2 - 18.82A_3^2 \quad (1)$$

$$Y_2 = 852.38 + 13.73B_1 + 6.44B_2 - 14.51B_1B_2 - 367.91B_1^2 - 355.07B_2^2 \quad (2)$$

The fit of the models was expressed by the coefficient of determination, $R^2$, which were 0.9946 and 0.9895, respectively. The $R^2$ value closer to 1, shows the stronger model and better predict the responses. The values of the adjusted determination coefficient (Adj $R^2$) (0.9896 and 0.9819) were also high, advocating the significance of each model, with the predicted $R^2$ (0.9655 and 0.9296) in reasonable agreement with

the Adj $R^2$. In addition, the statistical significance of the second-order model equations was checked using the F-test analysis of variance (ANOVA) (Table 3-3). These results, indicates the good agreement between the experimental and predicted values for the medium composition and also culture conditions.

Table 3-3. Analysis of variance for the effects of medium composition (A), pH and temperature (B) using coded values

| Source | SS | DF | MS | F-value | Prob (P) > F |
|---|---|---|---|---|---|
| Effect of medium composition | | | | | |
| : [$R^2$=0.9946, Adj $R^2$=0.9896, Predicted $R^2$=0.9655, CV=2.61%] | | | | | |
| Model | 5.383E+005 | 9 | 59812.28 | 202.02 | <0.0001 |
| Residual (error) | 2960.75 | 10 | 296.07 | | |
| Lack of Fit | 2599.74 | 5 | 519.95 | 7.20 | 0.0246 |
| Pure Error | 361.01 | 5 | 72.20 | | |
| Total | 5.413E+0.005 | 19 | | | |
| Effect of pH and temperature | | | | | |
| : [$R^2$=0.9895, Adj $R^2$=0.9819, Predicted $R^2$=0.9296, CV=14.33%] | | | | | |
| Model | 1.883E+006 | 5 | 3.766E+005 | 131.32 | <0.0001 |
| Residual (error) | 20074.78 | 7 | 2867.83 | | |
| Lack of Fit | 18361.47 | 3 | 6120.49 | 14.29 | 0.0133 |
| Pure Error | 1713.31 | 4 | 428.33 | | |
| Total | 1.903E+006 | 12 | | | |

SS, sum of squares; DF, degrees of freedom; MS, mean squares

## 3-4. Optimization of medium composition

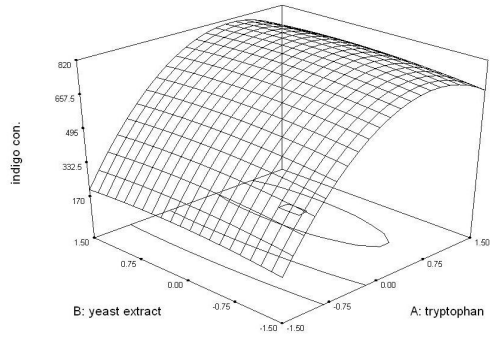
Using the response surface plot, the interaction between two variables and their optimum levels can be easily understood and located. The plots showing the interaction between tryptophan and yeast extract, tryptophan and sodium chloride, and yeast extract and sodium chloride are depicted in Fig. 3-5. An increase in the tryptophan with yeast extract up to the optimum point increased the production of indigo, with a further increase trending to cause the reverse (Fig. 3-5(A)). An increase in the tryptophan and sodium chloride gradually increased the production of indigo, however, with further increment in the tryptophan with sodium chloride concentration shows reverse trend (Fig. 3-5(B)). The numerical method given by Myers and Montgomery was used to solve the model equation of (1). The optimal values of the test variables in the coded units were as follows: $A_1$ = 0.36, $A_2$ = -0.54, and $A_3$ = -0.72. The natural values obtained by putting the respective values of Ai into Eqn (1) were: tryptophan = 2.4 g/L, yeast extract = 4.5 g/L and sodium chloride = 11.4 g/L, with indigo production = 781 mg/L. This model was confirmed experimentally. The maximum production of indigo obtained experimentally was 782 mg/L at pH 7.0 and temperature 30℃, which was in close agreement with the model prediction.

## 3-5. Optimization of temperature and pH

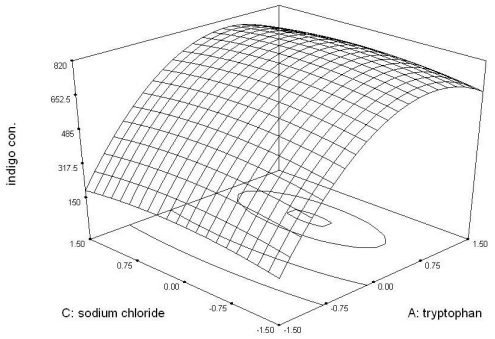
Fig. 3-6 shows the response and contour plots of the pH and temperature of a flask culture. The optimum points for maximum production of indigo lie near the center point of the pH and temperature ranges. The optimal theoretical values of the test variables in the coded units of Eqn (2) were solved as follows: $B_1$ = 0.02 and $B_2$ =

0.01 with the corresponding $Y_2$ = 853 mg/L. Experimental value of $Y_2$ was confirmed to be 880 mg/L. The natural values obtained by substituting the respective values of $B_i$ in Eqn (2) were: pH = 7.0 and temperature = 30℃. To confirm the predicted optimization conditions, the flask culture was performed using the above conditions proposed by the optimization model. Similarly, the experimental indigo production was 880 mg/L in the optimized tryptophan medium. The observed experimental results was close agreement with the values of model prediction. Experimentally obtained indigo production after the optimization process was more than 30% higher than previous optimization.

**[A]**




**[B]**




**[C]**

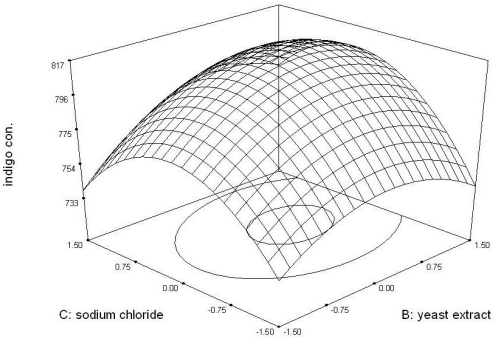



Fig. 3-5. Response surface plot showing the effect of yeast extract, sodium chloride and tryptophan on the production of indigo (mg/L).

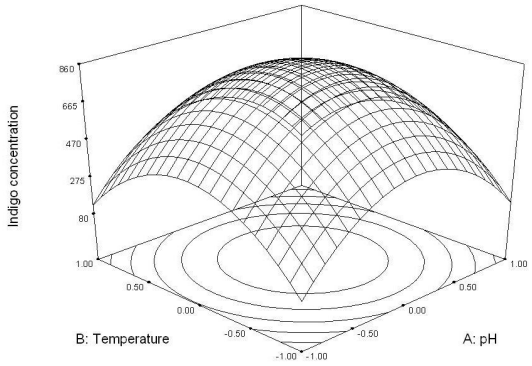



Fig. 3-6. Response surface plot showing the effects of pH and temperature as well as their mutual effects on the production of indigo (mg/L).

## 3-6. Indigo production under optimized conditions

The possibility to fine tune the optimum medium concentration, pH and temperature in addition to gene deletion for maximizing the bio-indigo production may be considered the key results of this study. With the optimized media compositions (2.4 g/L tryptophan, 4.5 g/L yeast extract and 11.4 g/L sodium chloride) and culture conditions of the flask culture (pH 7.0, temperature = 30.0℃) obtained from flask culture, a batch fermentation in 5 L fermentor (KoBioTech model: KF-5l) was performed using recombinant *E. coli* DH5α with pBlue 1.7. A maximum amount of 920 mg/L of indigo was produced from the recombinant *E. coli* cells (Fig. 3-7).

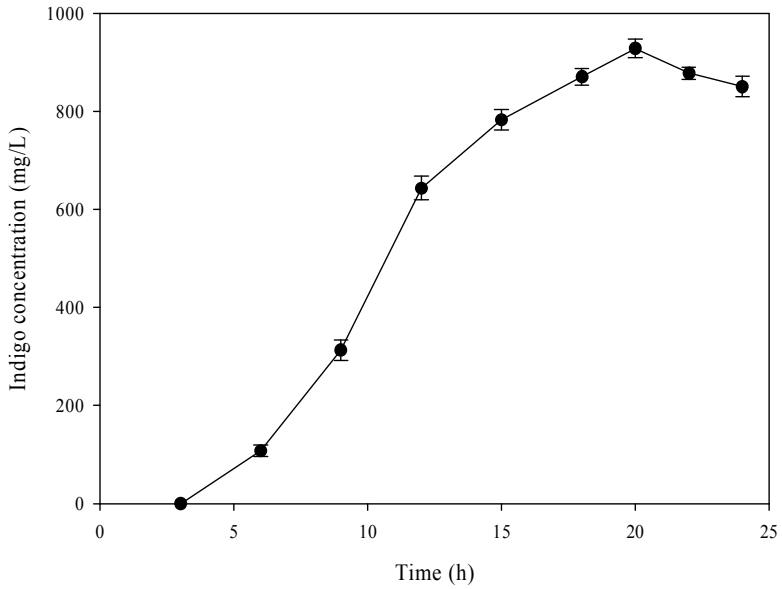



Fig. 3-7. Production of indigo by recombinant *E. coli* in the 5 L fermenter under the optimum conditions.

# 4. Discussion

In this study, we demonstrated indigo production systems using a recombinant *E. coli* strain harboring a novel *fmo* gene derived from *M. aminisulfidivorans* $MP^T$, which can be used commercially to produce indigo dye. This is the first report on the gram-scale production of indigo using microbial culture with tryptophan medium. Truncation of inserts resulted in a 413% increase in the production of indigo; from 160 to 662 mg/L, in the flask culture of recombinant *E coli* with plasmid pBlue 1.7 in the tryptophan medium (Table 3-4). RSM and the use of a central composite factorial design were able to determine optimized conditions for the maximum production of indigo. The optimal production of indigo was determined as follows: tryptophan medium (2.4 g/L tryptophan, 4.5 g/L yeast extract and 11.4 g/L sodium chloride) and culture conditions (pH 7.0, temperature = 30.0℃). The model predicted that the maximum production of indigo able to be obtained using the combination of the above optimal conditions was theoretically 853 mg/L and experimentally 880 mg/L based on the flask culture (Table 3-4). The experimental result for the production of indigo was 920 mg/L using a 5 L fermenter. For the optimization of the production of indigo, a molecular approach and experimental design have to be simultaneously considered to obtain the best results. The optimization results were summarized in Table 3-4. For an industrially available bioproduct, the concentrations of indigo obtained are so high that the proposed production system ensures industrial viability. In conclusion, these indigo production systems could be used for industrial applications to overcome the environmental problems associated with synthetic indigo production and to meet the demand for natural indigo in the dye and textile industries.

Table 3-4. Summary of optimization methods and results

| Optimization methods | | Conditions | Indigo production (mg/L) |
|---|---|---|---|
| Truncated insert | pBlue 2.0 | Medium (tryptophan 2 g/L, YE 5 g/L, NaCl 10 g/L) Culture condition (30℃, pH 7.0) | 160 |
| | pBlue 1.7 | Medium (tryptophan 2 g/L, YE 5 g/L, NaCl 10 g/L) Culture condition (30℃, pH 7.0) | 662 |
| RSM predicted conditions | predicted optimized medium | Medium (tryptophan 2.4 g/L, YE 4.5 g/L, NaCl 11.4 g/L) Culture condition (30℃, initial pH 7.0(final pH 8.2)) | 781 |
| | Fermentation | Optimized medium and culture condition (with temp & pH controlled, 20h culture) | 920 |

# Chapter IV

# Mass production of indigo in two types of fermentation systems using recombinant *E. coli* harboring *fmo* gene

This chapter is published in Process Biochemistry 46 (3), 788-791 (2011)

# 1. Introduction

Indigo is one of the oldest blue dyes used by man. Microbially-produced indigo can be exploited as an alternative to chemical synthesis. Although several studies on the biological production of indigo from indole using recombinant microorganisms expressing mono- or di-oxygenase have been reported [Choi, 2003; Doukyu, 2002; Ensley, 1983; McClay, 2005; Murdock, 1993; Rui, 2005], no attempt has yet been made for large scale and continuous production of indigo utilizing recombinant microorganisms. In previous studies, tryptophan was used as a substrate for indigo production using recombinant *E. coli* DH5α cells harboring a novel flavin-containing monooxygenase gene. In these cells, tryptophan was converted to indole by the *E. coli* tryptophanase, then indole was oxidized to 2-hydroxyindole, 3-hydroxyindole, and isatin by FMO. The indole derivatives were then dimerized to indigo and indirubin, depending on the oxygen concentration [Choi, 2003; Eaton, 1995; Maugard, 2002].

In this study, two systems for industrial bio-indigo production, 3,000-L batch and 5-L continuous fermentations were demonstrated, and different culture conditions for process optimization were investigated. This is the first reported work on a large scale or continuous production system for indigo using recombinant *E. coli* harboring a novel *fmo* gene. Important factors such as product toxicity, plasmid stability, medium oxygen concentration and the volumetric productivity of indigo were also described.

## 2. Materials and methods

### 2-1. Strains, media and fermenters

Recombinant *E. coli* strain DH5α, harboring a pBlue 1.7 construct with a novel bacterial *fmo* gene derived from *M. aminisulfidivorans* $MP^T$ ($KCTC\ 12909^T$ = $JCM14647^T$), was used in this study. All chemicals were of analytical grade and commercially available. All fermentations were performed in tryptophan medium(tryptophan 2 g/L, yeast extract 5 g/L, sodium chloride 10 g/L). The recombinant bacteria cells were grown on LB agar plates at 37℃ for overnight. A single colony was inoculated in 100 mL of tryptophan medium (seed culture) and incubated at 30℃ for 16h. The seed culture was inoculated in a 10 L fermenter (Bioflow 3000 model, New Brunswick) with 2L working volume. The cultured cells were transferred into a 100 L fermenter (tank height/tank diameter = 900/500 mm, impeller diameter = 300 mm, speed range 20 - 300 rpm, Rushton-turbine impeller, KSB 6231 model, KoBioTech) for the 3,000 L commercial batch culture seed (tank height/tank diameter = 3780/1800 mm, speed range 20 – 300 rpm, top-driven type, 6-bladed disc turbine impeller) using a made-to-order fermenter (Best Korea).

### 2-2. Batch and repeated-batch fermentation

Large-scale batch fermentation was carried out in a 3,000 L commercial reactor. The culture prepared in the 100 L bioreactor was transferred into the 3,000 L bioreactor containing 2000 L of fresh tryptophan medium. The basic operation conditions were derived from the results of batch fermentation experiments, which carried out at a 100 L fermenter by varying the culture conditions of temperature, pH,

agitation speed (100 to 250 rpm), and aeration (1 to 3 vvm). Based on the initial results, the final batch fermentation was performed for 24 h in a 3,000 L fermenter. After 24 h of fermentation, 95% of the culture fluid was drained from the reactor, and the same volume of freshly sterilized tryptophan medium was supplied to continue the repeated batch fermentation.

## 2-3. Continuous fermentation

Continuous fermentation was performed in a 5 L fermenter (KF-5L, KoBioTech) with an initial 2 L volume of sterile tryptophan medium. The indigo production under different fermentation conditions, such as temperature (30 to 37℃), agitation speed (100 to 300 rpm), dilution rate (D), and tryptophan concentration (2 to 4 g/L), were determined during operation. The continuous culture was started at the time when the indigo concentration in the medium reached 800 mg/L. Then, the fresh tryptophan medium was added to the fermenter at different dilution rates and tryptophan concentrations, and the product was drained into a separate vessel connected to the main instrument.

## 2-4. Analytical methods

*Indigo concentration*: Indigo was precipitated by centrifuging the culture broth at 10,000 × *g* for 1 min, washed at least five times with distilled water, and dried under a vacuum for 6 h. The amount and purity of indigo was estimated as described previous study [Han, 2008; Han, 2011].

*Tryptophan consumption*: Fermentation broth (10 mL) was centrifuged for 10

min at 10,000×*g*, and the supernatant was filtered through a 0.45 μm membrane filter (Millipore). The filtered sample was subjected to high performance liquid chromatography (HPLC) (Agilent 1200 HPLC) to measure the tryptophan concentration. The HPLC operation conditions were as follows: Symmetry ODS $C_{18}$ column (Agilent); solvent mixture containing 0.1% aqueous trifluoroacetic acid and acetonitrile, 90:10(v/v); flow rate, 1.0 mL/min; columntemperature, 25℃; and UV detection, 340 nm [2]. The indigo production yield (Y%) was calculated using the following equation:

$$Y(\%) = I/Ti\text{-}Tr \times 100,$$

where *I* represents the total amount of indigo produced (g/L), and *Ti* and *Tr* are the initial and residual tryptophan concentrations (g/L) in the medium, respectively.

