#### Dyes and Pigments 140 (2017) 29-35



Contents lists available at ScienceDirect

### Dyes and Pigments

journal homepage: www.elsevier.com/locate/dyepig

# Biosynthesis of indigo in *Escherichia coli* expressing self-sufficient CYP102A from *Streptomyces cattleya*



PIGMENTS

Hee-Jung Kim <sup>a</sup>, Seyoung Jang <sup>a</sup>, Joonwon Kim <sup>b</sup>, Yung-Hun Yang <sup>c</sup>, Yun-Gon Kim <sup>d</sup>, Byung-Gee Kim <sup>b</sup>, Kwon-Young Choi <sup>a, \*</sup>

<sup>a</sup> Department of Environmental Engineering, College of Engineering, Ajou University, Suwon, Gyeonggi-do, South Korea

<sup>b</sup> Department of Chemical and Biological Engineering, Seoul National University, Seoul, South Korea

<sup>c</sup> Department of Microbial Engineering, College of Engineering, Konkuk University, Seoul, South Korea

<sup>d</sup> Department of Chemical Engineering, College of Engineering, Soongsil University, Seoul, South Korea

#### ARTICLE INFO

Article history: Received 5 May 2016 Received in revised form 10 December 2016 Accepted 12 January 2017 Available online 16 January 2017

Keywords: Cytochrome P450 monooxygenases Self-sufficient CYP102A Indigo Eco-friendly dye Indole L-tryptophan

#### 1. Introduction

### АВЅТКАСТ

Cytochrome P450 monooxygenases (CYP) are a superfamily of heme-thiolate proteins which catalyze the incorporation of oxygen atoms into substrates. Here, a self-sufficient CYP102A from *Streptomyces cattleya* (CYP102A\_scat) was cloned, produced recombinantly in *Escherichia coli* strain BL21 (DE3), and the characteristic features were investigated. However, unlike other self-sufficient CYP102A enzymes that have been reported, CYP102A\_scat was found to be able to catalyze intracellular hydroxylation of indole molecules with 3-C specific regioselectivity. Consequently, *E. coli* strains producing CYP102A\_scat could synthesize approximately 1.0 g/L of indigo in LB media. Optimization of indigo synthesis was investigated through additional feeding of indole precursors such as glucose, L-tryptophan, and indole. Indigo production reached up to 3.8  $\pm$  0.1 g/L by adding 20  $\mu$ M of extracellular indole and 0.2 mM of L-tryptophan to the LB media. To our knowledge, this is a record and the highest yield achieved so far.

© 2017 Elsevier Ltd. All rights reserved.

Indigo-blue, which has a deep dark-blue color, is one of the oldest and most popular dyes [1]. It is as a natural dye, originally extracted from plants such as *Indigofera tinctoria* [2]. Novel chemical and biological routes for the synthesis of indigo were sought as the demand for indigo started to increase. For example, Baeyer-Drewson synthesis is the most well-known organic reaction used for indigo production [3]. In that synthetic route, 2-nitrobenzaldehyde and acetone undergo an aldol condensation reaction, followed by cyclization and oxidative dimerization to produce indigo. This synthetic method was once one of the most useful routes to produce indigo. However, the Baeyer-Drewson synthesis was a relatively impractical process for the synthesis of indigo, using feedstock [7–10].

Indigo at the industrial-scale, and was eventually replaced by the preparation of indigo from aniline. Aniline is readily obtained and can be used to synthesize N-(2-carboxyphenyl)glycine, an indigo precursor [4]. The production of indigo from N-(2-carboxyphenyl)

glycine provided a new and economically attractive route. Approximately 20 million kilograms of indigo are produced annually *via* this synthetic route, and that indigo is used mainly for dyeing blue jeans, and as a colorant in the food industry. Although indigo has a low oral toxicity, with an LD<sub>50</sub> of 5.0 g/kg in mammals, large spills of blue cloth dyes, which include indigo, are nonetheless hazardous because they contain harmful chemical byproducts [5,6].

With the increasing need to develop eco-friendly processes and the growing interest in natural pigments, various attempts have been made to synthesize natural dyes and pigments biologically. One of the most promising of such processes is the use of microbial enzymes within engineered metabolic pathways for the synthesis of indigo, using any of several renewable biomass sources as a feedstock [7–10]. Compared to production routes based on synthetic chemistry, this biological route is relatively simple and fast. Given appropriate biologically-sourced starting-chemicals, biological conversions such as biotransformation and enzymatic modifications of structure can result in the effective synthesis of indigo dye with excellent production yields [11,12]. For example, several attempts for indigo production using flavin-containing monooxygenases (FMO) have been reported [13,14]. Recent research on

<sup>\*</sup> Corresponding author. E-mail address: kychoi@ajou.ac.kr (K.-Y. Choi).

