Supporting information

Elucidating Cysteine-Assisted Synthesis of Indirubin by a Flavin-Containing Monooxygenase

Joonwon Kim[†,‡,§], Jeongchan Lee[†,‡], Pyung-gang Lee[‡,§], Eun-Jung Kim[‡,∥], Wolfgang Kroutil[#], Byung-Gee Kim[*,†,‡,§,∥,⊥]

[†]Department of Chemical and Biological Engineering, [‡]Institute of Molecular Biology and Genetics,[§]Institute of Engineering Research, [∥]Bio-MAX institute, [⊥]Interdisciplinary Program of Bioengineering, Seoul National University, Seoul, 08826, Republic of Korea

[#]Institute of Chemistry, Organic and Bioorganic Chemistry, University of Graz, Graz, 8074, Austria.

Corresponding author: *Byung-Gee Kim, byungkim@snu.ac.kr

Experimental procedures

Expression and purification of flavin-containing monooxygenase from *Methylophaga aminisulfidivorans*

The gene encoding the flavin-containing monooxygenase from *M. aminisulfidivorans* (KCTC12909)[1-2] was amplified using the forward primer MaFMO_NheI_F (5'-atatgctagcatggcaactcgtattgcg-3') and the reverse primer MaFMO_XhoI_R (5'-atatctcgagttaagcttctttagccacag-3') from genomic DNA of *M. aminisulfidivorans*. The PCR product was cloned into the *NheI* and *XhoI* site of pET28a, resulting pET28a::MaFMO.

The plasmid pET28a::MaFMO was transformed into *E. coli* BL21(DE3) and screened on a LB agar plate with kanamycin (30 μg/ml). Single colony was inoculated into LB broth supplemented with 0.15 % glucose and 30 μg/ml kanamycin, and grown overnight (14 hr) at 37 ℃ with shaking (200 rpm). For the scale-up process, 1 ml of the grown culture was inoculated into 50 ml of TB medium containing kanamycin (30 μg/ml) and grown at 37 ℃ with shaking (200 rpm) until OD_{600} = 0.7 ~ 0.8. Expression of MaFMO was induced by adding 0.1 mM of IPTG with incubation at 30 ℃ for 6 hr with agitation (200 rpm). The cells were harvested by centrifugation (x4,000g) for 10 min, and washed twice using PBS buffer solution. The cell pellet was resuspended in 10 ml of Tris-HCl pH 8.0 10 mM buffer solution, unless it was annotated in the context.

For purification of MaFMO, the resuspended cell pellet were disrupted by ultrasonication in ice-cold water 20 min (10s on, 12s off). The soluble fraction was harvested after centrifugation at 12,000g for 20 min at 4 ℃. MaFMO was separated from the soluble fraction using Ni-NTA His-tag purification kit (QIAGEN Korea LTD., Seoul, Korea). The soluble fraction was loaded into a column pre-filled with Ni-NTA resin. The resin was washed once with 100 mM potassium phosphate buffer (pH 7.5) and washed twice with the same buffer containing 10 mM imidazole. MaFMO bound at the resin was eluted by loading elution buffer containing 100 mM potassium phosphate buffer (pH 7.5) and 150 mM imidazole. Imidazole in the eluted solution was removed using centrifugal ultrafiltration (cutoff MW 30,000) until imidazole concentration was less than 0.1 μM. The enzymes were stored at -80 ℃ after immediate freezing with liquid N_2. Addition of glycerol reduced the activity of MaFMO.

Expression and purification of β-glucosidase

For deglycosylation of indican, β-glucosidase from *Sulfolobus acidocaldarius* (Sa-βglu) was used[3]. The gene encoding the glucosidase was amplified, cloned and expressed as the previous study except that the gene was cloned into pET28a plasmid, resulting in pET28a::Sa-βglu. The plasmid was transformed into *E. coli* BL21(DE3) and screened on LB kanamycin agar plate. A single colony was inoculated into LB broth with 30 μg/ml kanamycin and grown for 14 hr at 37 ℃ and 200 rpm. 1ml of the cell soup was inoculated into 50ml of LB medium containing kanamycin (30 μg/ml) and grown at 37 ℃ and 200 rpm until OD_{600} = 0.6 ~ 0.8. IPTG was added to the medium at a final concentration of 0.1 mM to induce expression of Sa-βglu. The expression of Sa-βglu was at 30℃ and 200rpm for 6hr or at 18℃ and 200rpm for 18hr.