*Indigo toxicity:* Exponentially growing *E. coli* cells in 50 mL of tryptophan medium were treated with different concentrations of purified indigo (0.5 - 2.0 g/L) and incubated at 37℃. The survival rate of the recombinant *E. coli* strain was immediately evaluated by plating the samples onto LB agar plates and counting the colony-forming units (CFUs).

*Plasmid stability:* The stability of the pBlue 1.7 construct in the recombinant *E. coli* DH5α cells during fermentation was determined [Friehs, 2004]. After dilution, the cells were spread on nonselective (ampicillin-negative) and selective (ampicillin-positive) LB agar plates, and incubated at 37℃ for 14 to 16 h. The percentage of plasmid containing cells, P(+), was estimated by calculating the number of colonies formed on the selective plate divided by the number of colonies formed on the nonselective plate and multiplying by 100.

# 3. Results

## 3-1. Effect of aeration and agitation rate on the profiles of dissolved oxygen, cell growth and indigo production

As described in the introduction, oxygen concentration in the medium is an important factor for indigo production [Eaton, 1995; Ensley, 1983; Maugard, 2002]. Fig. 4-1(A) and (B) shows the growth-dependent dissolved oxygen concentrations in the batch condition. To study the effect of aeration rate on cell growth and indigo production, three different amounts of air was supplied to the 100 L fermenter. When 1 vvm of air was supplied to the culture medium, the oxygen concentration decreased significantly and was kept at a low level, based on cell growth. When 2 vvm of air was supplied, the dissolved oxygen also decreased significantly but stayed at a higher level than that of 1 vvm. Thus a higher cell number and more indigo could be obtained. However, when 3 vvm of air was fed into the medium, the oxygen concentration decreased sharply until 12 h of cultivation and then recovered to a saturated level thereafter. The fact that the total viable cell number was decreased after 18 h of cultivation indicates that the energy source was depleted in the medium. However, a maximum of 910 mg/L of indigo was obtained under this condition (Fig. 4-2).

The effects of both aeration and agitation rates on indigo production were also investigated. As the rates increased, indigo production increased as follows: 640 ± 40 mg/L at 100 rpm and 1 vvm, 870 ± 50 mg/L at 150 rpm and 2 vvm, and 910 ± 50 mg/L at 200 rpm and 3 vvm. These results imply that the oxygen was used to not

only grow cells but also supply enough oxygen for the conversion of indole to indigo [Eaton, 1995; Maugard, 2002]. Therefore, 200 rpm and 3 vvm of air were supplied to the culture medium in all subsequent fermentation procedures.

A 100 L fermentation was performed with a 70 L working volume of tryptophan medium at 30℃, pH 7.0, an agitation speed of 200 rpm, and an aeration rate of 3 vvm. As shown in Fig. 4-3, indigo production increased as the tryptophan concentration in the medium decreased. Although a slight decrease in the tryptophan consumption rate was observed at around 12 h of cultivation, and matched the oxygen concentration in medium, maximum indigo production (912 ± 50 mg/L) and cell number were obtained after 18 h incubation. Afterwards, no further increases in indigo production were observed and the number of viable cell significantly decreased, probably due to the high concentration of indigo in the medium.

[A]

[B]

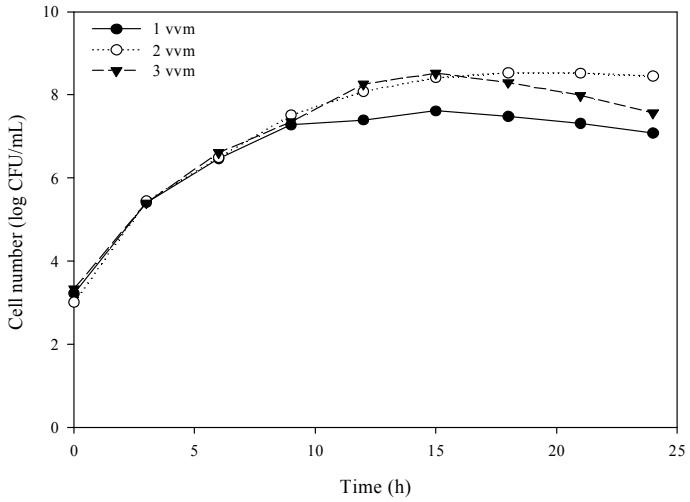



Fig. 4-1. The effect of aeration rate on the profiles of dissolved oxygen. Profiles of oxygen concentration (A) and cell number (B) under different air supply conditions. Symbols: (●) 1 vvm, (○) 2 vvm, and (▼) 3 vvm.

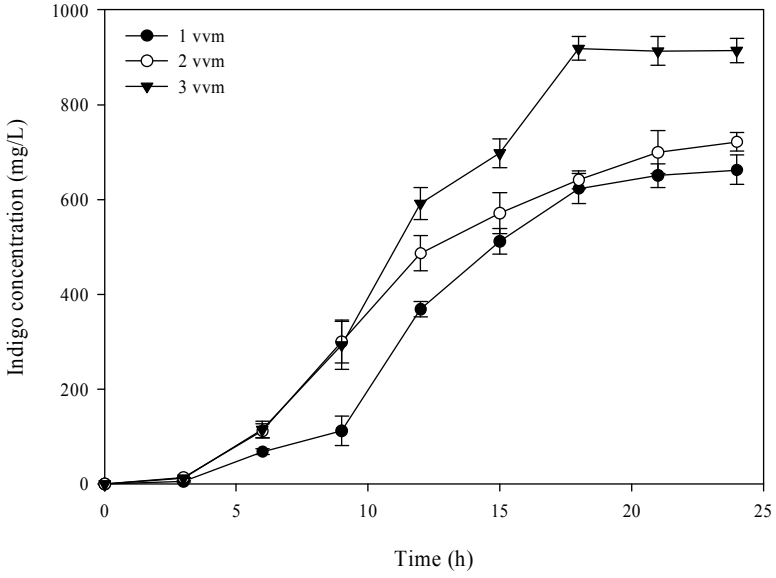



Fig. 4-2. The effect of aeration rate on the indigo production under different air supply conditions. Symbols: (●) 1 vvm, (○) 2 vvm, and (▼) 3 vvm.

Fig. 4-3. The indigo production profile during batch fermentation in a 100 L fermenter. Symbols: (●) indigo concentration, (□) tryptophan concentration, and (△) plasmid stability

## 3-2. 3,000-L batch and repeated-batch fermentation

A pilot-scale batch fermentation was carried out in a 3,000 L commercial fermenter with a working volume of 2,000 L. At this volume, under the conditions described above, 911 ± 22 mg/L of indigo was produced during the initial 24 h of fermentation (Fig. 4-4). Repeated batch fermentations (5 fermentations, a total of 120 h) were also carried out in a 3,000 L fermenter to increase the volumetric productivity per fermenter (Fig. 4-4). After each 24 h cultivation, 95% of the culture fluid was drained from the fermenter, and the same amount of fresh tryptophan medium was added to the reactor to repeat cultivation. Interestingly, the outputs of second and third batches decreased to 870 ± 48 and 702 ± 66 mg/L, respectively, and the fifth fermentation only produced 240 ± 60 mg/L. Given the operational stability of indigo production, the repeated batch fermentations should be carried out in fewer than three runs.

Indigo purity is a very important factor for industrial applications. Traditionally, indigo was extracted from plants by calcium oxide treatment process. However, this traditional purification method resulted in low efficiency and bacterial contamination in our system. To achieve a higher purity of indigo, total fermentation broth was initially heated to 60℃ for 45 min to kill the recombinant cells. The culture broth was then centrifuged at 2,000×*g* and the precipitate was washed with distilled water at least five times to remove the cell debris and residue from the medium. Determination of the indigo production yield (Y%) was performed in a 3,000 L batch culture. Totally 97% of the initial tryptophan (4 kg) was consumed within 24 h, and 1.82 kg of indigo was produced (Fig. 15). Therefore, the production yield was estimated to be 46.9%. Berry *et al*. produced 23.0 g/L of indigo from 200 g/L of

glucose using the naphthalene dioxygenase system [Eaton, 1995]. In this case, the volumetric productivity could not be calculated because the experiment was done in fed-batch mode.

Fig. 4-4. Repeated batch fermentation for indigo production in a 3,000 L reactor. Symbols: (●) indigo concentration, (△) plasmid stability, and (↓) the time that fresh medium was fed into the reactor for the next batch fermentation.

## 3-3. Continuous fermentation

The dilution rate (D) and indigo concentration were important factors for indigo accumulation, where accumulation of indigo at more than 900 mg/L inhibited the cell growth. When the concentration of indigo reached 800 mg/L in batch mode, 4-L of fresh medium containing different concentrations of tryptophan (2, 3 or 4 g/L) was added to the fermenter for 24 h at a constant feeding rate of 0.167 L/h. The initial operation conditions were as follows: temperature 30℃, pH 7.0, aeration rate 3 vvm, agitation speed 200 rpm, and dilution rate 0.084. However, when the indigo concentration in the culture medium reached 800 mg/L, the operating temperature was changed to 35℃ to increase productivity. Under these conditions, a maximum of 937.4 ± 24 mg/L of indigo was produced in 24 h of fermentation, and the final accumulated indigo and volumetric productivity were 23 g and 11.3 mg/Lh, respectively, after 110 h of fermentation (Fig. 4-5). To the best of our knowledge, these are among the highest values ever reported (Table 4-1).

No significant mortality was observed after 5 to 60 min of incubation in the medium containing 0.5 to 1.0 g/L of previously purified indigo. In the presence of 1.5 g/L indigo in the medium, cell growth was inhibited as a function of incubation time (e.g., approximately 80% growth inhibition was observed after 60 min of incubation). When the cells were incubated with 2.0 g/L of indigo, only 15% of the initial cell survived after 40 min, and 100% mortality was observed after 60 min incubation (Fig. 4-6).

Fig. 4-5. Indigo production by continuous fermentation in a 5-L fermenter. When the indigo concentration reached 800 mg/L, 2.8 mL/min of fresh medium containing 3 g/L tryptophan was added. Symbols: (●) indigo concentration, (▲) total indigo accumulation over 140 h, and (△) plasmid stability of recombinant *E. coli* DH5α.

Fig. 4-6. Inhibitory effect of accumulated indigo on the cell growth during cultivation.

## 3-4. Plasmid stability

In the batch culture, the percentage of plasmid-containing P(+) cells was counted every 3 h. As shown in Fig. 2, a slight reduction in the P(+) cell percentage was observed after 24 h in a 100-L fermenter. However, a dramatic decrease in the P(+) cell percentage was observed after repeated batch fermentation in a 3,000-L fermenter (Fig. 4-4). Although plasmid stability was kept constant at 95% until 42 h of cultivation, plasmid stability decreased significantly each time the culture medium was replaced with fresh medium. A sudden decrease in the P(+) cell percentage was observed between 18 and 24h after each medium replacement and start of cultivation, which is the time when the maximum amount of indigo accumulated in the medium. In the continuous culture, plasmid stability was maintained up to 110 h, and indigo was continuously produced (Fig. 4-5). As the number of P(+) cells decreased after 110 h of incubation, indigo production also decreased significantly. The reason why plasmid stability decreased seems to be that the cells without the plasmid have a growth advantage, while the cells with the plasmid grow poorly due to the indigo inside the cell. The plasmid stability data gave us several insights regarding how to best choose the appropriate fermentation mode for indigo production. For indigo productivity, we recommend a continuous culture system, because this system exhibited a high level of indigo productivity and long- term plasmid stability.

# 4. Discussion

In this study, we demonstrated two types of indigo production systems using a recombinant *E. coli* strain harboring a novel *fmo* gene derived from *M. aminisulfidivorans* $MP^T$, which can be used commercially to produce indigo dye. Recombinant microbial indigo production could serve as an alternative to synthetic and plant derived indigo. Several investigators have reported on the microbial production of indigo using monooxygenases or dioxygenases (Table 4-1). There is an important correlation between oxygenase and indigo production in fed-batch type fermentation, with a naphthalene dioxygenase system producing a high concentration of indigo. Glucose is the most economical substrate choice; however, the metabolic pathway for the production of indigo from glucose is very complicated. Berry *et al*. produced 23.0 g/L of indigo from 200 g/L glucose using the naphthalene dioxygenase system [Berry, 2002]. However, the volumetric productivity could not be calculated because the experiment was done by fed-batch mode. While, we achieved 23.1 g/L of indigo concentration and 233.3 mg/L/h of volumetric productivity from 2 g/L tryptophan which were the first and the highest values ever reported.

Generally mutants tend to revert to the wild-type; therefore, stability of the strain is very important. As our FMO system is very simple when compared to glucose/tryptophan utilizing system described by Berry *et al*., plasmid stability of the strain is very high. Although our FMO system is unique, more optimization is needed to obtain the most efficient conversion of the high substrate concentration of tryptophan.

Table 4-1 Summary of indigo production from different strains and substrates

| No | Enzyme/Host strain | Substrate | Indigo concentration (scale/type) | Reference |
|---|---|---|---|---|
| 1 | Naphthalene dioxygenase/*E. coli* | Tryptophan | 25 mg/L (500 mL/batch) | Ensley, 1983 |
| 2 | Naphthalene dioxygenase/*E. coli* | Glucose | 135 mg/L (500 mL/batch) | Murdock, 1993 |
| 3 | Styrene monooxygenase/*P. putida* S12 | Styrene/Indole | 17 mg/L (500 mL/batch) | O'Connor, 1997 |
| 4 | Unknown/*Acinetobacter* sp. ST-550 | Indole | 292 mg/L (500 mL/batch) | McClay, 2005 |
| 5 | Naphthalene dioxygenase/*E. coli* | Glucose/Tryptophan | 23.0 g/L ( 6 l/fed-batch) | Berry, 2002 |
| 6 | Toluene *ortho*-monooxygenase/*E. coli* | Indole | 69.9 mg/L (500 mL/batch) | Rui, 2005 |
| 7 | Cytochrome P450 BM3 & glucose dehydrogenase/*E. coli* | Indole | 759.8 mg/L (500 mL/batch) | Yan, 2007 |
| 8 | Flavin-containing monooxygenase/*E. coli* | Tryptophan | 911±22 mg/L (3,000L/batch)<br>23.1 g/L (5 L/continuous, productivity was 233.3 mg/Lh) | This study |

In this system, we inserted only a single gene (*fmo*) into *E. coli* DH5α strain and checked the plasmid stability. we observed a slight decrease in plasmid stability in batch mode and a significant loss of plasmid stability after 48 h of repeated batch fermentation. However, the plasmid stability was maintained over 110 h in a continuous system when the dilution rate was 0.084 $h^{-1}$. The plasmid stability data gave us several insights for selection of the best fermentation mode for indigo production. We recommend a continuous culture system because this system resulted in high levels of indigo productivity and long-term plasmid stability. Here we demonstrated a valuable recombinant *E. coli* system with a simple conversion of tryptophan to indigo and no need to addition of isopropyl-D-1-thiogalactopyranoside (IPTG) as an inducer. In addition, the factors determining indigo production rates in the 3,000 L batch and 5 L continuous fermentation processes were also investigated. Under optimal conditions, the batch process in a 3,000 L fermenter produced a 46.9% conversion rate. These results suggest that our indigo fermentation system, including medium composition, recombinant cell type, and culture conditions, was very effective for industrial application. During the indigo production process, high concentrations of accumulated indigo in the batch process inhibits the cell growth and indigo production. Similarly, in the repeated batch fermentation process, the recombinant cells starts to loose their plasmid from the $2^{nd}$ batch cycle. The indigo production rate is directly proportional to the plasmid stability in the recombinant cells. So, there are 20% of plasmid stability and 20% indigo production rate was decreased. From the results, it was reaveled that the recombinant cells release their plasmids to survive in highly toxic concentrations of indigo. To overcome the indigo toxicity, an alternative continuous fermentation process was

designed and successfully produced 937.4 ± 24 mg/L indigo over 110h of fermentation. The production yield was estimated to be 233.3 mg/L.h. Despite the sudden drop of P(+) cell number after 110 h, plasmid stability was maintained during the fermentation in the absence of antibiotic markers. In previous studies, plasmid stability in recombinant strains was attained only when IPTG or antibiotics were present in the cultures [Chen, 2008; Royo, 2005]. Our batch and continuous culture process produced significant amounts of indigo during 24 and 110 h fermentations, respectively, relative to other studies. For an industrially available bioproduct, the concentrations of indigo obtained are so high that the proposed production system ensures industrial viability. In conclusion, these indigo production systems could be used for industrial applications to overcome the environmental problems associated with synthetic indigo production and to meet the demand for natural indigo in the dye and textile industries.

# Chapter V

# Production of indirubin using recombinant *E. coli* harboring *fmo* gene

This chapter is submitted to Biotechnology and Bioengineering

# 1. Introduction

Indigoid compounds such as indirubin and indigo have been employed as natural drugs and dyes since ancient times [Kohda, 1990; Xia, 1992]. Traditionally, Danggui Longhui wan, which contains indirubin, has been used for the treatment of numerous chronic diseases including chronic granulocytic leukemia [Han, 2008; Nam, 2005; Tang, 1992]. Recently, indirubin was found to be a potent inhibitor of cyclin dependent kinases (CDKs) and glycogen synthase kinase-3β (GSK-3β), which suggests that indirubin may play an important role in the treatments of leukemia and Alzheimer's disease [Hoessel, 1999; Leclerc, 2001].