FMO-dependent indigo production has shown that *Escherichia coli* (*E. coli*) that recombinantly produce FMO from *Corynebacterium* glutamicum can produce 685 mg/L of indigo and 103 mg/L of indirubin from 2.5 g/L of L-tryptophan [15].

Here, we demonstrate biosynthesis of indigo dye from glucose using a recombinant E. coli strain producing a cytochrome P450 monooxygenase (CYP). CYP enzymes are heme-thiolate proteins. which catalyze the introduction of one oxygen atom from molecular oxygen to a target substrate molecule in a regioselective and stereoselective manner [16–18]. A few examples of CYP-dependent production of indigo dye have been reported. Some CYP102A1 BM3 mutants were generated through rational design and sitesaturation mutagenesis for indole hydroxylation activity [9]. A colorimetric colony method, based on the blue pigment generated from indigo, was utilized as a high-throughput screening assay to find active variants with enzymatic activity towards aromatic compounds [19]. Until now, indigo production using native CYP enzymes in recombinant microbial hosts has never been reported. A few attempts have been made using mutant CYPs [20,21]. In this report we demonstrate, for the first time, indigo production using a recombinant E. coli strain producing the wild-type CYP102A\_scat enzyme isolated from Streptomyces cattleya (S. cattleya). Furthermore, increased indigo production was achieved by optimal feeding of indigo precursors such as glucose, indole, and L-tryptophan.

#### 2. Materials and methods

#### 2.1. Chemical reagents

All chemical reagents used in this study were of analytical grade or higher. Indigo, indole, glucose, and L-tryptophan were purchased from Sigma-Aldrich Korea (Suwon, South Korea).

#### 2.2. Phylogenetic analyses

Amino acid sequences were aligned using the *ClustalW2* program *via* the European Bioinformatics Institute website (http:// www.ebi.ac.uk/Tools/msa/clustalw2/). Alignments were visualized with *Bio-edit* software (Fig. 1).

## 2.3. Heterologous expression of the gene encoding CYP102A\_scat in *E.* coli

The DNA sequence information for CYP102A\_scat was obtained from the open-access website: (http://avermitilis.ls.kitasato-u.ac. jp/). The encoding gene was amplified from its genome and the PCR product was cloned into a pET-28a(+) expression vector (Novagen, Madison, WI). The pET-28a(+)-CYP102A\_scat plasmid was subsequently transformed into E. coli BL21(DE3), after which, the transformants were grown in LB medium containing 50 µg/ml of kanamycin at 37 °C, until an OD<sub>600</sub> of 0.8 was attained. At that point, isopropyl-  $\beta$ -D-thiogalactopyranoside (IPTG) was added to obtain a final concentration of 12.5  $\mu$ M, along with 0.25 mM  $\delta$ aminolevulinic acid as a heme precursor, after which the cells were incubated at 30 °C for 12 h. After induction, 1 mL samples of every point were collected from the cell culture and prepared for indigo production analysis. At the same time the cells were subsequently harvested by centrifugation, washed twice with ice-cold PBS buffer, resuspended in 50 mM potassium phosphate buffer (pH 7.0) and used for subsequent CYP102A\_scat protein preparation and in-vitro assay.

#### 2.4. Spectral features of CYP102A\_scat

Absorption spectra of CO-bound CYP102A\_scat enzyme samples

were measured following reduction by sodium dithionite, *via* UV/ vis spectrophotometry (scanning wavelengths from 400 to 500 nm) [22]. Briefly, binding spectra were recorded following reduction of CYP102A\_scat protein (oxidized form) by sodium dithionite (reduced form), followed in turn by the bubbling of carbon monoxide gas into the enzyme solution (CO-bound form). The protein concentration was subsequently estimated using reduced CO *versus* reduced difference spectra. Using an extinction coefficient of 91.9 mM<sup>-1</sup> cm<sup>-1</sup>, at a wavelength of 450 nm, the recombinantly produced protein content was determined [22,23].

#### 2.5. Indigo production using E. coli expressing CYP102A\_scat

Cells producing CYP102A\_scat were cultured in LB media and the whole cell reactions were initiated by adding IPTG solution (final concentration: 12.5  $\mu$ M) to the media, along with the heme precursor  $\delta$ -aminolevulinic acid (final concentration: 0.25 mM). Whole cell production proceeded at 30 °C for 48 h in a high speed incubator (200 rpm), after which the reaction was quenched by the addition of an equal volume of DMSO, followed by vigorous vortexing. The mixtures were then centrifuged at 13000 rpm for 10 min, after which the blue-colored DMSO layer containing indigo was separated. The prepared samples were structurally and quantitatively analyzed using GC/MS, <sup>1</sup>H NMR spectroscopy, and other spectroscopic techniques.