The cells were harvested with centrifugation (4,000xg, 4℃) for 15 min. The cell pellet was resuspended in 100 mM potassium phosphate buffer (pH 7.5) and disrupted by ultrasonication in ice-cold water for 10 min (5s on, 8s off). The soluble fraction was heated at 75 ℃ for 10 min and centrifuged at 15,000xg at 4 ℃ for 20 min. The supernatant was harvested and concentrated using centrifugal ultrafiltration (cutoff MW 30,000). Buffer solution exchange was performed using centrifugal ultrafiltration (cutoff MW 30,000).

Condition for whole-cell reaction

The whole-cell reaction was performed in glass vials. The resuspended cell pellet was diluted to OD_{600} = 10 in Tris-HCl buffer with the final concentration 100mM. The prescribed pH was applied when pH variation has been performed. 1 mM substrates (indoles, 2-oxindoles, isatin), 5 mM cysteine (if required), and 0.6 % glucose were added to initiate the reaction unless the changes in concentration was noted. To verify the effect of supplementation effect of 2-oxindole and isatin, the supplements were added after 30 min. The reaction pot was incubated in 30 ℃ room at 200 rpm. Yield and conversion were determined based on the additional concentration of indole. Experiments were performed in triplicate.

Condition for *in vitro* reaction

In vitro reaction were performed in 2ml eppendorf tubes. For oxidation of indole using MaFMO, the reactions were performed in 100 mM Tris-HCl pH 8.0 unless stated otherwise. 1mM of substrates (indoles, 2-oxindoles, isatin), 5 mM additives (amino acids and reducing agents if required), 0.6 % glucose, 200 μM NADP, 5 U/ml GDH, and 2 μM MaFMO were added to initiate the reaction unless it was stated.

When indoxyl was generated by deglycosylation of indican, pH of the reaction was adjusted to pH 7.0 for maximum Sa-βglu activity. To initiate combine the deglycosylation and the oxidation reactions, 2.5 mM indican and 2 μM of Sa-βglu were added to the reaction pot for oxidation. For HPLC or TLC analysis, the samples were extracted with ethyl acetate or chloroform and evaporated with fast evaporator. For LC-MS/MS analysis, the samples were filtered using 3k centrifugal filtration system and used for analysis.

Condition for non-enzymatic reaction with 2-cysteinylindoleninone

To perform non-enzymatic reaction with 2-cysteinylindoleninone *in vitro*, the reactions were performed for 6 hr with identical condition as stated above except the concentration of indole and cysteine were 5 mM and 2.5 mM, respectively, with solution buffer being 20 mM Tris-HCl pH 7.0. The concentration of cysteine was decreased to make cysteine as a limiting agent, minimizing the residual cysteine after the reaction. Extraction using organic solvents (once with ethyl acetate and once with chloroform removed hydrophobic compounds including 2-oxindole, isatin, and indigoids. After extraction, the water layer remained yellow, indicating the presence of 2-cysteinylindoleninone. pH variation was performed by adding 1 M Tris-HCl buffer with desired pH. We confirmed that the pH corresponded to the pH of the solution that was added. 1 mM of 2-oxindole or isatin was added to the solution along with 5 mM cysteine if required. After 12 hour incubation in 30 ℃, indirubin was extracted with the same volume of ethyl acetate. The extracted indirubin sample was evaporated using fast evaporator followed by addition of ethyl acetate for further analysis.