In general, indirubin and indigo compounds are extracted from the plant cell cultures of Indigofera tinctoria, Isatis tinctoria, Polygonum tinctorium, and Lonchocarpus cyanescens [Ensley, 1983; Fitzhugh, 1997; Seldes, 1999]. Recently, several investigations have reported that indirubin and indigo can be produced from recombinant microorganisms expressing oxygenase [Doukyu, 2002; Ensley, 1983; Madsen, 1988; McClay, 2005; Murdock, 1993; Rui, 2005]. Our previous studies showed that recombinant *E. coli* DH5α cells harboring a flavin-containing monooxygenase (FMO) gene from *Methylophaga aminisulfidivorans* $MP^T$ [Choi, 2003] produced indirubin (≤5 mg/L) and indigo (920 mg/L) in a 5 L fermenter containing tryptophan medium [Han, 2008; Han, 2011].

FMOs belong to a family of FAD, NADPH and molecular oxygen-dependent enzymes. FMOs are involved in a wide range of oxidative biological processes, including drug detoxification and the biodegradation of aromatic compounds by the

catalyzation of the oxygenation of many nitrogen-, sulfur-, phosphorous-, selenium- and other nucleophilic heteroatom containing chemicals and drugs [Krueger, 2005]. In addition to cytochrome P450s (CYPs), FMOs are considered to be important monooxygenase enzymes for metabolism in both prokaryotes and eukaryotes [Zhou, 2006].

A modified synthetic pathway for several indigoid compounds in the recombinant *E. coli* containing FMO is represented in Chapter III Fig. 3-1. In a tryptophan rich condition, extracellular tryptophan is transported into the cell by tryptophan permease and thereafter converted into indole, pyruvate, and ammonia by tryptophanase (TnaA; EC 4.1.99.1) [Newton, 1965]. FMO catalyzes the hydroxylation of indole to 2-hydroxyindole, 3-hydroxyindole and isatin by using the reducing power of NADPH in the presence of oxygen [Choi, 2003; Eaton, 1995; Maugard, 2002]. In the modified synthetic pathway, indigo is produced from the combination of two 3-hydroxyindole molecules, whereas indirubin is made from the dimerization of 2-hydroxyindole and 3-hydroxyindole [Berry, 2002; Cho, 2011; Choi, 2003; Eaton, 1995; Maugard, 2002; McClay, 2005]. Currently, it is thought that the disproportional production of indigo and indirubin is due to random combinations of the various hydroxyindole molecules under different conditions, but any enzyme activity in the combination reaction has not been discovered yet.

In this study, to enhanced production of indirubin through the supplementation of tryptophan medium with cysteine for the growth of recombinant *E. coli* DH5α cells harboring the *fmo* gene. In addition, the effects of cysteine on the growth of *E. coli* and the FMO activity were investigated. Finally, the optimal conditions for maximum

indirubin production are described. This is the first report on a mass production for indirubin using recombinant *E. coli* pBlue1.7.

## 2. Materials and methods

### 2-1. Media and chemicals

Tryptophan medium (0.2% tryptophan, 0.5% yeast extract, 1.0% sodium chloride, w/v, and 50 μg/ml ampicillin) was used for the production of indigoid derivatives. Dimethyl sulfoxide (DMSO) was purchased from Sigma-Aldrich. Restriction enzymes, T4 DNA ligase and *Taq* polymerase were purchased from Takara, and indirubin standard were from Biomol (Biomol Research Laboratories, CC-206). All other chemicals used were analytical grade.

### 2-2. Cloning, expression and purification of FMO

The *fmo* from *Methylophaga aminisulfidivorans* $MP^T$ was subcloned into the *Nde*I and *Xho*I restriction sites of the expression vector pET30a(+). The expression of *fmo* was induced by the addition of 0.2 mM isopropyl-D-1-thiogalactopyranoside (IPTG) to the bacterial culture medium and incubation at 18℃ for 20 h. The cells were harvested (10,000×*g*), disrupted by French Press (30,000 psi, French Pressure Cells 40k, Thermo Electron Corporation) and centrifuged at 12,000×*g*. Then, the recombinant FMO protein in the supernatant fraction was purified through Ni-NTA affinity column chromatography (Qiagen) and Superdex G-200 gel filtration column chromatography (GE Healthcare). The purity of the FMO was determined by SDS-PAGE analysis, and the purified protein was concentrated to 45 mg/mL in 40mM Tris-HCl (pH8.0) by centrifugal ultrafiltration (100,000Da, Vivaspin, Satourius Stedim) and stored at –20℃ for further study.

## 2-3. Optimization of medium composition and culture conditions for indirubin production

The indirubin production conditions were optimized using a response surface methodology (RSM) based on the $2^n$ factorial central composite design. The first experiment was based on a $2^3$ factorial central composite experimental plan with three medium components, i.e, cysteine, tryptophan and yeast extract. The range and levels of the experimental variables investigated in this study are given in Table 5-1. A set of 20 experiments was carried out with three variables, and each variable was studied at five levels ($\alpha$=2.0). The second experimental design was a $2^2$ factorial central composite experimental plan with two culture conditions; pH and temperature. A set of 13 experiments was conducted with two variables, and each variable was studied at five levels ($\alpha$=2.0). For statistical calculations including the solutions of the second-order polynomial model equation, the Design Expert (version 7.0, Stat-Ease Inc., Minneapolis) software package was used.

Table 5-1. Experimental design (range, levels and response) to assess the effects of medium composition (Experiment 1), pH and temperature (Experiment 2)

| Variable | Symbol code | Range and levels | | | | |
|---|---|---|---|---|---|---|
| | | -2 | -1 | 0 | 1 | 2 |
| Experimeent 1 medium composition | | | | | | |
| cysteine (mM) | $M_1$ | 1 | 2 | 3 | 4 | 5 |
| tryptophan (g/L) | $M_2$ | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |
| yeast extract (g/L) | $M_3$ | 3 | 4 | 5 | 6 | 7 |
| Experimeent 2 culture condition | | | | | | |
| pH | $C_1$ | 6 | 7 | 8 | 9 | 10 |
| temperature (℃) | $C_2$ | 31 | 33 | 35 | 37 | 39 |

## 2-4. Fermentation

To produce the seed culture of recombinant *E. coli* pBlue1.7, a single colony from tryptophan solid medium was used to inoculate 100 ml tryptophan medium, and incubation was continued for 16 h at 30°C. The fermentation was carried out in a 10 L fermenter (Bioflow 3000 model, New Brunswick, Enfield) with a 5 L working volume of tryptophan medium containing 3 mM cysteine at 35℃, pH 8.0, and at an agitation speed of 180 rpm. To investigate the effect of oxygen on the indirubin production, three different cultivations were conducted with different oxygen levels. For anoxic cultivation, air in the fermenter was replaced with nitrogen gas three times and the cells were cultured without agitation for 70 h. To prepare a sequential partial aerobic and anoxic condition, air (3 vvm) was supplied to the reactor for the initial 12 h cultivation, and then the oxygen in the fermenter was replaced with nitrogen gas three times and the cells were cultured without agitation for 58 h. For aerobic cultivation, air (3 vvm) was supplied continuously to the medium with an agitation speed of 180 rpm during the growth.

## 2-5. Analytical methods

### (1) Enzyme assay and Western blot analysis of bacterial FMO

The FMO activity was determined spectrophotometrically by the use of trimethylamine and thiourea as substrates. The assay mixture (1 mL) contained 0.1 mM EDTA and 0.1 mM NADPH dissolved in 0.1 M Tricine/KOH buffer (pH 8.5), and an aliquot of the purified FMO solution was added to the mixture. The reaction rates of the enzyme was monitored by following NADPH oxidation at 340 nm. Protein

concentrations were determined by the Bradford method [9]. For Western blot analysis, the harvested cells ($1 \times 10^9$) were resuspended in 300$\mu$l SDS gel–loading buffer (50 mM Tris-HCl [pH 6.8], 100 mM dithiothreitol, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 10% (v/v) glycerol, and 8 mM $MgCl_2$) and placed in boiling water for 3min. The resulting crude cell lysate (10μl) was loaded into 8% polyacrylamide gels and electrophoresed, and Western blotting was performed as described in the manual [Sambrook, 2001]. The antibody against FMO was prepared and provided by Abfrontier Co. (South Korea).

## (2) Determination of tryptophan and cysteine concentrations

For the quantitative analyses of tryptophan and cysteine, high performance liquid chromatography (HPLC) with a Pico-tag column (4 μm, 3.9 × 300 mm, Waters) was used. The HPLC pump (Waters 510) was used for the time-dependent mixing of the mobile phases of solvent A (140 mM sodium acetate and 6% acetonitrile) and solvent B (60% acetonitrile). The gradient program consisted of two solvent mixtures: 0 to 9 min, 100% solvent A and 0% solvent B; 9 to 9.2 min, 86% A and 14% B; 9.2 to 17.5 min, 80% A and 20% B; 17.5 to 17.7 min, 54% A and 46% B; 17.7 to 21.0 min, 100% B; 21.0 to 30.0 min, 100% A. The flow rate was 1.0 ml/min. All solvents used for the mobile phase in HPLC were filtered through a 0.45 μm cellulose membrane filter (Life Sciences) and degassed in the ultrasonic bath. The peak responses from each chemical were monitored at 254 nm by a variable wavelength photodiode UV detector (Waters 2487 UV).

### (3) Effect of reducing agents or amino acids on indirubin synthesis

To examine the effect of several reducing agents and amino acids on indirubin synthesis, variable concentrations (0~5mM) of cysteine, methionine, serine, tyrosine, phenyalanine, arginine, aspartic acid, glutathion, dithiothreitol and thioglycolic acid were added to the tryptophan medium.

### (4) Protein preparation and two-dimensional gel electrophoresis

For 2-DE analysis, pH 3–10 IPG strips (Amersham Biosciences) were rehydrated in swelling buffer containing 7 M urea, 2 M thiourea, 0.4% (w/v) DTT, and 4% (w/v) CHAPS. The gel image using a GS-710 imaging calibrated densitometer (Bio-Rad, Hercules, CA, USA). Protein spot detection and 2-D pattern matching were carried out using Image Master$^{TM}$ 2D Platinum software (Amersham Biosciences). Protein bands of interest were excised and digested in-gel with sequencing grade, modified trypsin (Promega). The resulting tryptic peptides were separated and analyzed using reversed phase capillary HPLC directly coupled to a Finnigan LCQ ion trap mass spectrometer (LC-MS/MS). Both of the 0.1 × 20 mm trapping and a 0.075 × 130mm resolving column were packed with Vydac 218MS low trifluoroactic acid $C_{18}$ beads (5um in size, 300Å in pore size; Vydac) and placed in-line. Following the peptides were bound to the trapping column for 10 min at 5% (v/v) aqueous acetonitrile containing 0.1% (v/v) formic acid, then the bound peptides were eluted with a 50-min gradient of 5-80% (v/v) acetonitrile containing 0.1% (v/v) formic acid at a flow rate of 0.2 μℓ/min. For tandem mass spectrometry, a full mass scan range mode was m/z = 450 - 2000Da. After determination of the charge states of an ion on zoom scans,

product ion spectra were acquired in MS/MS mode with relative collision energy of 55%.

The individual spectra from MS/MS were processed using the Turbo SEQUEST software (Thermo Quest). The generated peak list files were used to query either MSDB database or NCBI using the MASCOT program (http://www.matrixscience.com). Modifications of methionine and cysteine, peptide mass tolerance at 2 Da, MS/MS ion mass tolerance at 0.8 Da, allowance of missed cleavage at 2, and charge states (+1, +2, and +3) were taken into account. Only significant hits as defined by MASCOT probability analysis were considered initially.

### (5) Indigoid compounds analyses

The 2-hydroxyindole and 3-hydroxyindole formation were detected by a luminescence spectrometer (excitation at 365 nm and emission at 470 nm, Perkin Elmer LS 45, Perkin Elmer) [Woo, 2000]. The in vitro assay was performed in a 100 mM Tricine-KOH buffer (pH 8.5) containing various concentrations of cysteine, 0.1 mM NADPH, 10 mM indole, and the purified FMO (5 g) at room temperature. For in vivo assay, the concentrations of 2- or 3-hydroxyindole were measured by HPLC (Agilent 1200 HPLC). The compounds were separated on a reverse-phase $C_{18}$ column (5 μm, 4.5 × 150 mm, YMC) and detected with a diode array detector. Solvent A (0.1% formic acid) and solvent B (50% methanol and 50% acetonitrile) were applied as the mobile phase according to the following timetable: 0 to 8 min, 85% solvent A and 15% solvent B, flow rate of 1 ml/min; 8 to 10 min, 65% A and 35% B, flow rate of 2 ml/min; 10 to 15 min, 65% A and 35% B, flow rate of 2 ml/min. For the analysis,

commercially available standards for isatin, 2- and 3- hydroxyindole, and indole were used, and 3-hydroxyindole was prepared by dephosphorylating 3-hydroxyindole phosphate with alkaline phosphatase under anaerobic conditions [Meyer, 2002].

The concentration of indigo and indirubin were determined by HPLC analysis. Each cell culture broth from flasks or fermenters was centrifuged at 10,000×g for 1 min to precipitate a dark pellet that was then washed twice with distilled water. The precipitant was resuspended in 10 mL of DMSO and then subjected to repeated 1 min sonication (Branson) with a microprobe. The amount of indirubin or indigo in the supernatant was determined by comparison of each standard indigoid solution dissolved in DMSO. The HPLC procedure was conducted with a photodiode array detector (Agilent 1200 HPLC, Santa Clara, CA, USA) and an effluent flow rate of 1 mL/min; indirubin sample was monitored at 540 and indigo was at 620 nm, respectively. For the analysis, an isocratic elution system of methanol and water (80:20, v/v) was applied and the YMC-pack ODS-A ($C_{18}$, 250 × 4.6 mm) column was used [Perpete, 2006].

### (6) Electron microscopic observation of indigoid compounds in cells

The samples were taken from two different culture medium. The morphology of indigoid compounds were observed by scanning electron microscopy (SEM) (Hitach, S-4800) and transmission electron microscope (TEM) (JEOL Ltd., JEM-1400).

# 3. Results

## 3-1. Effect of reducing agents on indirubin synthesis

It is known that recombinant *E. coli* harboring oxygenases can produce indigoid compounds from glucose or indole. Glucose is readily available, but its metabolic pathway is so complicated that the efficiency of the indigoid compound production is low [Berry, 2002]. Indole (>1-6 mM) inhibits the growth of *E. coli* [Chant, 2007], thus the indole concentration in the culture medium must be maintained at a low level to avoid toxicity. Therefore, tryptophan was used in this study as an alternative substrate for indirubin synthesis to overcome the aforementioned problems.

In batch fermentation, it was found that recombinant *E. coli* pBlue 1.7 cells harboring the *fmo* gene produced indigo and indirubin by utilizing tryptophan as a substrate under high oxygen level (3 vvm of air supply) [Han, 2008; Han, 2011]. Under this condition, the major product (>95%) was indigo and a small amount of indirubin was produced. Interestingly, when the indigo fermentation was performed in an aerobic condition for 12 h to cultivate the cells and then an anoxic condition, it was observed that some of the blue indigo compound that was produced in the medium turned into pink indirubin. Based on the above result, we expected the possibility that the addition of any reducing agents, including amino acids, could increase the indirubin synthesis in the recombinant *E. coli*. Among the various chemicals tested, only cysteine shows enhanced indirubin production (Fig. 5-1).

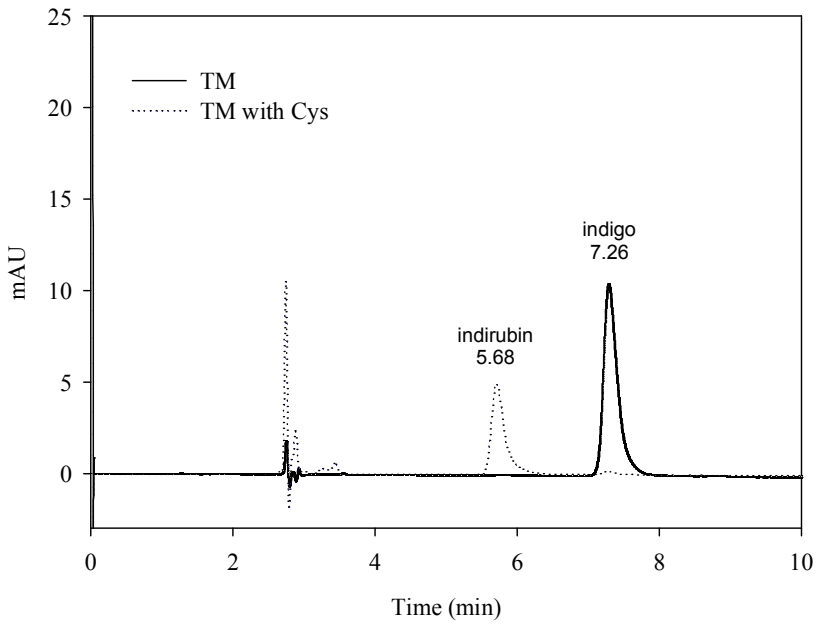



Fig. 5-1. Analysis of indigo and indirubin by HPLC. The control condition was produced indigo (major) and indirubin (minor) in tryptophan medium, however cysteine contributed to increase of indirubin production and decrease of indigo production in tryptophan medium containing cysteine.