# 2.6. Structural and quantitative analysis of biosynthetically produced indigo

The fractions collected from the indigo producing cell culture were separated by centrifugation and further taken for TLC, HPLC, and <sup>1</sup>H NMR. For TLC analysis, the mobile phase was composed of chloroform: hexane: methanol (5:4:1). The indigo standard solution was prepared by dissolving 262 mg of synthetic indigo in 1 mL of DMSO and diluted up to several µM ranges using methanol. Indigo products were also separated with HPLC equipped with a C18 reverse phase column (Zorbax extend-C18 Waters, 250 mm  $\times$  4.6 mm, 3.5  $\mu$ m, Agilent, USA) and eluted at 1.0 mL/min with ACN/Water (50:50 v/v). The absorbance of the eluent was monitored at 540 nm.

For quantitative analysis of the reaction products, the absorption intensity of the extracted solution was monitored at 610 nm by UV/ vis spectrometry. The production yield of indigo was further determined using a standard calibration curve obtained using the same quantification methods with commercially available synthetic indigo. In addition, the collected indigo product was analyzed by <sup>1</sup>H NMR spectroscopy, and the resulting chemical shifts were compared to that of synthetic indigo.

#### 3. Results and discussion

#### 3.1. Self-sufficient CYP102A\_scat and their sequence analysis

Cytochrome P450 (CYP) enzymes are generally classified into several family, depending on their protein sequence identities. One of the most studied CYP family is CYP102A family, such as CYP102A1 BM3, CYP102A2, CYP102A3, CYP102A5, and CYP102A7, which are reported as long-chain fatty acid hydroxylase [24–28]. Likewise, the novel CYP that we are demonstrating in this manuscript, CYP102A\_scat, also belongs to the CYP102A family and has very high sequence identities with CYP102A enzymes. As was expected, CYP102A\_scat has several unique motifs, such as a heme binding domain, a dioxygen binding domain, and a substrate binding domain (Fig. 1). Among the CYP102A family, CYP102A\_scat has highest sequence identity with CYP102A1 BM3 (41.3%), and

(A)