Expression of CYP102G4

Procedure for preparing *E. coli* BL21(DE3) expressing CYP102G4 was identical to the previous study.[4] The cells were harvested by centrifugation (x4,000g) for 15 min and washed twice using 100 mM potassium phosphate buffer (pH 7.5). The cell pellets were resuspended in 8 ml of the buffer solution. For the whole-cell reaction, the resuspended cell pellet was diluted to OD_{600} = 10 in 100 mM potassium phosphate buffer. 1 mM substrates (indole, 2-oxindole), 2. 5 mM cysteine, and 0.6 % glucose were added to initiate the reaction. The reaction mixture was incubated in 30 ℃ room at 200 rpm for 12 hrs.

HPLC and TLC analysis of indirubin production

The reaction sample for indirubin production was analyzed with HPLC (YoungLin, South Kroea) equipped with PFP column (100 x 4.6 mm; 2.6 μM; Thermofisher). 20 μl of reaction sample was injected for the HPLC analysis. The condition for pump and detector was as follows: Eluent A – HPLC water with 0.1 % TFA; Eluent B – HPLC ACN with 0.1 % TFA; 0.700 ml/min; 0-3 min, A 90%; 10-20 min A 5%; 21-30 min A 90%; 540 nm. The retention time for indirubin was around 11.2 min.

Silica gel TLC was used to analyze the distribution of indigoids. The dried sample was resolved in ethyl acetate and the solution was repeatedly spotted on TLC. Mixture of hexane and ethyl acetate (50:50) was used as a mobile phase. The solution was poured in the TLC chamber followed by 10 minutes of equilibration. The TLC plate was loaded immediately and removed when the mobile phase almost reached the end.

HPLC analysis for full spectra.

Due to 2-oxindole and isatin did not separate effectively, the same samples were analyzed using C18 Hypersil gold column (150 x 4.6 mm; 5.0 μM; Thermofisher) with a different gradient condition. The samples were quenched by adding 1 volume of MeOH containing phenol, which was used as an internal standard. The condition for separating 2-oxindole and isatin was as follows: Eluent A – HPLC water; Eluent B – HPLC ACN; 0.700 ml/min; 0-2 min, A 95 %; 18 min A 60 %; 19-28 min A 20 %; 28.5-35 min A 95 %; 254 nm.

Quantification of indigo

Indigo was quantified separately due to its low solubility in water. The samples from the whole-cell reactions were harvested and centrifuged at x10,000g for 20 minutes. The cell pellets containing indigo were further lyophilized to remove water component. The pellets were dissolved with DMSO and placed into the sonication water bath for 10 minutes. If the pellets with indigo remains, heating in 55 ℃ oven and the sonication were repeated. When there was no trace of indigo left in the pellet, the sample was centrifuged at x10,000g for 5 minutes. The supernatant was moved to a glass cuvette and analyzed with UV-vis spectrometry. Concentration of indigo was measured at 660 nm.

LC-MS/MS anlaysis of indoleninone

The filtrated sample was analyzed with LC connected with tandem mass spectrometry (Thermofisher TSQ) using C18 Hypersil gold column (150 x 4.6 mm; 5.0 μM; Thermofisher). 10 μl of reaction sample was injected to the LC. The condition for pump and mass spectrometry was as follows: Eluent A – HPLC water with 0.1 % formate; 400 μl/min; 0-10 min, A 85 %; 22-29 min, A 10 %; 30-35 min, A 85 %; autosampler temperature 20 ℃; column temperature 25 ℃; positive mode; capillary voltage 3.8 kV; vaporizer temperature, 300 ℃; capillary temperature, 320 ℃; sheath gas pressure, 30 psi; aux gas pressure 10 psi; Scan event 1, scan range 200-400; Scan event 2, dependent scan from scan event 1, 1st most intense ion higher than signal threshold (10^5), collision energy 20 V; Scan event 3, dependent scan from scan event 1, 2nd most intense ion higher than signal threshold (10^5), collision energy 20 V.