## 3-2. Effect of cysteine on the growth and *fmo* expression in the recombinant *E. coli*

To investigate whether cysteine supplementation influences the growth of recombinant *E. coli* pBlue1.7, the cell numbers of the control *E. coli* DH5α and the recombinant *E. coli* pBlue1.7 at different growth phases were compared. For the control strain, the *E. coli* harboring pBSK that contains no inserted gene was used. Fig. 5-2 showed that addition of cysteine inhibited the growth of both the control and the recombinant strains. The cell numbers of the recombinant *E. coli* pBlue 1.7 cultured in tryptophan media with (3 mM) and without cysteine after 15 h cultivation at 35°C were $(4.11 \pm 1.25) \times 10^7$ and $(4.05 \pm 1.58) \times 10^8$, respectively. The numbers of cells of the control *E. coli* strains with and without cysteine at the same incubation time were $(4.12 \pm 1.16) \times 10^7$ and $(3.28 \pm 1.31) \times 10^8$, respectively. These experiments showed that cysteine significantly inhibited the growth of *E. coli*.

The Western blot analysis was performed to investigate the effect of cysteine on the expression of *fmo*. A comparison of the band intensities from the samples indicated that the expression level of FMO was not significantly reduced by the supplemented cysteine in the recombinant *E. coli*. The enzyme assays with purified FMO also showed that cysteine did not change the activity of FMO, but its higher concentration inhibited indigo production (Fig. 5-3). From these results, it was concluded that FMO is not involved in the enhanced indirubin production of recombinant *E. coli*. In addition, we observed that the minimum concentration of cysteine required for indirubin production was 0.5 mM in the tryptophan medium.

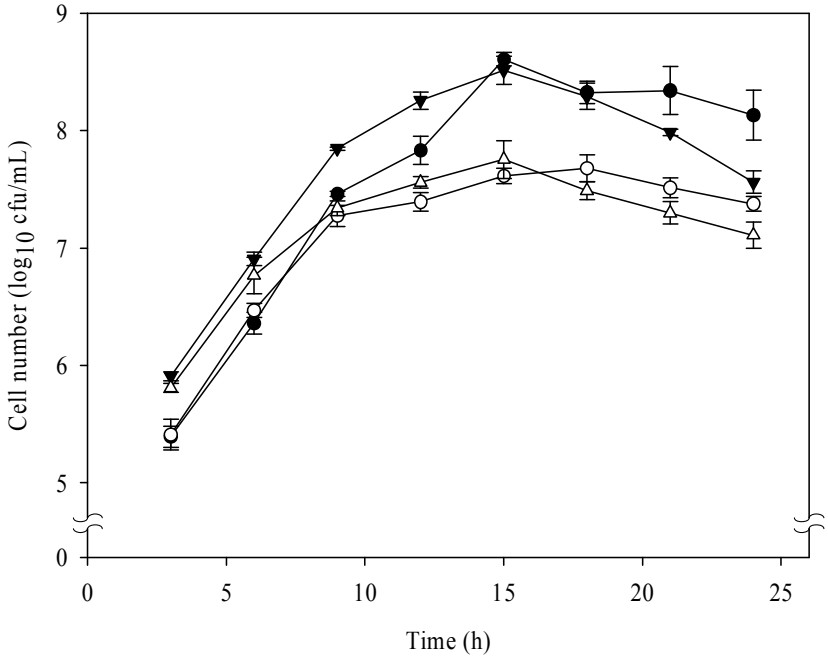



Fig. 5-2. Effect of cysteine on the growth of different types of recombinant *E. coli*. symbol: (●), *E. coli* pBSK in the tryptophan medium. (▼), *E. coli* pBlue 1.7 in tryptophan medium; (○), *E. coli* pBSK in tryptophan plus cysteine medium; (△), *E. coli* pBlue 1.7 in tryptophan plus cysteine medium.

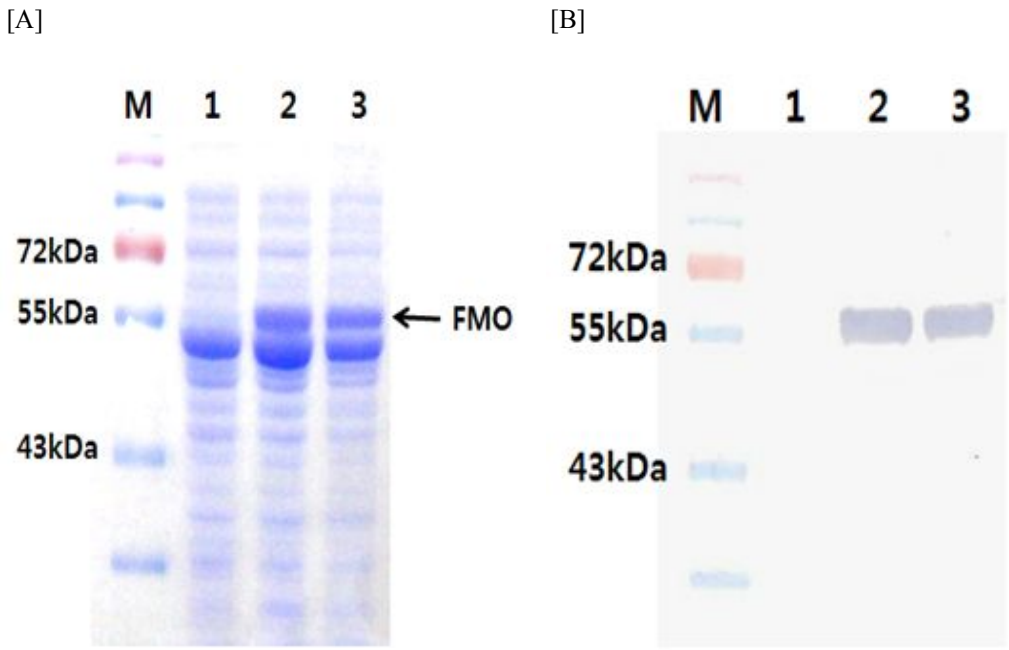



Fig. 5-3. Comparison of SDS-PAGE (A) and western blot (B) analysis of *E. coli* pBlue1.7 grown on different medium composition. Symbol: M, molecular weight marker; lane 1, *E. coli* DH5α cells containing no plasmid; lane 2, *E. coli* cells harboring pBlue 1.7 grown on tryptophan medium; lane 3, *E. coli* cells harboring pBlue 1.7 grown on tryptophan plus 3 mM cysteine medium.

## 3-3. Comparative protein expression pattern in *E. coli*

The 2-DE patterns confirm the prediction that the vast majority of proteins have their counterparts in all strains. However, between these strains, there are clear differences in spot intensity, presence or absence and position of the spots. We concentrated our evaluation on readily detectable spot variations, which were consistent in all 2-DE patterns. From over 1,000 spots on the 2D gels (Fig. 5-4), we identified 26 total proteins by comparing them with the *E. coli* poly acrylamide gel electrophoresis database (MSDB database, http://www.matrixscience.com) or by MALDI-TOF mass spectrometry analysis (Table 5-2). Although there were more than 22 protein spots expressed at lower levels in the tryptophan medium cultured *E. coli* than in the tryptophan medium containing 3 mM cysteine cultured *E. coli*. Interestingly, recombinant *E. coli* grown in tryptophan medium containing 3mM cysteine, the expression of tryptophanase (*trp*ABCED) decreased.

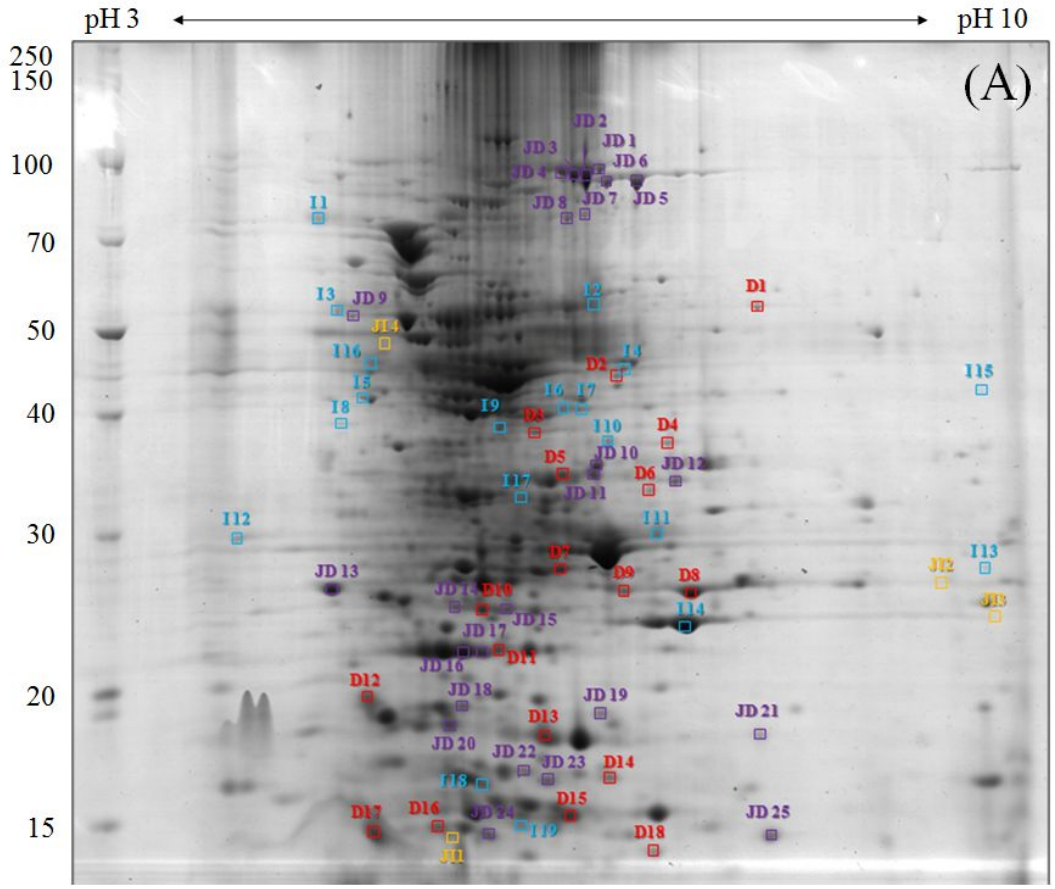



Fig. 5-4. Comparison of 2-D gels containing soluble proteins (800 mg) isolated from two different culture medium such as tryptophan medium (A) and tryptophan medium containing 3 mM cysteine (B). Total soluble proteins (800 mg) separated by their pH3-10 IPG strips and by their molecular weight on 8-16% (v/v) linear gradient SDS-PAGE. Identified protein spots are numbered and their identity is listed in Table 8. I increase, D decrease, JI just increase, JD just decrease.

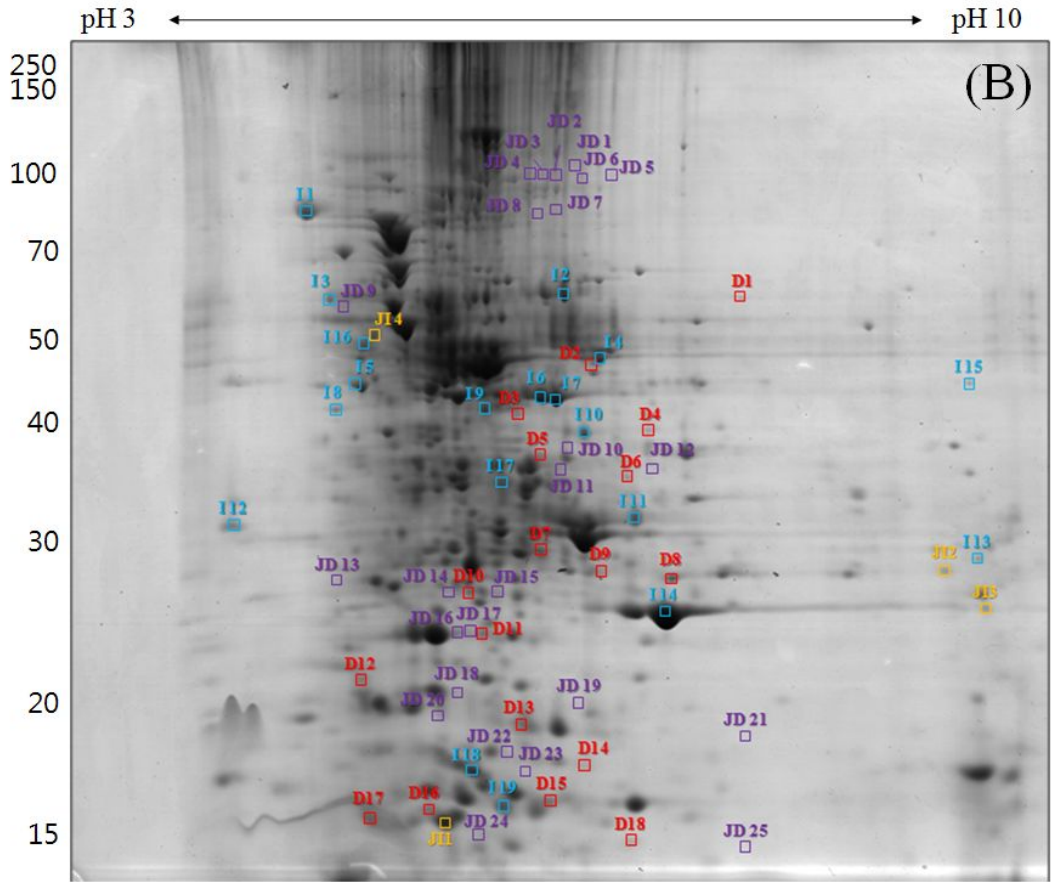



Fig. 5-4. Continuous.

Table 5-2. Summary of MS/MS data for protein spots showing expression levels on 2-DE gels for *E. coli* grown on two different media

| Spot NO. | Protein name | Accession number (NCBI) | Calculated pI | Mr/Da | Mascot Score | Sequence coverage (%) |
|---|---|---|---|---|---|---|
| JD1 | tryptophanse | Gi54401932 | 6.00 | 53445 | 95 | 29 |
| JD2 | Putative protease inhibitor | Gi284922219 | 5.91 | 166714 | 67 | 10 |
| JD3 | tryptophanse | Gi147997 | 5.88 | 52699 | 114 | 24 |
| JD4 | tryptophanse | Gi147997 | 5.88 | 52699 | 79 | 21 |
| JD5 | phage tail tape measure protein | Gi 300917761 | 9.05 | 92851 | 97 | 15 |
| JD6 | tryptophanse | Gi147997 | 5.88 | 52699 | 115 | 27 |
| JD7 | tryptophanse | Gi41936 | 5.88 | 52699 | 66 | 17 |
| JD8 | Trigger factor | Gi301049642 | 4.81 | 48088 | 83 | 6 |
| JD9 | cysK protein | Gi145686 | 5.64 | 34496 | 150 | 47 |
| JD10 | Thiosulfate binding protein | Gi209764118 | 7.78 | 37605 | 175 | 52 |
| JD11 | Rhs family protein | Gi2920637 | 6.17 | 157568 | 52 | 10 |
| JD12 | phage integrase family protein | Gi191168581 | 9.85 | 42648 | 137 | 32 |

Table 5-2. continuous

| Spot NO. | Protein name | Accession number (NCBI) | Calculated pI | Mr/Da | Mascot Score | Sequence coverage (%) |
|---|---|---|---|---|---|---|
| JD13 | predicted hydrolase | Gi1128948 | 5.20 | 23086 | 245 | 24 |
| JD14 | Alkyl hydroperoxide reductase C22 protein | Gi170172436 | 4.92 | 19135 | 93 | 60 |
| JD15 | Alkyl hydroperoxide reductase C22 protein | Gi170172436 | 4.92 | 19135 | 105 | 60 |
| JD16 | Alkyl hydroperoxide reductase C22 protein | Gi170172436 | 4.92 | 19135 | 105 | 59 |
| JD17 | cysK protein | Gi145686 | 5.64 | 34496 | 124 | 49 |
| JD18 | ParB-like nuclease | Gi260741955 | 5.11 | 71798 | 122 | 28 |
| JD19 | relaxase | Gi302651678 | 7.09 | 103796 | 65 | 15 |
| JD20 | Initmin THETA-2 | Gi215489834 | 8.73 | 101647 | 91 | 17 |
| JD21 | cellulose synthase operon protein C | Gi284923560 | 5.74 | 125784 | 89 | 16 |
| JD22 | Alkyl hydroperoxide reductase C22 protein | Gi170172436 | 4.92 | 19135 | 85 | 54 |
| JI1 | autonomous glycyl radical cofactor Grc A | Gi15832699 | 5.09 | 14275 | 178 | 76 |
| JI2 | Minor tail protein precursor H | Gi218700210 | 8.96 | 92575 | 69 | 11 |
| JI3 | 50S ribosomal protein L1 | Gi297517724 | 9.27 | 20745 | 94 | 52 |
| JI4 | trigger factor | Gi254160505 | 4.82 | 48144 | 319 | 57 |

## 3-4. Hydroxyindole formation *in vitro* and *in vivo*

A modified synthetic pathway for several indigoid compounds in the recombinant *E. coli* containing FMO is represented in Fig. 3-1. In a tryptophan rich condition, extracellular tryptophan is transported into the cell by tryptophan permease and thereafter converted into indole, pyruvate, and ammonia by tryptophanase [Newton, 1965]. The FMO catalyzes the hydroxylation of indole to 2-hydroxyindole, 3-hydroxyindole and isatin by using the reducing power of NADPH in the presence of oxygen [Eaton, 1995; Maugard, 2002]. In the modified synthetic pathway, indigo is produced from the combination of two 3-hydroxyindole molecules, whereas indirubin is made from the dimerization of 2-hydroxyindole and 3-hydroxyindole [Berry, 2002; Cho, 2011; Eaton, 1995; Maugard, 2002; McClay, 2005]. Currently, it is thought that the disproportional production of indigo and indirubin is due to random combinations of the various hydroxyindole molecules under different conditions, but any enzyme activity in the combination reaction has not been discovered yet.