	10	20	30	40	50	60	70	<u>,                                    </u>	90	100	110	120
CYP102A2 CYP102A5 CYP102A3	1 MH 1 MEH 1 MH	KETSPIPQPKTF KKVSAIPQPKTY KQASAIPQPKTY	GPLGNLPLIDK - GPLGNLPLIDK - GPLKNLPHLEK -	DKPTL: DKPTL: EQLSQ:	SLIKLAEEQO SFIRIAEEYO SLWRIADELO	SPIFQIHTPA SPIFQIQTLSI SPIFRFDFPG	GTTIVVSGHE DTIIVVSGHE VSSVFVSGHN	LVKEVCDEE LVAEVCDET ILVAEVCDEK	RFDKSIEGAL RFDKSIEGAL RFDKNLGKGL	EKVRAFSGDO AKVRAFAGDO QKVREFGGDO	SLFTSWTHEP SLFTSETHEP SLFTSWTHEP	WRKAHNILMPTFSQ111 WKKAHNILMPTFSQ112 WQKAHRILLPSFSQ111
CYP102A1 CYP102A_scat	1 MSPTPHSASG	MTIKEMPQPKTF TTGAAAATPGAA	GELKNLPLLNT - SPAPPVPVADIS	DTGFGTTPIQ	ALMEIADELO QAMALAREHO	GE I FKFEAPGI GPVFRRRFGTI	RVTRYLSSQF FESLLVGSVE	AVTELCODES	RFDKNLSQAL RFVKAVGPVL	KFVRDFAGDO TNVRQ I AGDO	LF TSWTHEK BLF TAYNDE PI	WKKAHNILLPSFSQ110 WAKAHDILLPAFAL126
	130	140	150 1	60 17	0 1	180 -	190	200	210	220	230	240 250
CYP102A2 CYP102A5	112 RAMKDYHEKM	VDIAVQLIQKWA VDIAVQLVQKWA	RLNPN - EAVDVP RLNPN - ENVDVP	EDMTRLTLDT	IGLCGFNYRF	FNSYYRETPHI FNSFYRETPHI	PFINSMVRAL PFITSMTRAL	DEAMHQMQR DEAMHQLQR	LDVQDKLMVR	TKRQFRYDIG TKRQFQHDIG	RTMFSLVDSI RSMFSLVDNII	AERRANGDQDEKDL 236
CYP102A3	112 KAMKGYHSMML	DIATQLIQKWS	RLNPN - EE I DVA			FNSFYRDSQHI						AERKANPDEN I KDL 236
CYP102A_scat	127 SSMHTYHPTML	RVAKRLIAAWD	TALADGAPVDVA	DDMTRMTLDT	IGLAGFGYDF	FGSFRRGEPHI	PFVAAMVRGL	LHSQALLSR	KADDGVDHSA	ADEAFRADNA		EARRASGETGTDDL 252
	260	270	280	290	300	310	320	B) <sub>330</sub>	340	350	360	370
CYP102A2	237 LARMLNVEDP	TGEKLDDENIR	FQIITFLIAGHE	TTSGLLSFAT	YFLLKHPDKI	LKKAYEEVDR	VLTD AAF	TYKQVLELT	YIRMILNESL	RLWPTAPAFS	SLYPKEDTVI	GKFPITTN - DRISV358
CYP102A3	237 LSLMLYAKDP	TGETLDDENIR	YQIITFLIAGHE	TTSGLLSFAI	YCLLTHPEKI		VLTD · · · DTF	EYKQIQQLK	YIRMVLNETL	RLYPTAPAFS	SLYAKEDTVL	GEYPISKG - QPVTV358
CYP102A1 CYP102A_scat	235 LTHMLNGKDPE 253 LGLMLGAPHPS	TGEPLDDENIR DGTPLDAANIR	YQIITFLIAGHE NQVITFLIAGHE	TTSGLLSFAL'	YFLVKNPHVI YYLAKNPAVI	LOKAAEEAAR	VLVDPVF LWGDDPDPEF	SYKQVKQLK DYTDVGRLT	YVGMVLNEAL	RLWPTAPAFS RLWPTAAAFG	SLYAKEDTVLC SFQAV1DTVLC	GEYPLEKG-DELMV356 DGRVPMRAG-DTALV377
	380	390 40	00 410	420	430	(C) 440	450	0 4	60 47	70 4	80 4	90 500
CYP102A2	359 LIPQLHRDRD	AWGKDAEEFRPE	RF EHQDQYPH	HAYKPFGNGQI	RACIGMQFAL	LHEATLVLGM	ILKYFTLIDH	IENYELDIKO	TLTLKPGDFH	ISVQSRHQEA	AIHADVQA - AE	KAAPDEQKEKTEAK481
CYP102A5 CYP102A3	361 LIPQLHRDKD/ 369 LIPKLHRDQN/	AWGDNVEEFQPE AWGPDAEDFRPE	RF EELDKYPH RF EDPSS PH	HAYKPFGNGQI Haykpfgngqi	RACIGMQFAL RACIGMQFAL	LHEATLVMGMI LQEATMVLGL'	LLQHFELIDY VLKHFELINH	TGYELKIKE	ALTIKPGDFK	ITVKPRKTAA	I SHPTVLAPTE A I NVQRKEQAD	EDKLKNDEIKQHVQK484 DIKAETKPKETKPKH482
CYP102A1 CYP102A scat	357 LIPQLHRDKT				RACIG <mark>Q</mark> QFAL					VKAKSKKIPL	GGIPSPS-TE	EQSAKKVRKKAENAH479
	510	520	630	540	550	560	570	590	500	600	610	620
CYP102A2	482 GASVIGLNNR	PLLVLYGSDTGT	AEGVARELADTA	SLHGVRTKTA	PLNDR - I GKI	LPKEGAVVIV	TSSYNGKPPS	NAGQEVQWL	QEIKPGELEG	VHYAVFGCGE	OHNWASTYQY	PRFIDEQLAEKGAT606