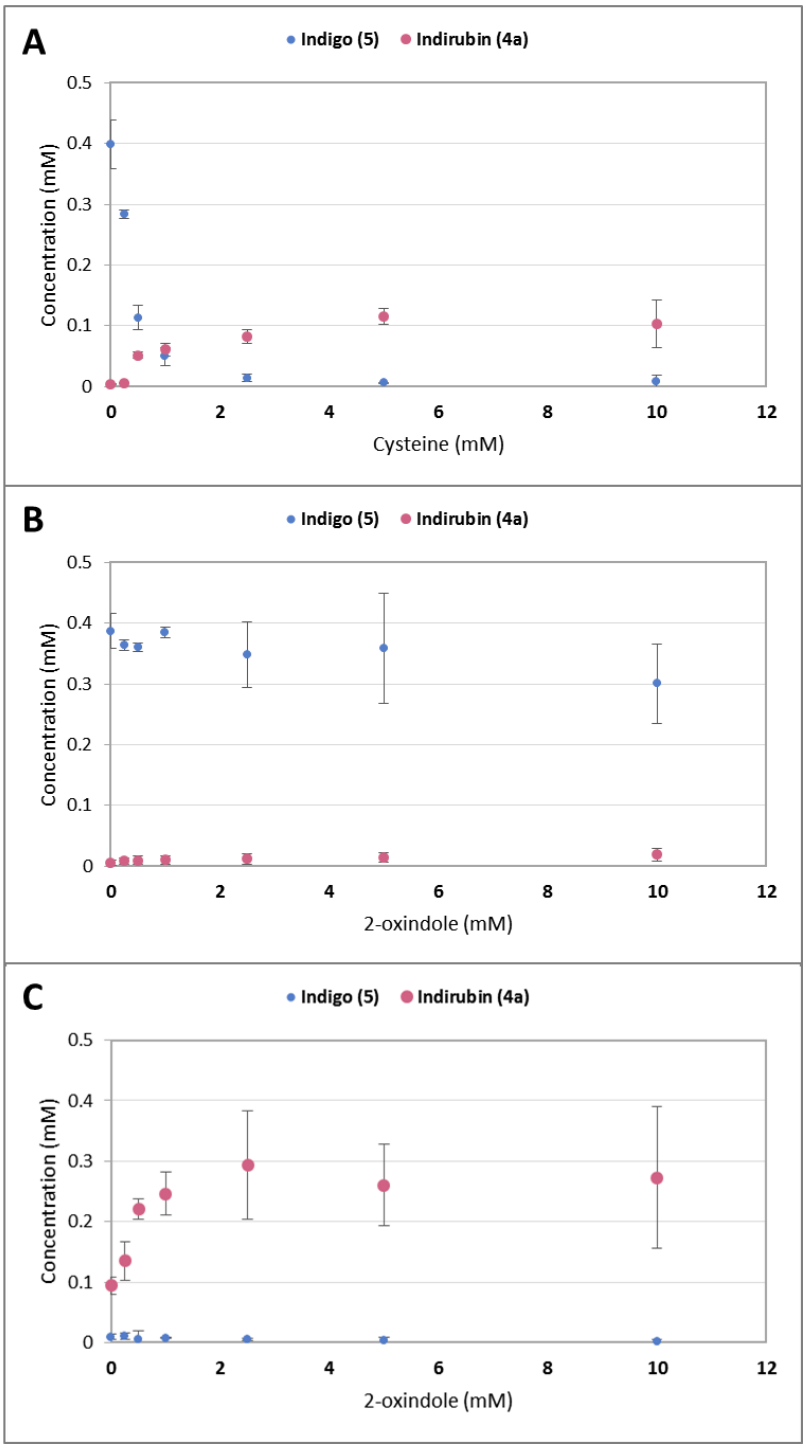

Figure S1.
Quantification of indigo and indirubin after the whole-cell reaction with 1.0 mM indole. (A) Effect of L-cysteine supplementation on the production of indirubin and indigo. (B) Effect of 2-oxindole supplementation without L-cysteine on the production of indirubin and indigo. (C) Effect of 2-oxindole supplementation with L-cysteine 5 mM on the production of indirubin and indigo.