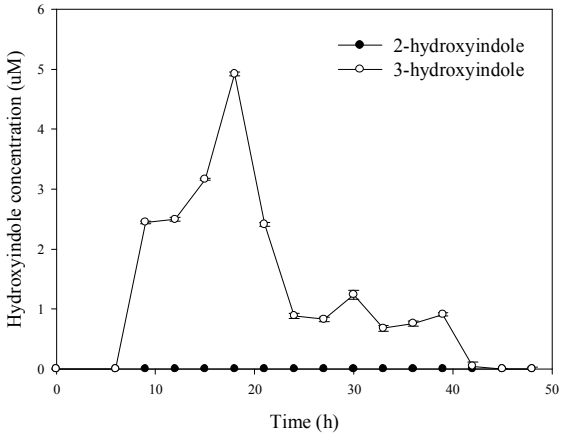
To investigate the modified processes of indigo and indirubin biosynthesis after cysteine supplementation, different hydroxyindole formations were determined by in vitro and in vivo assays. An in vitro assay was conducted in a Tricine-KOH buffer mixture containing FMO protein (5 g) with or without cysteine. When cysteine was not added to the assay mixture, indole was hydroxylated to 3-hydroxyindole by FMO under air.

As shown in Fig. 5-5(A), the concentration of 3-hydroxyindole increased only for the initial 6 min and, thereafter, almost the same concentration was steadily maintained and the color of assay mixture turned blue. From this result, it is assumed that a limited concentration of 3-hydroxyindole continuously dimerized to indigo under air. When

cysteine was added, the 3-hydroxyindole concentration increased continuously for 30 min, but the assay mixture did not change color. This finding indicates that cysteine inhibits the dimerization of 3-hydroxyindole into indigo in the in vitro assay.

The effect of cysteine on the conversion of indole to hydroxylindole in vivo was also investigated. The same numbers ($1.7 \times 10^3$) of recombinant *E. coli* cells were grown on tryptophan medium with or without cysteine, and the concentrations of 2- or 3-hydroxyindole synthesized in the cells were compared. In the cells grown on tryptophan medium without cysteine, 3-hydroxyindole concentration increased up to 18 h, subsequently, a considerable decreased trend was observed (Fig. 5-5(A)). During the incubation period, 2-hydroxyindole was not detected and the indigo concentration was maximal because indigo synthesis is proportional to the concentration of 3-hydroxyindole. However, when cysteine was added to the medium, similar levels of 2- and 3-hydroxyindole were synthesized from indole (Fig. 5-5(B)). As described above, both 2- and 3-hydroxyindole are essential for the dimerization process to produce indirubin. Based on these data, it was concluded that cysteine influences both the hydroxylation level of indole to 2- or 3-hydroxyindole and the synthesis of indirubin in the cell. Therefore, it is possible to produce a high concentration of indirubin in a tryptophan medium by controlling the addition of cysteine.

[A]




[B]

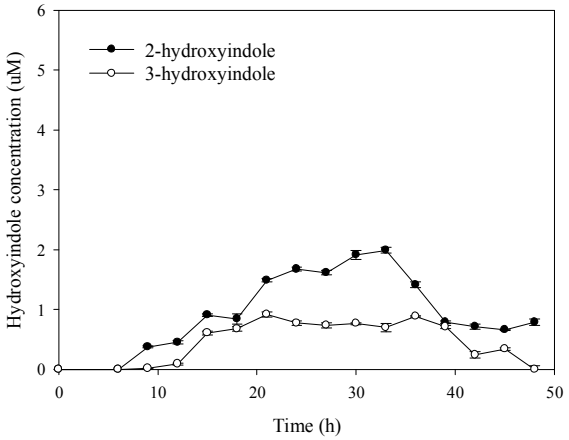



Fig. 5-5. Effects of cysteine on the hydroxylation of indole to 2- and 3-hydroxyindole *in vivo* assays. The time-dependent 2- and 3-hydroxyindole concentrations in the cells grown on tryptophan medium (A) and with cysteine (B) were determined.

## 3-5. Optimization of medium and culture conditions for indirubin production.

Using the RSM, the interaction between two variables and their optimum levels can be easily understood and located. Fig. 5-6(A) shows the interaction between tryptophan and yeast extract, tryptophan and cysteine, and yeast extract and cysteine. Fig. 5-6(B) shows the responses and curves for plots of the pH and temperature. The optimum points for the maximum production of indirubin are near the center point of the pH and temperature ranges. The numerical method given by Myers and Montgomery was used to solve model equations (3) and (4) [50]. The optimal values of the medium composition test variables in the coded units were: $M_1$ (tryptophan) = 0.22, $M_2$ (yeast extract) = – 0.26 and $M_3$ (cysteine) = – 0.10. The natural values were obtained by putting the respective values of $M_i$ into Eqn (3). The optimal values of the culture condition test variables in the coded units were: $C_1$ (pH) = – 0.05 and $C_2$ (temperature) = 0.17. The natural values were obtained by putting the respective values of $C_i$ into Eqn (4).

Eqn (3)

$$\text{Indirubin concentration} = 48.983 + 0.250\ M_1 + 0.626M_2 - 0.623M_3 - 1.571M_{12} - 3.096M_{13} + 1.411M_{23} - 8.863M_1^2 - 2.172M_2^2 - 0.774M_3^2$$

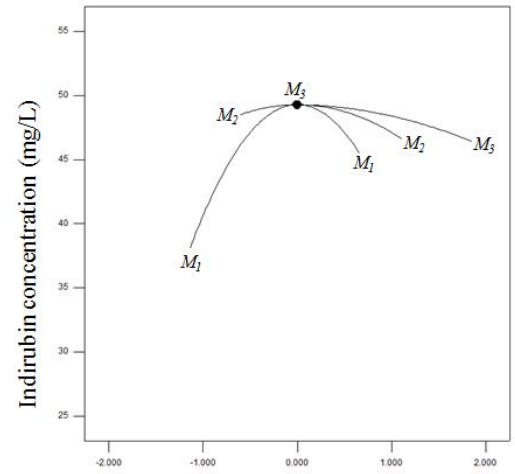
Eqn (4)

$$\text{Indirubin concentration} = 49.07 - 0.24\ C_1 + 0.90C_2 - 0.17C_{12} - 2.74C_1^2 - 2.71C_2^2 - 0.54M_1^2M_2 - 1.25C_1C_2^2 - 1.07C_{12}^2$$

The fit of the models was expressed by the coefficients of determination, $R^2$, which were 0.9464 and 0.9877, respectively. The closer the $R^2$ value is to 1, the stronger the model and the better the prediction of the response. The values of the adjusted determination coefficient (Adj $R^2$) (0.8981 and 0.9631) were also high in reasonable agreement with the Adj $R^2$. In addition, the statistical significance of the second-order model equations was assessed using the *F*-test analysis of variance (ANOVA), which indicates a good agreement between the experimental and predicted values for the medium composition and culture conditions (Table 5-3). The RSM analysis showed that cysteine concentration was a critical factor for indirubin production. The RSM data shown were obtained from the tryptophan medium containing 3 mM cysteine for which the initial pH was adjusted. The analysis determined that tryptophan concentration was another important factor for indirubin production. The RSM result showed that the tryptophan concentration for the highest indirubin production was 2.0 g/L.

In addition, the optimal temperature and initial pH for indirubin production were 35℃ and pH 8.0 according to the analysis. The experiment indicated that even a slight change of the pH greatly affected the indirubin production level. The RSM model equation was confirmed experimentally. The maximum production rate of indirubin (93.1 mg/L) was obtained in modified tryptophan medium, which consists of tryptophan 2 g/L, yeast extract 5 g/L, sodium chloride 10 g/L and 3 mM cysteine in a 300 mL flask. The initial pH and temperature was set to 8.0 and 35℃, respectively at 48 h cultivation, these results shows a closer agreement with the prediction of model.

(A)




(B)

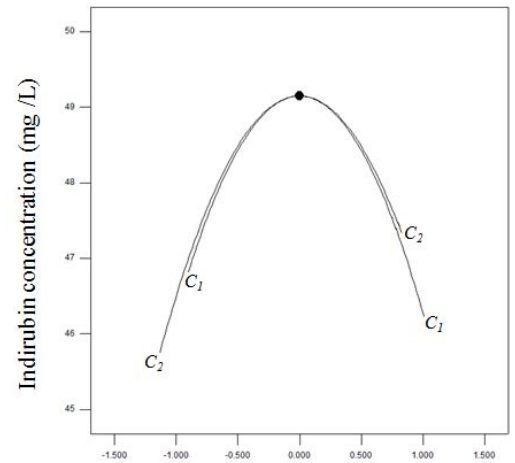



Fig. 5-6. Perturbation plot shows the effect of (A) medium composition (tryptophan, yeast extract, and cysteine) and (B) culture conditions (pH and temperature) on the production of indirubin. The optimal coded units were: $M_1$ (tryptophan) = 0.22, $M_2$ (yeast extract) = −0.26, $M_3$ (cysteine) = −0.10, $C_1$ (pH) = −0.05 and $C_2$ (temperature) = 0.17.

Table 5-3. Analysis of variance with coded values for the effects of medium composition (Experiment 1), pH and temperature (Experiment 2).

| Source | SS | DF | MS | F-value | Prob (P) > F |
|---|---|---|---|---|---|
| Experiment 1 (medium composition) | | | | | |
| Model | 203.08 | 9 | 226.45 | 19.60 | <0.0001 |
| Residual (error) | 115.52 | 10 | 11.55 | | |
| Lack of fit | 62.76 | 5 | 12.55 | 1.19 | 0.4268 |
| Pure error | 52.76 | 5 | 10.55 | | |
| Total | 2153.60 | 10 | | | |
| R-squared | 0.9463 | | | | |
| Adj R-squared | 0.8980 | | | | |
| Experiment 2 (culture condition) | | | | | |
| Model | 287.45 | 8 | 35.93 | 40.15 | <0.0015 |
| Pure error | 3.58 | 4 | 0.89 | | |
| Total | 291.02 | 12 | | | |
| R-squared | 0.9877 | | | | |
| Adj R-squared | 0.9631 | | | | |

SS, sum of squares; DF, degrees of freedom; MS, mean squares

## 3-6. Indirubin fermentation

The cultivation of the recombinant *E. coli* in flasks was performed at optimum conditions (medium containing 2 g/L tryptophan, 5 g/L yeast extract and 3 mM cysteine, pH 8.0, 35℃). The indirubin production rate was obtained as 75.4 mg/L. The indirubin production was also carried out in a 10-L fermenter (Bioflow 3000 model, New Brunswick, NJ, USA) with a 5 L working volume of tryptophan medium containing 3mM cysteine at 35℃, pH 8.0, agitation speed of 200 rpm. To investigate effect of air on the indirubin production, three different culturing conditions were applied for the fermentation. The first culturing condition was no air supply during fermentation. The second culturing condition was air supply (3 vvm) for 12 h and blocking of the supply until the end of fermentation. The last condition was constant air supply throughout the fermentation period. The indirubin production was estimated from the fermentations was 82.5 mg/L (the first condition), 153 mg/L (the second condition) and 223.6 mg/L (the third condition). During the fermentation process at second condition, the indirubin production was decreased after blocking the air supply. In the third condition, the recombinant *E. coli* utilize and convert 1.61g/L of tryptophan into 0.245 g/L of indirubinin by consuming 0.73 mM cysteine in 48 h incubation (Fig. 5-7). These results demonstrates, that the presence of air is a crucial factor for the indirubin production.

Fig. 5-7. Relationship between tryptophan and cysteine consumption and indirubin synthesis during cell growth in a 10 L fermenter with a 3 vvm air supply. Symbols: (●) indirubin; (▼) tryptophan; (■) cysteine.

## 3-7. SEM and TEM analysis for recombinant *E. coli* pBlue1.7

After fermentation, the indirubin was recovered by centrifuge (5,000×g, 10 min). We observed two phase in collection tube that upper phase (white, *E. coli*) and lower phase (dark red, indirubin). The SEM images showed the *E. coli* and crystallized indigo (control) and indirubin (Fig. 5-8). Fig. 5-9 shows the accumulation of synthesized indirubin in the recombinant *E. coli* cells. In the initial growth stage, the synthesized indirubin were distributed throughout the cellular membrane and thereafter it starts slowly accumulated inside the cells. Most of the cells were disrupted due to the indigoid toxicity, and the compounds were secreted and accumulated in the medium. These results indicated that none of the membrane transporters in the *E. coli* strain pumps out the indigoid compounds. The intracellular accumulation of the indigoid crystals might stress the cell, which may be a main reason for the poor growth of the recombinant *E. coli* pBlue1.7. Additionally, a comparison of the growth rates implied that indirubin may be more toxic than indigo to the *E. coli* strain. However, the identification and overexpression of a proper transporter that expels the indigoid compounds into the medium may improve the growth of the recombinant *E. coli* in the future.

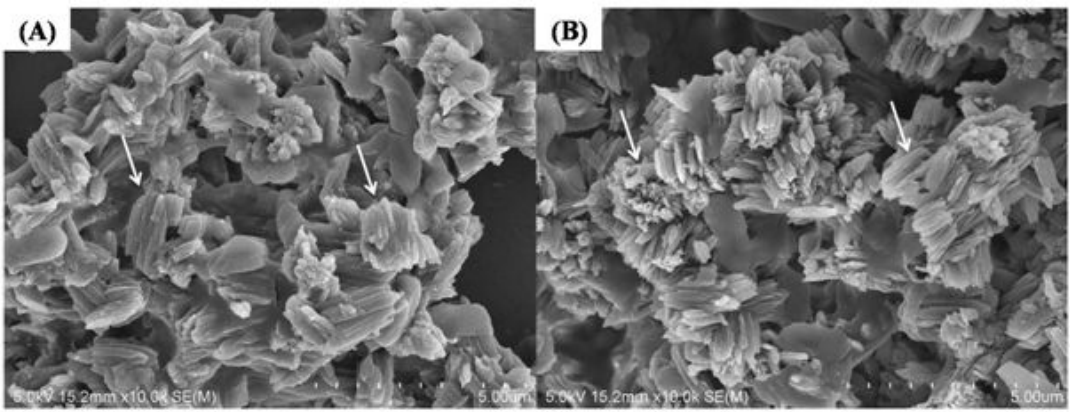



Fig. 5-8. SEM images of indigo and indirubin pigment from *E. coli* pBlue 1.7 on the cultured two different medium (10,000×). (A) Indigo (dark blue), after culture 24h; (B) Indirubin (dark red), after culture 48h. Symbol: crystallization of indigo and indirubin (→). Length scale is indicated by bars.