CYP102A5 CYP102A3	485 TPSIIGADNLS	SLLVLYGSDTGV	AEGIARELADTA	SLEGVQTEVV	ALNDR - IGSL	LPKEGAVLIV	TSSYNGKPPS	NAGQ F VQWL	EELKPDELKG			PRYIDEQMAQKGAT609
CYP102A1	480 NT	PLLVLYGSNMGT	AEGTARDLADIA	MSKGFAPQVA	TLDSH - AGNI	LPREGAVLIV	TASYNGHPPE	NAKQFVDWL	DQASADEVKG	VRYSVFGCGE	DKNWATTYQK	PAFIDETLAAKGAE596
CTFT02A_3Cut	503 RPG	0 650	GEFAAGLADLG	670	680	690	700	710	EQAGPGAADG 720	VRYAVLGVGL 730	740	750
CYP102A2	607 RF SARGEGDV		KSMWADA I KAFG		-RSTLSLQF		RSYEASHASI	AENRELQSA	DS DRSTRH			PKNSQTNVSRILHR729
CYP102A3	600 RLTAIGEGDA	ADDFESHRESWE	NRFWKETMDAF -	-DINEIAQKE	DRPSLSITFI	LSEATETPVA	KAYGAFEGIN	LENRELQTA	AS TRSTRH	IELEIPAGKT	YKEGDHIGIL	PKNSREL VQRVL SR721
CYP102A1 CYP102A_scat	697 NIADRGEADA	SDDFEGTYEEWR AGDLAGTVRGFG	EHMWSDVAAYFN	ILD I ENSEDNK DPDSVGAVAG	STLSLQF	VDSAADMPLAI VTGGPLDALA	KMHGAFSTN\ ARHEVVAMT\	/VASKELQQP /TETGDLADL	GS ARSTRH	LEIELPKEAS VRLALPDGAT	SYQEGDHLGVI YRTGDHLAVL	PRNYEG I VNRVTAR718 PANDPAL VERAARL 748
	760	770	780 7	90 80	0 8	B10 8	320	830	840	850	860	870 880
CYP102A2	730 FGLKGTDQVT	LSASGRSAGHLP	LGRPVSLHDLLS	YSVEVQEAAT	RAQIRELAS	FTVCPPHRREI	EELS-AEG	YYQEQ I LKKR	ISMLDLLEKY	EACDMPFER -	FLELLRPL	KPRYYSI SSSPRVN851
CYP102A5 CYP102A3	722 FGLQSNHVIK	VSGS-AHMAHLP	MDRPIKVVDLLS	SYVELQEPAS	RLQLRELAS	YTVCPPHQKE	LEQLVSDDG	YKEQVLAKR	LTMLDFLEDY	PACEMPFER -	···FLALLPAL	KPRYYSISSSPLVA854
CYP102A1 CYP102A scat	719 FGLDASQQIR 749 LGADPDTVLG	LEAEEEKLAHLP VRARRPGRGTLP	LAKTVSVEELLC	Y-VELQDPVT	RTQLRAMAAH AAQIAVLADE	KTVCPPHKVEI RNPCPPEQAEI	LEALL - EKQA LKKLA	YKEQVLAK	ASVLDLVERY	PACEMKFSE - PALTGRLDWP	FIALLPSI	RPRYYSISSSPRVD839 RIRHYSVSSSPAVS863
_	890	900	910	920	930	940	950	960	970	980	990	1000
CYP102A2	852 PRQASITVOV	VRGPAWSGRGEY	RGVASNDLAERO		TPESRFQLP	KDPETPIIMV	GPGTGVAPFF	GFLQARDVL	KREGKTLGEA	HLYFGCRN-D	RDFIYRDELE	RFEKDGIVTVHTAF976
CYP102A5 CYP102A3	855 HNRLSITVGV 844 ANIVSMTVGV	VNAPAWSGEGTY VKASAWSGRGEY	EGVASNYLAQRH RGVASNYLAELN	INKDEIICFIR ITGDAAACFIR	TPQSNFELPH TPQSGFQMPN	KDPETPIIMV(	GPGTG I APFF GPGTG I APFF	GFLQARRVQ	KQKGMNLGQA	HLYFGCRHPE	EKDYLYRTELE DHDDLYREELD	NDERDGLISLHTAF980
CYP102A1	840 EKQASITVSV	VSGEAWSGYGEY	KGIASNYLAELO	EGDTITCFIS	TPQSEFTLP	KDPETPLIMV	GPGTGVAPF	GFVQARKQL	KEQGQSLGEA	HLYFGCRSPH	IEDYLYQEELE	NAQSEGIITLHTAF 965
CTPT02A_3Cat	1040	4000 4000	BGHEHKVR		PUREAFRIA		AGIGLAPP		NSAGRELAPA	LIFGUNPE		GAEAAGAVSLRPAP983
CYP102A2	977 SRKEGMPKTY	VQHLMADQADTL	ISILDRGGRLYV	CGDGSKMAPD	VEAALQKAY	QAVHGTGEQE			AG I			106'
CYP102A5 CYP102A3	981 SRLEGHPKTY	VQHLIKQDRINL	I SLLDNGAHLY I	CGDGSKMAPD	VEDTLCQAY	RETHEVSER		EGRYGKDV	AGI			106
CYP102A1	966 SRMPNQPKTY	VQHVMEQDGKKL	IELLDQGAHFYI	CGDGSQMAPD	VEATLMKSY	ADVHQVSEAD		KGRYAKDV	AG -			1054
CTP102A_scat	984 SAAPDGDVRF	VQHRIAAEADEV	WSLLKGGARVYV	CGDGSRMAPG	VREAFTALY	ASRTGATAEQ	AAGWLADLVA	RGRYVEDVY	AAG			1060

Fig. 1. Multiple sequence alignment of heme domain and reductase domain of the self-sufficient CYP, CYP102A\_scat, with other CYP102A family members. (A) Key amino acid residues in the CYP102A\_scat substrate binding pocket and that affect the opening of the substrate access channel; (B) the mechanistically important I-helix region; and (C) the heme-binding region. Distinctive differences in the active sites are indicated with red boxes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lowest identity with CYP102A5 (39.1%). Since CYP102A enzymes exhibit very narrow substrate specificities for long-chain fatty acids, CYP102A\_scat was also expected to have fatty acid hydrox-ylation activity. Consequently, an attempt was made to characterize the enzymatic activity of CYP102A\_scat following its heterologous production in *E. coli* BL21(DE3).

#### 3.2. Preparation of recombinant E. coli BL21(DE3) producing a selfsufficient CYP102A\_scat

Recombinant *E. coli* cells capable of expressing the gene encoding CYP102A\_scat were prepared as demonstrated in materials and method section. There was a good yield of protein. Recombinant CYP102A\_scat has an apparent mass of 101.2 kDa, according to an SDS-PAGE analysis (Fig. 2A). The presence of an activated CYP fraction was confirmed by a CO (99.99%) binding assay; CYP102A\_scat was reduced by sodium dithionite, and after that, the sample had CO gas bubbled through it. Scanning the absorption wavelength from 400 to 550 nm in a UV/vis spectrophotometer revealed that the highest absorbance peak occurred at 450 nm, which is a characteristic feature of CYP enzymes (Fig. 2C, dotted line). In the initial stages of CYP102A\_scat enzyme characterization, the media changed color, becoming blue when CYP102A\_scat was produced (Fig. 2B). At this stage, we did not realize that the color arose from indigo dye. After several repeated rounds of careful CYP102A\_scat production, ensuring definite autoclave sterilization, it was finally proved that the blue pigment arose from the heterologous CYP102A\_scat production, not from any contamination source. The unknown pigment was finally identified upon extraction of the blue pigments from the growth media, followed by structural analyses using <sup>1</sup>H NMR spectroscopy.

#### 3.3. Structural identification of the unknown blue pigment

The unknown blue pigment was extracted from 1 mL of growth medium using an equal volume of DMSO. After 10 min of centrifugation, the deep blue supernatant was separated and prepared for analysis by TLC, HPLC, and <sup>1</sup>H NMR spectroscopy. A verified sample of synthetic indigo was purchased from Sigma-Aldrich Korea and analyzed alongside the unknown pigment, to aid in the precise analysis of its structure. The product separated by TLC reveled a single indigo product while indigo standard has two spots



**Fig. 2.** Production of functional CYP102A\_scat. (A) SDS-PAGE results; M: marker, T: total proteins, S: proteins in the soluble fraction, I: proteins in the insoluble fraction. (B) Formation of indigo dye, following production of CYP102A\_scat, adding various concentrations of certain intermediates involved in tryptophan synthesis. (C) Oxidized, reduced, and CO-bound spectral features of CYP102A\_scat.

of indigo and indirubin in Fig. 4A. The single spot exactly matched with the blue indigo spot separated from indigo standard. Also, HPLC separation of produced blue pigment was attempted and resulted in a single peak with same retention time of corresponding standard indigo as well (Fig. 4A). <sup>1</sup>H NMR spectroscopic analysis was attempted to further the detailed identification. The assignment of <sup>1</sup>H NMR spectrum for blue pigment is shown in Fig. 4B and the <sup>1</sup>H NMR spectrum gave peaks at  $\delta$  6.551 ppm,  $\delta$  6.990 ppm, and  $\delta$  7.496 ppm corresponding to aromatic ring protons, while at  $\delta$  10.465 ppm corresponding to aromatic C-NH. The chemical shifts observed were well-matched with those of the reference, synthetic indigo. In GC/MS analysis, a 262 m/z reading for the isolated blue compound demonstrated that its molecular weight matched that of the verified, synthetic indigo (data not shown). Using these structural techniques, the blue pigment generated by heterologous production of CYP102A\_scat was finally identified as indigo dye.

# 3.4. Optimization of indigo production by adding synthetic precursors of indigo

In order to quantify the production of indigo by the cells producing CYP102A\_scat, the time-dependence of indigo production was monitored in LB medium. The highest indigo production (~1.0 g/L) was recorded around 21 h following the induction by IPTG (Fig. 5). The glucose concentration-dependence of indigo production was investigated, in order to examine the effects of providing an additional carbon source. Media containing three different concentrations of glucose were examined for indigo production. Unanticipated production profiles resulted. The higher the concentration of glucose added, the lower was the production of indigo (Fig. 5). The increased flux from glucose did not improve indigo production.

*E. coli* cells synthesize the amino acid tryptophan through the shikimate pathway. Glucose is converted into shikimate, then indole, and finally into tryptophan, by the action of several enzymes (Fig. 3). The most important intermediate in the synthesis of tryptophan is indole, which is an intercellular signaling molecule. Indole has been reported to regulate certain physiological processes of cells, such as sporulation, biofilm formation, plasmid stability, and virulence. In order to investigate the dependency of indigo production on the extracellular feeding of indole, different concentrations of indole were added to the growth media and the resulting indigo production was monitored over time. Interestingly, the higher the indole concentration in the media, the higher the indigo production that was achieved. With the provision of 20  $\mu$ M of additional indole, the indigo production reached 2.0  $\pm$  0.1 g/L



Fig. 3. Synthetic pathways of indigo production in *E. coli* cells. Cells utilize the shikimate pathway for indole and tryptophan synthesis. Intracellular indole becomes intercepted by heterologously produced CYP102A\_scat, which catalyzes C-5 specific hydroxylation of indole, resulting in indigo production.



Fig. 4. Separation and structural analysis of produced indigo. (A) Separation of generated indigo by TLC and HPLC. (B) Structural analysis of generated indigo by <sup>1</sup>H NMR analysis.