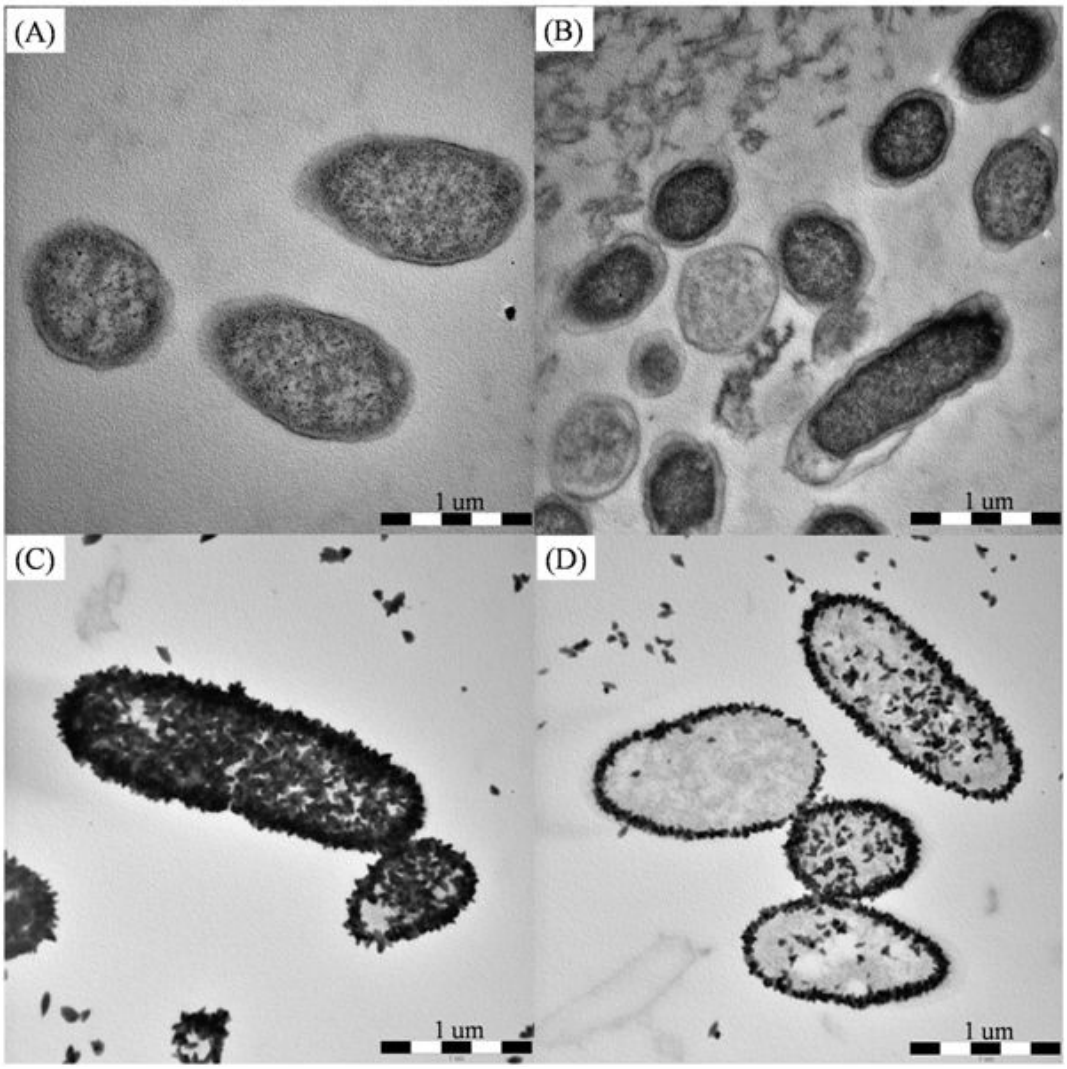



Fig. 5-9. TEM images of the indigo and indirubin formation in *E. coli* pBlue 1.7. (A) Indigo, after culture 12 h ( x 40k), (B) Indigo, after culture 24 h ( x 40k), (C) Indirubin, after culture 24 h ( x 40k) and (D) Indirubin, after culture 48 h ( x 40k). Length scale is indicated by bars.

# 4. Discussion

The recombinant *E. coli* containing the oxygenases can produce indigoid compound by using glucose and indole as substrates. Glucose is readily available, but its metabolic pathway is so complicated [Berry, 2002] and indole (>1-6 mM) inhibits cell growths of *E. coli* [Chant, 2007]. Therefore, the indole concentration in the culture medium must be maintained at low level to avoid the toxicity. In this study, tryptophan was used as substrate for indirubin synthesis to overcome the problems.

The synthesis of indigoid compound pathway by the recombinant *E. coli* containing flavin-containing monooxygenase was described above. In *E. coli*, indole is produced from tryptophan using tryptophan metabolism (tryptophan operon (*trpABCDE*) and operon (*tnaCAB*)) [Jung, 2006; Lee, 2007; Lee, 2011; Yanofsky, 1991]. In tryptophan rich conditions, the extracellular tryptophan was transported into *E. coli* by three permeases (Mtr, TnaB, and AroP). The TnaB permease is critical for tryptophan uptake. Thereafter, tryptophan is converted into indole, pyruvate, and ammonia by TnaA [Newton, 1965]. FMO catalyzes the hydroxylation of indole to produce 2-hydroxyindole, 3-hydroxyindole and isatin by using the reducing power of NADPH in the presence of oxygen. From this pathway, it is predicted that indigo is produced from combination of two 3-hydroxyindoles while indirubin is made from one molecule of 2-hydroxyindole and one molecule of 3-hydroxyindole. Previously, it was found that this biological process produced complex mixture containing more indigo (920 mg/L) than indirubin (5 mg/L) [Han, 2008; Han, 2011].

In this study, indirubin production was greatly enhanced by the addition of cysteine.

The supplementation of cysteine into medium enabled the recombinant *E. coli* to produce more indirubin (93.1 mg/L) than indigo (5 mg/L). Our data confirmed that the addition of cystein did not affect activity and expression level of FMO, but it reduced formation of the indoxyl compounds *in vivo* and *in vitro*. It was assumed that the cysteine is involved in two different reaction mechanisms to enhance indirubin production in the recombinant *E. coli*. Firstly, the cysteine reduces oxidation potential of the conversion of indigo to indirubin by working as reducing agent. In the presence of cystein, the dimerization reaction of 3-hydroxyindole maybe inhibited, thus it will reduce rate of indigo formation. As a result, more indirubin could be produced. Secondly, the cysteine may stimulate production of isatin and 2-hydroxyindole but not 3-hydroxyindole in the indole conversion reaction, which is related to preferential increase of indirubin formation.

The SEM and TEM results showed that the biosynthesis of indigo and indirubin may occur in intracellular of *E. coli*. During the initial production, small particles were formed at intracellular. At the later stage, the particles were increased and the indigoid compounds were accumulated inside the recombinant *E. coli*. Moreover, the *E. coli* cell membrane was destructed by crystallized indigoid compounds deposited throughout the membrane. Therefore, the membrane transport functions were stopped such as indole transporter (AceEF) [Pinero-Fernandez, 2011]. It is presumed that the accumulation of the indigoid compounds explain poor growth of the recombinant *E. coli* pBlue1.7.

In conclusion, the growth of the recombinant *E. coli* was inhibited by cysteine. However, the expression and activity of FMO were not influenced. The optimum culture conditions for indirubin production in tryptophan medium were as follows: 2

g/L tryptophan, 5 g/L yeast extract, 10 g/L NaCl, 3 mM cysteine, pH 8.0 and temperature at 35℃. Under these conditions, the recombinant *E. coli* produced 223.6 mg/L of indirubin in the 10L fermenter. This study provides, a strong evidence for production of high purity indirubin from tryptophan in the presence of cysteine.

# Appendix

# Draft genome sequence of *Methylophaga aminisulfidivorans* MP$^T$

This chapter is published in Journal of Bacteriology 193 (3), 4265 (2011)

# 1. Introduction

The genus *Methylophaga* is a unique group of aerobic, halophilic, and non-methane-utilizing methylotrophs [Janvier, 1995]. Members of this genus have been isolated from various marine sediments and soda lakes, and play an important role in the coastal environments participating in biogeochemical cycling of one-carbon substrates containing nitrogen, sulfur or halogens [Anthony, 1982; Kim, 2007]. They use ribulose monophosphate (RuMP) pathway for one-carbon compound assimilation. The genus *Methylophaga* currently has seven species with validly published names while only the genome sequence of *M. thiooxydans* (strainDMS010T) is reported [Boden, 2011]. *Methylophaga aminisulfidivorans* MP$^T$ (KCTC 12909$^T$ = JCM14647$^T$) was isolated from the seawater at Mokpo, South Korea [Choi, 2003]. This strain is aerobic, neutrophilic, halophilic, and grows well on methanol, methylated amines, dimethysulfide and DMSO. Only fructose is utilized as a multi-carbon source. Whole genome sequence information provided in the present study should facilitate more metabolic diversity investigation of the *Methylophaga* genus in future. In this study, the whole genome sequence of *M. aminisulfidivorans* MP$^T$ was investigated.

## 2. Materials and methods

### 2-1. Strain and media

*Methylophaga aminisulfidivorans* MP$^T$ (KCTC12909$^T$ = JCM14647$^T$) was cultivated in standard mineral base medium (SMM), comprised of 3% (w/v) sodium chloride and 1% (w/v) methanol, at 30℃ [Choi, 2003].

### 2-2. General DNA manipulation

Total genomic DNA from *M. aminisulfidivorans* MP$^T$ was prepared using an iNtRONBio G-spinTM Genomic DNA extraction kit (iNtRON). Plasmid DNA was obtained using a plasmid purification mini kit (NucleoGen). DNA ligation, transformation and other recombinant DNA manipulation methods followed the standard methods described by Sambrook *et al* [Sambrook, 2001]. The DNA sequencing and primer synthesis were performed at Bionics Inc (Table 6-1).

### 2-3. Sequence methods

The high coverage genome sequence of *M. aminisulfidivorans* MP$^T$ was determined using a combination of Illumina Genome Analyzer IIx (150 bp single end shotgun sequencing; 33,762,341 sequencing reads; 1365′ fold coverage) and Roche Genome Sequencer FLX system (paired-end with an insert size of 3 Kb; 909,213 reads; 60′ fold coverage). The assembly of Illumina sequencing reads was carried out the

CLC genomics wb4 (CLCbio) whereas Newbler assembler 2.3 (Roche) was used to generate scaffolds from a 454 paired-end library. Two assemblies were merged into single assembly using the CodonCode Aligner software (CodonCode Co.). The resultant genome sequence was uploaded into the RAST server [Aziz, 2008] to predict the open reading frames, tRNAs and rRNAs. The predicted ORFs were annotated by searching against clusters of orthologous group [Tatusov, 1997] and SEED databases [Disz, 2010].

Table 6-1. Primers used in genome sequencing

| Name | Sequence (5' → 3') |
|---|---|
| ctg2-5'R | 5' - CTCAAACGCATCATCCCATCATTG - 3' |
| ctg2-3'F | 5' - TTGCAGGAGTCTTCGATCAGCTTG - 3' |
| ctg3-5'R | 5' - AGTTTACAGCTCTGCGCAGGCAAA - 3' |
| ctg3-3'F | 5' - ATGAGTCACCAGATGACACTGATG - 3' |
| ctg15-5'R | 5' - TCGTCTGTTGTTGATGAGGTGCAA - 3' |
| ctg15-3'F | 5' - AATACATGGCTAGAAGACGGAGCC - 3' |
| ctg16-5'R | 5' - ACATTAGGTTTGACGGGTCACCAG - 3' |
| ctg16-3'F | 5' - AGTCGCATTTACCTCGATTCTTCC - 3' |
| ctg17-5'R | 5' - TTTTCAGCCATTAGAACGACTCCC - 3' |
| ctg17-3'F | 5' - TTGCAGCCGATCTATCCATCAAAG - 3' |
| ctg35-5'R | 5' - TGGGAGGTCTCAAAACTCTGGCTT - 3' |
| ctg35-3'F | 5' - TTATCACGTCAGGTTGGTGTACCT - 3' |
| ctg34-5'R | 5' - CCCATAGCATCAATGCGACCATCA - 3' |
| ctg34-3'F | 5' - CGATTCTGGTTGTAAACGCAGCTG - 3' |
| ctg33-5'R | 5' - CTTGCGCTTTTGCAAAGTGACCAG - 3' |
| ctg33-3'F | 5' - GGTTTGTCATTAGGCATGCACTTA - 3' |
| ctg32-5'R | 5' - CTACTTAGATGTTTCAGTTCGCCG - 3' |
| ctg32-3'F | 5' - ATTGAAGAGCCTGATGTCGCTTCA - 3' |
| ctg31-5'R | 5' - TTGTTTGCTCTATCTCTGGCTCGA - 3' |
| ctg31-3'F | 5' - GCTTGGTCATTTGTTTCGCACCTA - 3' |
| ctg30-5'R | 5' - AGGAACATATACAAGCCCAAGCTC - 3' |
| ctg30-3'F | 5' - CACACGATATGGTCGTCTGCTTGA - 3' |
| ctg29-5'R | 5' - GCAACCTCAATGACTACCTCATCA - 3' |
| ctg29-3'F | 5' - ACGATACTGATAGCGGCCGACAAA - 3' |
| ctg28-5'R | 5' - TGCTCATCGAGGTGGAATAGATAC - 3' |
| ctg28-3'F | 5' - CCTAGCTTCATCGTTGTTTGGATC - 3' |
| ctg27-5'R | 5' - CATCTATGCCTCGGAAACTAGTAG - 3' |
| ctg27-3'F | 5' - GGATATGCTGTATGCCGTCCATGT - 3' |

# 3. Results and discussion

## 3-1. Genome sequencing of *M. aminisulfidivorans* MP$^T$

The genome sequencing results showed 35 contigs by Illumina Genome Analyzer IIx and Roche Genome Sequencer FLX system. The gap-closing steps was performed standard PCR and Sanger sequencing. From the gap-closing results, *M. aminisulfidivorans* MP$^T$ genome sequence data was assembled 2 contigs (2,266,798 and 825,287bp, respectively; G+C mol%=43.5) representing a chromosome and one circular plasmid (169,875 bp; G+C mol%=41.20) (Table 6-2 and Fig. 6-1). The resultant genome sequence was uploaded into the RAST server [Aziz, 2008] to predict the open reading frames, tRNAs and rRNAs. The predicted ORFs were annotated by searching against clusters of orthologous group [Tatusov, 1997] and SEED databases [Disz, 2010].

The chromosome of *M.aminisulfidivorans* MP$^T$ harbors 2,984 protein coding genes, 3 rRNA operons, and 31 tRNAs whereas 190 coding genes and 9 tRNAs were found in the plasmid (Table 6-2). In addition, cluster of othologous group (COG) analysis was described Table 6-3.

The gene clusters responsible for methanol oxidation (*mxaFJGIRSACKL*) and for synthesis of the cofactor pyrroloquinoline quinone (*pqqBCDE*) were identified from the draft genome sequence. Five types of cytochromes and one cytochrome *c* oxidase ($cbb_3$) were also predicted. A gene encoding a novel bacterial flavin-containing monooxygenase (FMO) (MAMP_00532), which catalyzes nitrogen-containing compounds or indole by use of oxygen through an NADPH-dependent pathway [Choi, 2003] was identified. At least

four different gene clusters potentially encoding restriction modification (RM) systems, which include type I and type II, were predicted. Very low transformation efficiency of *M. aminisulfidivorans* $MP^{T,}$ due to the presence of multiple RM systems. The *M. aminisulfidivorans* $MP^{T}$ whole genome sequence information provided in the present study should facilitate future studies of the metabolic diversity of the genus *Methylophaga*.