**Fig. 5.** Indigo production over 29 h in LB medium ( $\Diamond$ line, without additional glucose) and the effect on indigo production of adding additional glucose to the LB medium ( $\bigcirc$ ; 20 mg/L,  $\blacksquare$ ; 100 mg/L,  $\bullet$ ; 200 mg/L).

which amounts to a doubling of the production yield (Fig. 6A). Along with indole supplementation, the effect of additional Ltryptophan in the LB media on the production of indigo was also investigated and resulted in increased indigo production up to



**Fig. 6.** (A) Indigo production with various concentrations of supplemental indole in the LB medium. ( $\diamond$ ; 0 µM,  $\bigcirc$ ; 1 µM,  $\blacksquare$ ; 10 µM,  $\blacklozenge$ ; 20 µM), (B) Indigo production with various concentration of supplemental L-tryptophan in the LB medium ( $\diamond$ ; 0 mM,  $\bigcirc$ ; 0.01 mM,  $\blacksquare$ ; 0.1 mM,  $\blacklozenge$ ; 0.2 mM), (C) Indigo production with various concentration of supplemental indole and 0.2 mM of L-tryptophan in the LB medium ( $\diamond$ ; 0 mM,  $\bigcirc$ ; 0.01 mM,  $\blacksquare$ ; 0.1 mM,  $\blacklozenge$ ; 0.2 mM of Interpret Protocomental indole and 0.2 mM of L-tryptophan in the LB medium ( $\diamond$ ; 0 mM,  $\bigcirc$ ; 0.01 mM,  $\blacksquare$ ; 0.1 mM,  $\blacklozenge$ ; 0.2 mM of indole).

1.6  $\pm$  0.3 g/L with the addition of 0.2 mM of L-tryptophan (Fig. 6B). Finally, the combinatorial effects from both of indole and Ltryptophan supplementation was investigated by varying concentration of indole while the L-tryptophan concentration being constant for 30 h. The optimum Indigo production reached up to 3.8  $\pm$  0.1 g/L by adding 20  $\mu$ M of extracellular indole and 0.2 mM of L-tryptophan to the LB media at 25 h after IPTG induction (Fig. 6C). And the production rate of indigo was measured as 0.15 g/L/hr with the same optimized condition. To our knowledge, this is a record and the highest yield achieved so far.

#### 4. Conclusions

Here we have isolated a unique self-sufficient monooxygenase from S. cattleya, termed 'CYP102A\_scat'. The gene encoding CYP102A\_scat was cloned, and then expressed in E. coli BL21(DE3), in order to investigate its characteristic features. The amino acid sequence of CYP102A\_scat is highly similar with other CYP102A family members. Unlike other CYP102A species, CYP102A\_scat uniquely produces a blue-dye following IPTG induction in LB medium. The isolated blue pigment (1.0 g/L) was eventually identified as indigo using TLC, HPLC separation, and <sup>1</sup>H NMR spectroscopic analysis. Additional feeding of glucose did not increase the productive yield of indigo. The highest productive yield of indigo was achieved by adding both 20 µM of indole and 0.2 mM of L-tryptophan to the culture medium, which boosted indigo production to 3.8 g/L. This novel, self-sufficient CYP102A enzyme, is the first selfsufficient CYP that exhibits indole hydroxylation activity, resulting in indigo production—without having to mutate the wild-type enzyme. Further investigation of a CYP102A\_scat structural model, together with indole docking simulations extends our understanding of the indole hydroxylation activity, and these insights could be applied to other self-sufficient CYP enzymes in the future.

#### Acknowledgements

The authors have declared no conflict of interest. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (NRF-2015R1A2A2A04006014, NRF-2015M1A5A1037196). And this work was also supported by Industrial Strategic Technology Development program (No.10044604), funded by the Ministry of Trade, Industry and Energy (MOTIE/KEIT, Korea).

#### References

- Stasiak N, Kukula-Koch W, Glowniak K. Modern industrial and pharmacological applications of indigo dye and its derivatives—a review. Acta Pol Pharm 2014;71(2):215–21.
- [2] Sarangi BK, Minami Y, Thul ST. RNA-Seq analysis for indigo biosynthesis pathway genes in Indigofera tinctoria and Polygonum tinctorium. Genom Data 2015;6:212-3.
- [3] Drewson ABaV. Darstellung von Indigblau aus Orthonitrobenzaldehyd. Berichte Dtsch Chem Ges 1882;15(2):2856–64.
- [4] Bailey GC. RSP. NOTES-synthesis of indigo from fumaric acid and aniline; preparation of dibrom-succinic acid; fusion of dianilido-succinic acid to indigo. J Am Chem Soc 1922;44(1):215–6.
- [5] Ferber KH. Toxicology of indigo. A review. J Environ Pathol Toxicol Oncol 1987;7(4):73–83.
- [6] Butterworth KR, Hoosen J, Gaunt IF, Kiss IS, Grasso P. Long-term toxicity of indigo carmine in mice. Food Cosmet Toxicol 1975;13(2):167–76.
- [7] Chia MA, Musa RI. Effect of indigo dye effluent on the growth, biomass production and phenotypic plasticity of Scenedesmus quadricauda (Chlorococcales). An Acad Bras Cienc 2014;86(1):419–28.
- [8] Vuorema A, John P, Keskitalo M, Mahon MF, Kulandainathan MA, Marken F. Anthraquinone catalysis in the glucose-driven reduction of indigo to leucoindigo. Phys Chem Chem Phys 2009;11(11):1816–24.
- [9] Lu Y, Mei L. 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 2007;34(3):247–53.