Table 6-2. Summary of genome information

| Name | Genome size | G + C ratio | CDS | rRNA | tRNA |
|---|---|---|---|---|---|
| *M. aminisulfivorans* MP$^T$ | 3,261,960 | 43.37 | 3174 | 3 | 40 |
| Contig 1 | 2,266,798 | 43.53 | 2172 | 3 | 31 |
| Contig 2 | 825,287 | 43.39 | 812 | | 9 |
| plasmid | 169,875 | 41.20 | 190 | | |

Table 6-3. Clusters of orthologous group (COG), distribution in the genome

| COG | Description | Count | % |
|---|---|---|---|
| J | Translation, ribosomal structure and biogenesis | 151 | 6.35 |
| K | Transcription | 117 | 4.92 |
| L | Replication, recombination and repair | 122 | 5.13 |
| D | Cell cycle control, cell division, chromosome partitioning | 33 | 1.39 |
| O | Posttranslational modification, protein turnover, chaperones | 113 | 4.75 |
| M | Cell wall/membrane/envelope biogenesis | 195 | 8.20 |
| N | Cell motility | 122 | 5.13 |
| P | Inorganic ion transport and metabolism | 175 | 7.36 |
| T | Signal transduction mechanisms | 176 | 7.40 |
| C | Energy production and conversion | 162 | 6.81 |
| G | Carbohydrate transport and metabolism | 94 | 3.95 |
| E | Amino acid transport and metabolism | 180 | 7.57 |
| F | Nucleotide transport and metabolism | 57 | 2.40 |
| H | Coenzyme transport and metabolism | 119 | 5.00 |
| I | Lipid transport and metabolism | 68 | 2.86 |
| Q | Secondary metabolites biosynthesis, transport and catabolism | 18 | 0.76 |
| R | General function prediction only | 237 | 9.97 |
| S | Function unknown | 239 | 10.05 |
| Total | | 2378 | 100 |

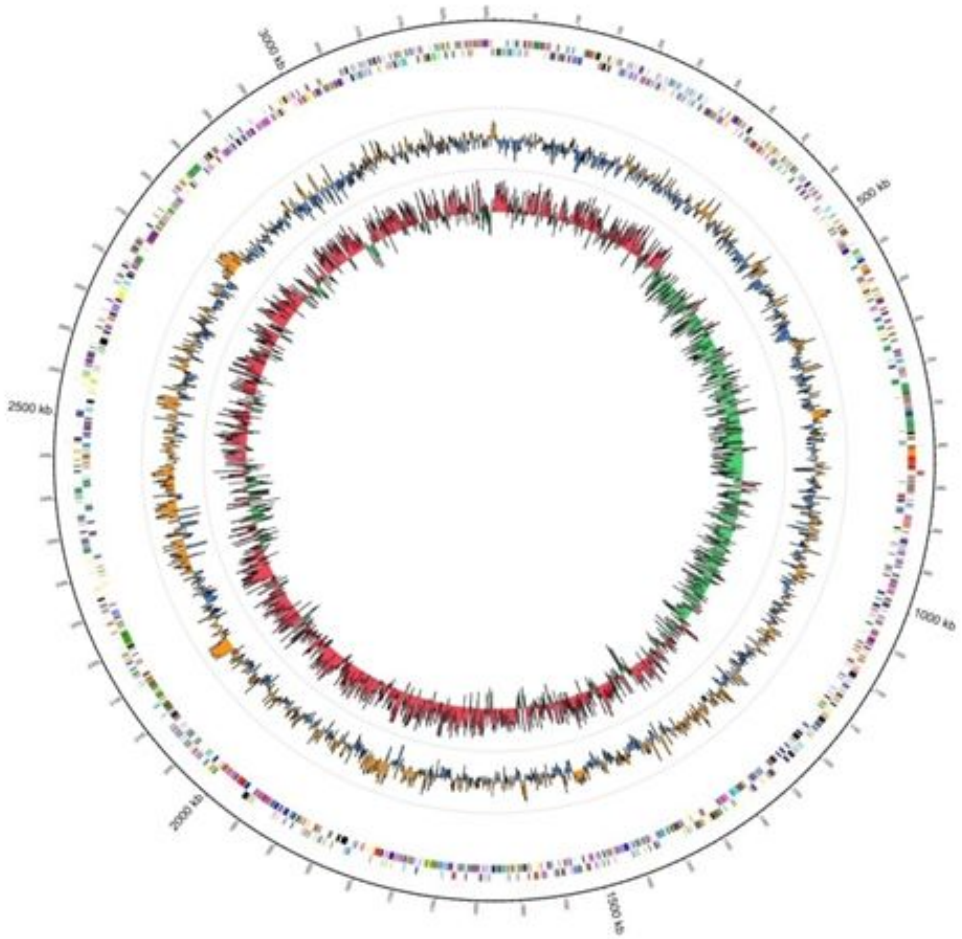



Fig. 6-1. Circular genome map of pseudo chromosome of 3 moleculs. 1. Coordinate, 2. RNA, 3. Reverse CDS, 4. Forward CDS, 5. GC Skew, 6. GC Ratio.

# Reference

Anthony, C. **1982**. The biochemistry of methylotrophs. Academic Press, London and New York.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O. **2008**. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, 9(75).

Bae, J.W., Han, J.H., Park, M.S., Lee, S.G., Lee, E.Y., Jeong, Y.J., Park, S.H. **2006**. Development of recombinant *Pseudomonas putida* containing homologous styrene monooxygenease genes for the production of (S)-styrene oxide. *Biotechnol Bioprocess Eng,* 11, 530-537.

Baeyer, A., Drewsen, V. **1882**. Darstellung von indigblau aus orthonitrobenzaldehyd. *Berichte der deutschen chemischen Gesellschaft*, 15(2), 2856-2864.

Beaty, N.B., Ballou, D.P. **1981**. The reductive half-reaction of liver microsomal FAD-containing monooxygenase. *Journal of Biological Chemistry*, 256(9), 4611-4618.

Berry, A., Dodge, T.C., Pepsin, M., Weyler, W. **2002**. Application of metabolic engineering to improve both the production and use of biotech indigo. *J Ind Microbiol Biotechnol*, 28, 127-133.

Biasiolo, M., Bertazzo, A., Costa, C., Beghetto, A., Allegri, G. **1995**. Determination of nonprotein tryptophan in yoghurts by selective fluorescence and HPLC. *Food Chem.*, 52, 87-92.

Boden, R., Ferriera, S., Johnson, J., Kelly, D.P., Murrell, J.C., Schafer, H. **2011**. Draft Genome Sequence of the Chemolithoheterotrophic, Halophilic Methylotroph *Methylophaga thiooxydans* DMS010. *J Bacterio.*, 193(12), 3154-3155.

Bradford, M.M. **1976**. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, 72(1-2), 248-254.

Cashman, J.R. **2004**. The implication of polymorphisms in mammalian flavin-containing monooxygenases in drug discovery and development. *Drug Discov Today*, 9, 574-581.

Chant, E.L., Summers, D.K. **2007**. Indole signalling contributes to the stable maintenance of *Escherichia coli* multicopy plasmids. *Mol Microbiol*, 63(1), 35-43.

Chen, S.J., Ke, B.S., Chiu, I.C. **2008**. A fermentation strategy for production of recombinant protein subjected to plasmid instability. *Kor J Chem Eng*, 25, 1110-1114.

Cho, H.J., Cho, H.Y., Kim, K.J., Kim, M.H., Kim, S.W., Kang, B.S. **2011**. Structural and functional analysis of bacterial flavin-containing monooxygenase reveals its ping-pong-type reaction mechanism. *J Struct Biol*, 175(1), 39-48.

Choi, H.S., Kim, J.K., Cho, E.H., Kim, Y.C., Kim, J.I., Kim, S.W. **2003**. A novel flavin-containing monooxygenase from *Methylophaga* sp. strain SK1 and its indigo synthesis in *Escherichia coli*. *Biochemical and Biophysical Research Communications*, 306(4), 930-936.

Disz, T., Akhter, S., Cuevas, D., Olson, R., Overbeek, R., Vonstein, V., Stevens, R., Edwards, R. **2010**. Accessing the SEED Genome Databases via Web Services API: Tools for Programmers. *BMC Bioinformatics*, 11(1), 319.

Doukyu, N., Aono, R. **1997**. Biodegradation of indole at high concentration by persolvent fermentation with *Pseudomonas* sp. ST-200. *Extremophiles*, 1(2), 100-105.

Doukyu, N., Nakano, T., Okuyama, Y., Aono, R. **2002**. Isolation of an *Acinetobacter* sp. ST-550 which produces a high level of indigo in a water-organic solvent two-phase system containing high levels of indole. *Applied Microbiology and Biotechnology*, 58(4), 543-546.

Eaton, R.W., Chapman, P.J. **1995**. Formation of indigo and related compounds from indolecarboxylic acids by aromatic acid-degrading bacteria: chromogenic reactions for cloning genes encoding dioxygenases that act on aromatic acids. *J Bacteriol*, 177(23), 6983-6988.

Ensley, B.D., Ratzkin, B.J., Osslund, T.D., Simon, M.J., Wackett, L.P., Gibson, D.T. **1983**. Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science*, 222(4620), 167-169.

Fitzhugh, E.W., Feller, R.L., Roy, A. **1997**. Artists' Pigments: A Handbook of Their History & Characteristics. 3 ed. Oxford University Press, Oxford.

Fraaije, M.W., Kamerbeek, N.M., van Berkel, W.J.H., Janssen, D.B. **2002**. Identification of a Baeyer-Villiger monooxygenase sequence motif. *FEBS Letters*, 518(1-3), 43-47.

Friehs, K. **2004**. Plasmid copy number and plasmid stability. *Advances in biochemical engineering/biotechnology*, 86, 47-82.

Gasparetti, C., Buzzini, P., Caramarossa, M.R., Turchetti, B., Pagnoni, U.M., Forti, L. **2006**. Application of the response surface methodology (RSM) for optimizing the production of volatile organic compounds (VOCs) by *Trichosporon moniliiforme*. *Enzyme Microb Technol*, l39, 1341−1346.

Han, G.H., Bang, S.E., Babu, B.K., Chang, M., Shin, H.J., Kim, S.W. **2011**. Bio-indigo production in two different fermentation systems using recombinant *Escherichia coli* cells harboring a flavin-containing monooxygenase gene (*fmo*). *Process Biochemistry*, 46(3), 788-791.

Han, G.H., Shin, H.J., Kim, S.W. **2008**. Optimization of bio-indigo production by recombinant *E. coli* harboring *fmo* gene. *Enzyme and Microbial Technology*, 42(7), 617-623.

Han, R. **1994**. Highlight on the studies of anticancer drugs derived from plants in china. *Stem Cells*, 12(1), 53-63.

Hoessel, R., Leclerc, S., Endicott, J.A., Nobel, M.E.M., Lawrie, A., Tunnah, P., Leost, M., Damiens, E., Marie, D., Marko, D., Niederberger, E., Tang, W., Eisenbrand, G., Meijer, L. **1999**. Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. *Nat Cell Biol*, 1(1), 60-67.

Janvier, M., Grimont, P.A.D. **1995**. The genus *Methylophaga*, a new line of descent within phylogenetic branch [gamma] of proteobacteria. *Research in Microbiology*, 146(7), 543-550.

Jung, K.H. **2006**. Continuous production of recombinant interferon-alpha in *Escherichia coli* via the derepression of *trp* promoter using casamino acid. *Process Biochemistry*, 41(4), 809-814.

Kim, H, Eom, H.J., Lee, J.S., Han, J.S., Han, N.S. **2004**, Statistical optimization of medium composition for growth of *Leuconostoc citreum*. *Biotechnol Bioprocess Eng*, 9, 278-284.

Kim, H.G., Doronina, N.V., Trotsenko, Y.A., Kim, S.W. **2007**. *Methylophaga aminisulfidivorans* sp. nov., a restricted facultatively methylotrophic marine bacterium. *International Journal of Systematic and Evolutionary Microbiology*, 57(9), 2096-2101.

Kohda, H., Niwa, A., Nakamoto, Y., Takeda, O. **1990**. Flavonoid glucosides from *Polygonum tinctorium*. *Chem Pharm Bull*, 38, 523-524.

Krueger, S.K., Williams, D.E. **2005**. Mammalian flavin-containing monooxygenases:

structure/function, genetic polymorphisms and role in drug metabolism. *Pharmacology & Therapeutics*, 106(3), 357-387.

Lattard, V., Zhang, J., Cashman, J.R., **2004**. Alternative processing of flavin-containing monooxygenase (FMO) pre-mRNA in human. *Mol Pharmacol*, 65, 1517-1525.

Lawton, M.P., Cashman, J.R., Cresteil, T., Dolphin, C.T., Elfarra, A.A., Hines, R.N., Hogson, E., Kimura, T., Ozols, J., Phillips, I.R., **1994**. A nomenclature for the mammalian flavin-containing monooxygenase gene family based on amino acid sequence identities. *Arch Biochem Biophys*, 308, 254-257.

Leclerc, S., Garnier, M., Hoessel, R., Marko, D., Bibb, J.A., Snyder, G.L., Greengard, P., Biernat, J., Wu, Y.-Z., Mandelkow, E.-M., Eisenbrand, G., Meijer, L. **2001**. Indirubins Inhibit Glycogen Synthase Kinase-3 and CDK5/P25, Two Protein Kinases Involved in Abnormal Tau Phosphorylation in Alzheimer's Disease. *Journal of Biological Chemistry*, 276(1), 251-260.

Lee, J., Jayaraman A., Wood T. **2007**. Indole is an inter-species biofilm signal mediated by SdiA. BMC Microbiology 7:42.

Lee, J.H., Lee, J. **2011**. Indole as an intercellular signal in microbial communities. *FEMS Microbiology Reviews*, 34(4), 426-444.

Luo, Z., Hines, R.N., **2001**. Regulation of flavin-containing monooxygenase 1 expression by Ying Yang 1 and hepatic nuclear factors 1 and 4. *Mol Pharmacol*, 60, 1421-1430.

Madsen, E.L., Bollag, J.M. **1988**. Pathway of indole metabolism by a denitrifying microbial community. *Archives of Microbiology*, 151(1), 71-76.

Marko, D., Schätzle, S., Friedel, A., Genzlinger, A., Zankl, H., Meijer, L., Eisenbrand, G.D., **2001**. Inhibition of cyclin-dependent kinase 1 (CDK 1) by indirubin derivatives in human tumor cells. *Brit J Cancer*, 84(2), 283-289.

Maugard, T., Enaud, E., Sayette, A., Choisy, P., Legoy MD. **2002**. $\beta$ -Glucosidase-Catalyzed Hydrolysis of Indican from Leaves of *Polygonum tinctorium*. *Biotechnology Progress*, 18(5), 1104-1108.

McClay, K., Boss, C., Keresztes, I., Steffan, R.J. **2005**. Mutations of Toluene-4-Monooxygenase That Alter Regiospecificity of Indole Oxidation and Lead to Production of Novel Indigoid Pigments. *Appl Environ Microbiol*, 71(9), 5476-5483.

Merkel, T.J., Boucher, P.E., Stibitz, S., Grippe, V.K. **2003**. Analysis of *bvg*R expression in *Bordetella pertussis*. *Journal of Bacteriology*, 185(23), 6902-6912.

Meyer, A., Würsten, M., Schmid, A., Kohler, H.P.E., Witholt, B. **2002**. Hydroxylation of indole by laboratory-evolved 2-hydroxybiphenyl 3-monooxygenase. *J Biol Chem*, 277(37), 34161-34167.

Murdock, D., Ensley, B.D., Serdar, C., Thalen, M. **1993**. Construction of Metabolic Operons Catalyzing the *De Novo* Biosynthesis of Indigo in *Escherichia coli*. *Nat Biotech*, 11(3), 381-386.

Myers, R.H., Montgomery, D.C. **2002**. Response surface methodology; process and product optimization using designed experiments. 2nd ed. NewYork: A Wiley-Interscience Publication.

Nam, S., Buettner, R., Turkson, J., Kim, D., Cheng, J.Q., Muehlbeyer, S., Hippe, F., Vatter, S., Merz, K.H., Eisenbrand, G., Jove, R. **2005**. Indirubin derivatives inhibit Stat3 signaling and induce apoptosis in human cancer cells. *Proc Natl Acad Sci*, 102(17), 5998-6003.

Newton, W. A., and E. E. Snell. **1965**. Formation and Interrelationships of Tryptophanase and Tryptophan Synthetases in *Escherichia coli*. *J Bacteriol*, 89,355-364.

Oae, S., Mikami, A., Matsuura, T., Ogawa-Asada, K., Watanabe, Y., Fujimori, K., Iyanagi, T. **1985**. Comparison of sulfide oxygenation mechanism for liver microsomal FAD-containing monooxygenase with that for cytochrome P-450. *Biochemical and Biophysical Research Communications*, 131(2), 567-573.

O'Connor, K.E., Dobson, A.D., Hartmans, S. **1997**. Indigo formation by microorganisms expressing styrene monooxygenase activity. *Appl Environ Microbiol*, 97(63), 4287-4291.

Perpete, E.A., Preat, J., Andre, J.M., Jacquemin, D. **2006**. An ab initio study of the absorption spectra of indirubin, isoindigo, and related derivatives. *J. Phys. Chem. A*, 110(17), 5629-5635.

Pinero-Fernandez, S., C. Chimerel, U. F. Keyser, and D. K. Summers. **2011**. Indole Transport across *Escherichia coli* Membranes. *J Bacteriol*, 193, 1793-1798.

Rettie, A.E., Fisher, M.B., **1999**. Handbook of drug metabolism (Woolf TF ed), Marcel Dekker, New York.

Roessler, A., Jin, X. **2003**. State of the art technologies and new electrochemical methods for the reduction of vat dyes. *Dyes and Pigments*, 59(3), 223-235.

Royo, J.L., Moreno-Ruiz, E., Cebolla, A., Santero, E. **2005**. Stable long-term indigo production by overexpression of dioxygenase genes using a chromosomal integrated cascade expression circuit. *J Biotechnol*, 116, 113-124.

Rui, L., Reardon, K.F., Wood, T.K. **2005**. Protein engineering of toluene *ortho*-monooxygenase of *Burkholderia cepacia* G4 for regiospecific hydroxylation of indole to form various indigoid compounds. *Appl Microbiol Biotechnol*, 66(4), 422-429.

Sambrook, J., Russell, D.W. **2001**. Molecular cloning : a laboratory manual. 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schhrott, W. **2000**. Denim wieder im Blickpunkt der Textilindustrie. *Melliand Textilber* 81(3), 190-194.

Seldes, A.M., Burucúa, J.E., Maier, M.S., Abad, G., Jáuregui, A., Siracusano, G. **1999**. Blue Pigments in South American Painting (1610-1780). *J. Am. Inst. Conserv*, 38(2), 100-123.

Suh, J.K., Poulsen, L.L., Ziegler, D.M., Robertus, J.D. **1999**. Yeast flavin-containing monooxygenase generates oxidizing equivalents that control protein folding in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences*, 96(6), 2687-2691.

Tang, W., Eisenbrand, G. **1992**. Chinese Drugs of Plant Origin: Chemistry, Pharmacology, and Use in Traditional and Modern Medicine 1ed. Springer, Berlin Heidelberg New York.

*van* Berkel, W.J.H., Kamerbeek, N.M., Fraaije, M.W. **2006**. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *Journal of Biotechnology*, 124(4), 670-689.

Walter, P., Weygand-Durasevic, I., Sanni, A., Ebel, J.P., Fasiolo, F. **1989**. Deletion analysis in the amino-terminal extension of methionyl-tRNA synthetase from *Saccharomyces cerevisiae* shows that a small region is important for the activity and stability of the enzyme. *The Journal of biological chemistry*, 264(29), 17126-30.

Woo, H.j., Sanseverino, J., Cox, C.D., Robinson, K.G., Sayler, G.S. **2000**. The measurement of toluene dioxygenase activity in biofilm culture of *Pseudomonas putida* F1. *J Microbiol Meth*, 40(2), 181-191.

Wu, G.Y., Fang, F.D., Liu, J.Z., Chang, A., Ho, Y.H., **1980**. Studies on the mechanism of action of indirubin in the treatment of chronic granulocytic leukemia. I. Effects on nucleic acid and protein synthesis in human leukemic cells. *Chinese Medical J*, 60, 451−454.

Xia, Z.Q., Zenk, M.H. **1992**. Biosynthesis of indigo precursors in higher plants. *Phytochemistry*, 31(8), 2695-2697.

Yan L, Lehe M. **2007**. Co-expression of P450 BM3 and glucose dehydrogenase by recombinant *Escherichia coli* and its application in an NADPH-dependent indigo production system. *J Ind Microbiol Biotechnol*, 34, 247-253.

Yanofsky, C., V. Horn, and P. Gollnick. **1991**. Physiological studies of tryptophan transport and tryptophanase operon induction in *Escherichia coli*. *J Bacteriol*, 173, 6009-6017.

Zhao, Y., Christensen, S.K., Fankhauser, C., Cashman, J.R., Cohen, J.D., Weigel, D., Chory, J. **2001**. A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science*, 291(5502), 306-309.

Zhou, J., Shephard, E.A. **2006**. Mutation, polymorphism and perspectives for the future of human flavin-containing monooxygenase 3. *Mutation Research/Reviews in Mutation Research*, 612(3), 165-171.

Ziegler, D.M., 1988. Flavin-containing monooxygenases: catalytic mechanisms and substrate specificities. *Drug Met Rev*, 19, 1-32.

# Overall conclusion

Indigo is one of the oldest blue dyes used by man. Furthemore, indirubin has been used as a natural drug and dye since ancient times. Microbially produced indigo and indirubin can be used as alternatives to chemically synthesized compounds. Although several studies using recombinant microorganisms expressing mono- or di-oxygenase for the biological production of indigo and indirubin from indole have been reported, no attempt has yet been made for the large-scale and continuous production of indigo with recombinant microorganisms. A bacterial FMO gene was cloned from *M. aminisulfidivorans* $MP^T$, and a plasmid pBlue 2.0 was constructed to express the bacterial *fmo* gene in *E. coli*. Recombinant *E. coli* pBlue 2.0 can effectively metabolize tryptophan for the production of indigo and indirubin.

Recombinant *E. coli* pBlue 2.0 (2,024bp) contained upstream region (663 bp) and *fmo* (1361 bp) structural gene which is cloned into pBluescript SKII(+) expression system. However, recombinant *E. coli* pBlue 2.0 did not require an inducer for FMO expression. In this study, the putative promoter and regulator regions of *fmo* in the plasmid were investigated in order to achieve increased indigo and indirubin production yields. The transcription of the *fmo* gene took place at T (+1), which is situated 279 bp (TSS1) upstream from the translation initiation codon of the *fmo* gene. On the basis of the sequenced data of the *fmo* transcripts, we identified the -35 region (5'-CTGGAA-3') and the -10 region (5'-TATCCT-3'). In addition, extended -10 promoter was considered as TG at -14, -15bp upstream from the TSS1. The recombinant *E. coli* harboring pBlue 2.0 plasmid produced 160 mg/L of indigo from 2 g/L tryptophan medium. To increase the production of indigo, upstream sequence length of the *fmo*

gene was optimized and response surface methodology was used. The pBlue 1.7 plasmid (1,686 bp) was prepared by deleting the sequence upstream of pBlue 2.0. The recombinant *E. coli* harboring the pBlue 1.7 plasmid produced 662 mg/L of indigo in tryptophan medium after 24h of flask cultivation.

Additionally, we investigated the optimum culture conditions and the optimal culture media for increasing indigo and indirubin compound yields. The production of indigo was optimized using a response surface methodology with a $2^n$ central composite design. The optimal combination of media constituents to produce maximum yields of indigo was determined to be as follows: 2.4 g/L tryptophan, 4.5 g/L yeast extract, and 11.4 g/L sodium chloride. In addition, the optimum culture temperature and pH were 30℃ and pH 7.0, respectively. Under the optimized conditions mentioned above, the recombinant *E. coli* harboring the pBlue 1.7 plasmid produced 920 mg/L in optimum tryptophan medium after 24 h of cultivation. The combination of the truncated insert sizes and culture optimization resulted in a 575% increase in the production of indigo. Indigo was produced in 3,000 L repeated batch and 5 L continuous fermentation of a recombinant *E. coli* DH5α harboring a novel gene encoding flavin-containing monooxygenase (FMO). Repeated batch fermentation in a 3,000 L fermenter produced 911 ± 22 mg/L of indigo from 2 g/L tryptophan as a substrate (yield, 46.9%) under the following culture conditions as follows: culture temperature 30°C; pH 7.0; agitation speed 200 rpm; and aeration 3 vvm. Sufficient oxygen (aeration rate and agitation speed) was critical for indigo production. For continuous fermentation in a 5 L fermenter, the volumetric productivity was found to be 11.3 mg/L/h up to 110 h (final accumulated indigo was 23 g) with a constant dilution rate (D) of 0.084/h (constant

feeding rate of 0.167 L/h with medium containing 3 g/L tryptophan). Recombinant *E. coli* pBlue 1.7 can withstand the toxicity of high concentrations of accumulated indigo during batch fermentation. During continuous fermentation, the recombinant cells exhibited high plasmid stability up to 110 h, after which the plasmid was lost from the cells. Such indigo-production systems using *E. coli* pBlue 1.7 have the potential be used for in industrial applications, thus overcoming the environmental problems related with synthetic indigo while meeting the demands for natural indigo in the dye and textile industry.

Recombinant *E. coli* pBlue 1.7 was used to produce indigo (920 mg/L) and indirubin (≤5 mg/L) in a 5 L fermenter containing tryptophan medium. In this study, an efficient indirubin production process was developed using recombinant *E. coli* DH5 α harboring an *fmo* gene. When 3 mM cysteine was added to the tryptophan medium, the production of indirubin from the cells significantly increased. The addition of cysteine did not influence the expression level or activity of FMO in the cell. However, cysteine was found to inhibit the cell growth of recombinant *E. coli* and to affect the synthetic pathway of 2- and 3-hydroxyindole from indole, which might increase indirubin production. The recombinant *E. coli* cells could produce 223.6 mg/L of indirubin under the following optimum culture conditions: 2 g/L tryptophan, 5 g/L yeast extract, 10 g/L NaCl, 0.363 g/L (3 mM) cysteine, and pH 8.0 at 35°C. The synthesized indirubin compounds accumulated inside the cells and were released later because of the disruption of the cells due to the toxicity of the accumulated indirubin compounds. This study shows that it is possible to produce the expensive indirubin in large amounts by simple fermentation with recombinant *E. coli* cells.

In addition, we investigated the draft genome sequence of *M. aminisulfidivorans* MP$^T$. *M. aminisulfidivorans* MP$^T$ (KCTC 12909T = JCM 14647T) has consists of a chromosome (3,092,085 bp) and a plasmid (16,875 bp). The *M. aminisulfidivorans* MP$^T$ chromosome was found to harbor 2,984 protein-coding genes, 3 rRNA operons, and 31 tRNAs, whereas the plasmid was found to contain 190 coding genes and 9 tRNAs. We identified a gene encoding a novel FMO (MAMP_00532) what catalyzes nitrogen-containing compounds or indole by the use of oxygen through an NADPH-dependent pathway. The draft genome sequence information obtained in the current study should facilitate future investigations into the metabolic diversity of the *Methylophaga* genus.

# 요 약 문

본 연구는 일원자탄소대사세균인 *Methylophaga aminisulfidivorans* MP$^T$에 내포된 flavin-containing monooxygenase 유전자의 upstream에 존재하는 promoter region을 분석하고, flavin-containing monooxygenase 유전자를 포함하는 재조합 대장균을 이용하여 인디고 및 인디루빈을 생산 하고자 하였다. 일반적으로 인디고 및 인디루빈은 화학합성 방법과 식물로부터 추출하는 방법으로 생산되어지고 있다. 이와 같은 생산방식은 생산 중 발생하는 부산물에 의한 환경오염 및 인체 유해성을 가지고 있으며, 생산성의 한계를 가지고 있다. 따라서 본 연구에서는 *M. aminisulfidivorans* MP$^T$로부터 유래된 flavin-containing monooxygenase를 이용하여 인체 및 환경에 무해하고, 고생산성을 가지는 생물학적 인디고 및 인디루빈 대량생산 공정을 개발하고자 하였다. 또한 *M. aminisulfidivorans* MP$^T$의 전체 유전자 서열을 분석하였다.

*M. aminisulfidivorans* MP$^T$ 유래 flavin-containing monooxygenase 유전자는 1361bp의 염기로 구성되어져 있다. 이 연구에서는 재조합 plasmid pBlue2.0으로부터 약 100bp 간격으로 upsteam region을 제거하여 이들 재조합 plasmid pBlue들을 포함하는 재조합 대장균 *E. coli* pBlues를 제작하였다. 이 중 upstream region의 길이가 315bp인 pBlue1.7에서 가장 많은 인디고가 생산되었다. pBlue1.7에 포함되어진 *fmo* 유전자의 전사시작점은 총 6개가 존재하였다. 인디고 생산에 가장 큰 영향을 미치는 전사시작점을 분석한 결과 *fmo* 유전자의 translation initiation codon으로부터 279번째에 nucleotide T(+1) 였다. *fmo* 전사시작점으로부터 8bp upstream 위치에 있는 5'-TATCCT-3' 과 14-15 upstream 위치에 있는 5'-TG-3'염기서열이 존재하는 고 있었으

며 이들 염기 서열이 *fmo*의 extended -10 promoter로 작동하였다.

인디고 생산에 적합한 숙주 및 발현시스템을 구축하기 위하여 *E. coli*, *B. subtils* 및 *C. glutamicum*을 숙주를 이용하여 재조합 *E. coli* pHCE IIB 1.7, *E. coli* pGEX 4T-1 1.7, *E. coli* pBSK 1.7 (pBlue1.7), *B. subtils* pHY300PLK 1.7, *C. glutamicum* pEKE 1.7를 제작하였다. 이들 재조합 미생물로부터 각각 400.2 mg/L, 638.5 mg/L, 672.3 mg/L, 182 mg/L, 377 mg/L의 인디고가 생산되었으며, 이 중 *E. coli* pBSK 1.7 (pBlue1.7)에서 가장 높은 인디고 생산성을 나타내었다. 인디고의 최적생산을 위하여 반응표면분석법 (response surface methodology, RSM) 중 중심합성계획법 (central composite design, CCD)을 이용하여 배양조건 및 배지의 최적화를 수행하였다. 최적 배양배지의 조성은 tryptophan 2g/L, yeast extract 5g/L, sodium chloride 10g/L이었으며, 최적 배양 조건은 pH 7.0, 30℃, 180rpm이었다. 최적배양조건에서 10L, 100L, 3,000L 배양기를 이용하여 회분식배양을 수행하였다. 그 결과 재조합 *E. coli* pBlue1.7은 각각 모든 배양기에서 920mg/L의 인디고를 생산하였다. 또한 3,000L 배양기를 이용한 반연속식 배양에서 각각 1차 배양 (911 ± 22 mg/L), 2차 배양 (870 ± 48), 3차 배양 (702 ± 66 mg/L)의 인디고를 생산하였다. 10L 배양기를 이용한 연속식 배양에서는 110시간 동안 23 g의 인디고를 생산하였으며, 생산율은 11.3 mg/L/h이었다.

재조합 *E. coli* pBlue1.7은 tryptophan 배지에서 주 생산물인 인디고 (920 mg/L) 와 소량의 인디루빈 (5 mg/L)을 생산하였다. 인디루빈의 생산을 증가시키기 위하여 tryptophan 배지에 3 mM cysteien을 첨가한 결과 재조합 *E. coli* pBlue1.7로부터 95 mg/L의 인디루빈이 생산되었으며 인디고는 5 mg/L이 생산되었다. 배지 내 cysteine 첨가에 따른 *E. coli* pBlue1.7의 생장률은 약 10배 감소하였으나, FMO의 활성 및 발현양은 영향을 미치지 않았다. 세포내 2-hydroxyindole 및 3-hydroxyindole의

농도를 측정한 결과 cysteine 첨가한 경우 기존에 검출되지 않았던 2-hydroxyindole이 30시간동안 2 μM이 생성되었으며, 3-hydroxyindole은 0.8 μM이 생산되었다. 5L 배양기를 이용한 회분식 최적배양을 통해 223.6 mg/L의 인디루빈을 생산하였다.

*M. aminisulfidivorans* MP$^T$의 전체 유전자 서열의 길이는 3,216,960 bp였으며, 2,984 protein coding genes, 3 rRNA operons, 31 tRNAs 을 포함하는 chromosome 유전자 서열과 190 coding genes, 9 tRNAs 를 포함하는 한개의 plasmid로 구성되어져 있었다. 특히 *M. aminisulfidivorans* MP$^T$는 메탄올 산화와 관련된 methanol oxidation (*mxaFJGIRSACKL*), pyrroloquinoline quinone (*pqqBCDE*) 유전자 클러스터 및 flavin-containing monooxygenase가 존재하였다. 또한 현재까지 알려진 sigma factor와 유사성이 낮은 Sigma 242C FecI-like2C ECF family, RNA polymerase sigma factor RpoD, RNA polymerase sigma-54 factor RpoN, RNA polymerase sigma factor RpoH가 포함되어져 있었다.

# Publication list

1. Hee Gon Kim, **Gui Hwan Han**, Chi-Yong Eom and Si Wouk Kim (2008) Isolation and taxonomic characterization of a novel type I methanotrophic bacterium. The Journal of Microbiology 46(1): 45-50.

2. **Gui Hwan Han**, Hyun-Jae Shin and Si Wouk Kim (2008) Optimization of bio-indigo production by recombinant *E. coli* harboring *fmo* gene. Enzyme and Microbial Technology 42(7): 617-623.

3. Jung Kon Kim, **Gui Hwan Han**, Baek Rock Oh, Young Nam Chun, Chi-Yong Eom and Si Wouk Kim (2008) Volumetric scale-up of a three stage fermentation system for food waste treatment. Bioresource Technology 99(10): 4394-4399.

4. Hee Gon Kim, **Gui Hwan Han** and Si Wouk Kim (2010) Optimization of lab scale methanol production by *Methylosinus trichosporium* OB3b. Biotechnology and Bioprocess Engineering 15(3), 476-480.

5. **Gui Hwan Han**, Seong Eun Bang, Bandamaravuri Kishore Babu, Man Chang, Hyun-Jae Shin and Si Wouk Kim (2011) Bio-indigo production in two different fermentation systems using recombinant *Escherichia coli* cells harboring a flavin-containing monooxygenase gene (*fmo*). Process Biochemistry 46(3): 788-791.

6. **Gui Hwan Han**, Wonduck Kim, Jongsik Chun, and Si Wouk Kim (2011) Draft genome sequence of *Methylophaga aminisulfidivorans* $MP^T$. Journal of Bacteriology 193(6): 4265.

7. Hee Gon Kim, **Gui Hwan Han**, Dockyu Kim, Jong-Soon Choi, Si Wouk Kim (2011) Comparative analysis of two types of methanol dehydrogenase from *Methylophaga aminisulfidivorans* $MP^T$ grown on methanol. Journal of Basic Microbiology 52(2): 141-149 (2012).

8. **Gui Hwan Han**, Geun Ho Gim, Wonduck Kim, Sun Il Seo and Si Wouk Kim (2012) Cysteine enhances indirubin production in the recombinant *Escherichia coli* cell harboring flavin-containing monooxygenase gene, will be submitted to Biotechnology and Bioengineering.