- [10] Qu Y, Zhang X, Ma Q, Deng J, Deng Y, Van Nostrand JD, et al. Microbial community dynamics and activity link to indigo production from indole in bioaugmented activated sludge systems. PLoS One 2015;10(9):e0138455.
- [11] Shi S, Ma F, Sun T, Li A, Zhou J, Qu Y. Biotransformation of indole to indigo by the whole cells of phenol hydroxylase engineered strain in biphasic systems. Appl Biochem Biotechnol 2013;169(4):1088–97.
- [12] Doukyu N, Nakano T, Okuyama Y, Aono R. Isolation of an Acinetobacter sp. ST-550 which produces a high level of indigo in a water-organic solvent twophase system containing high levels of indole. Appl Microbiol Biotechnol 2002;58(4):543-6.
- [13] Choi HS, Kim JK, Cho EH, Kim YC, Kim JI, Kim SW. A novel flavin-containing monooxygenase from Methylophaga sp strain SK1 and its indigo synthesis in Escherichia coli. Biochem Biophys Res Commun 2003;306(4):930–6.
- [14] Lin GH, Chen HP, Huang JH, Liu TT, Lin TK, Wang SJ, et al. Identification and characterization of an indigo-producing oxygenase involved in indole 3-acetic acid utilization by Acinetobacter baumannii. Ant Van Leeuwenhoek 2012;101(4):881–90.
- [15] Ameria SP, Jung HS, Kim HS, Han SS, Kim HS, Lee JH. Characterization of a flavin-containing monooxygenase from Corynebacterium glutamicum and its application to production of indigo and indirubin. Biotechnol Lett 2015;37(8): 1637–44.
- [16] McDonnell AM, Dang CH. Basic review of the cytochrome p450 system. J Adv Pract Oncol 2013;4(4):263-8.
- [17] Girvan HM, Munro AW. Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. Curr Opin Chem Biol 2016;31: 136–45.
- [18] Urlacher VB, Girhard M. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. Trends Biotechnol 2012;30(1):26–36.

- [19] Park SH, Kim DH, Kim D, Kim DH, Jung HC, Pan JG, et al. Engineering bacterial cytochrome P450 (P450) BM3 into a prototype with human P450 enzyme activity using indigo formation. Drug Metab Dispos 2010;38(5):732–9.
- [20] Pengpai Z, Sheng H, Lehe M, Yinlin L, Zhihua J, Guixiang H. Improving the activity of cytochrome P450 BM-3 catalyzing indole hydroxylation by directed evolution. Appl Biochem Biotechnol 2013;171(1):93–103.
- [21] Appel D, Lutz-Wahl S, Fischer P, Schwaneberg U, Schmid RD. A P450 BM-3 mutant hydroxylates alkanes, cycloalkanes, arenes and heteroarenes. J Biotechnol 2001;88(2):167–71.
- [22] Omura T, Sato R. A new cytochrome in liver microsomes. J Biol Chem 1962;237:1375–6.
- [23] Koley AP, Buters JT, Robinson RC, Markowitz A, Friedman FK. CO binding kinetics of human cytochrome P450 3A4. Specific interaction of substrates with kinetically distinguishable conformers. J Biol Chem 1995;270(10):5014–8.
- [24] Budde M, Maurer SC, Schmid RD, Urlacher VB. Cloning, expression and characterisation of CYP102A2, a self-sufficient P450 monooxygenase from Bacillus subtilis. Appl Microbiol Biotechnol 2004;66(2):180–6.
- [25] Chowdhary PK, Alemseghed M, Haines DC. Cloning, expression and characterization of a fast self-sufficient P450: CYP102A5 from Bacillus cereus. Arch Biochem Biophys 2007;468(1):32–43.
- [26] Dietrich M, Eiben S, Asta C, Do TA, Pleiss J, Urlacher VB. Cloning, expression and characterisation of CYP102A7, a self-sufficient P450 monooxygenase from Bacillus licheniformis. Appl Microbiol Biotechnol 2008;79(6):931–40.
- [27] Choi KY, Jung E, Jung DH, Pandey BP, Yun H, Park HY, et al. Cloning, expression and characterization of CYP102D1, a self-sufficient P450 monooxygenase from Streptomyces avermitilis. FEBS J 2012;279(9):1650–62.
- [28] De Mot R, Parret AH. A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes. Trends Microbiol 2002;10(11):502-8.