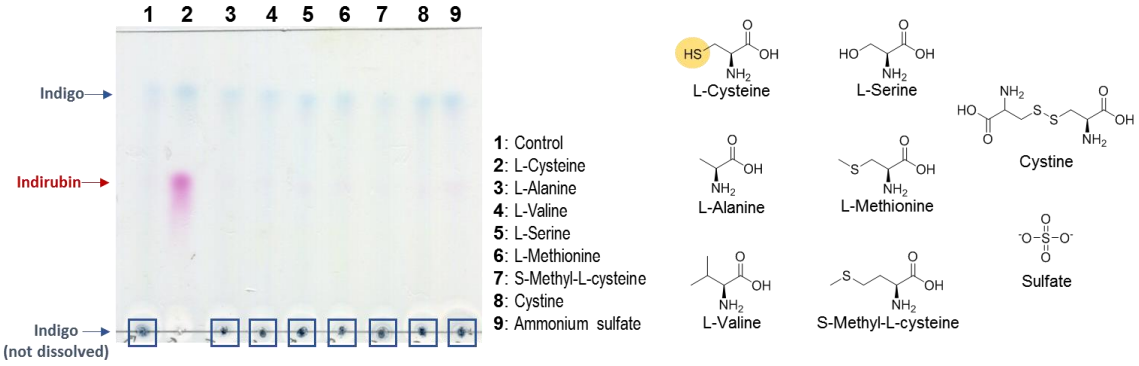

Figure S2.
TLC analysis of indirubin production with supplements including amino acids and reducing agents. The ability of additives for product selectivity change from indigo to indirubin was evaluated. 5 mM of each additives was supplemented into the reaction mixture for indole oxidation using MaFMO *in vitro*.

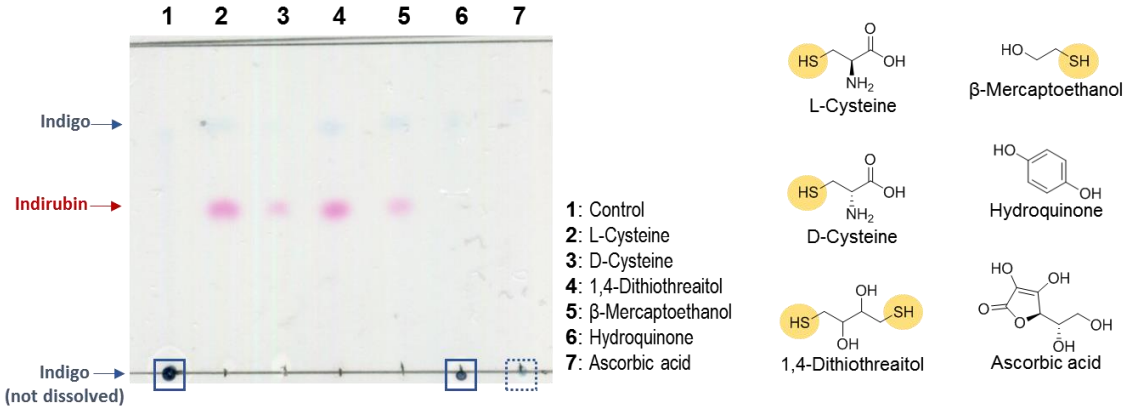

Figure S3.
TLC analysis of indirubin synthesis with supplementation of L-cysteine, D-cysteine, 1,4-dithiothreatitol (DTT), β-mercaptoethanol, hydroquinone and ascorbic acid (Figure S4).

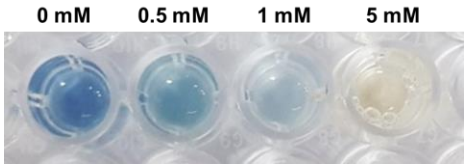

Figure S4.
Effect of ascorbic acid supplementation to the MaFMO *in vitro* reaction. Indigo production decreased according to ascorbic acid concentration, but no change in product selectivity.

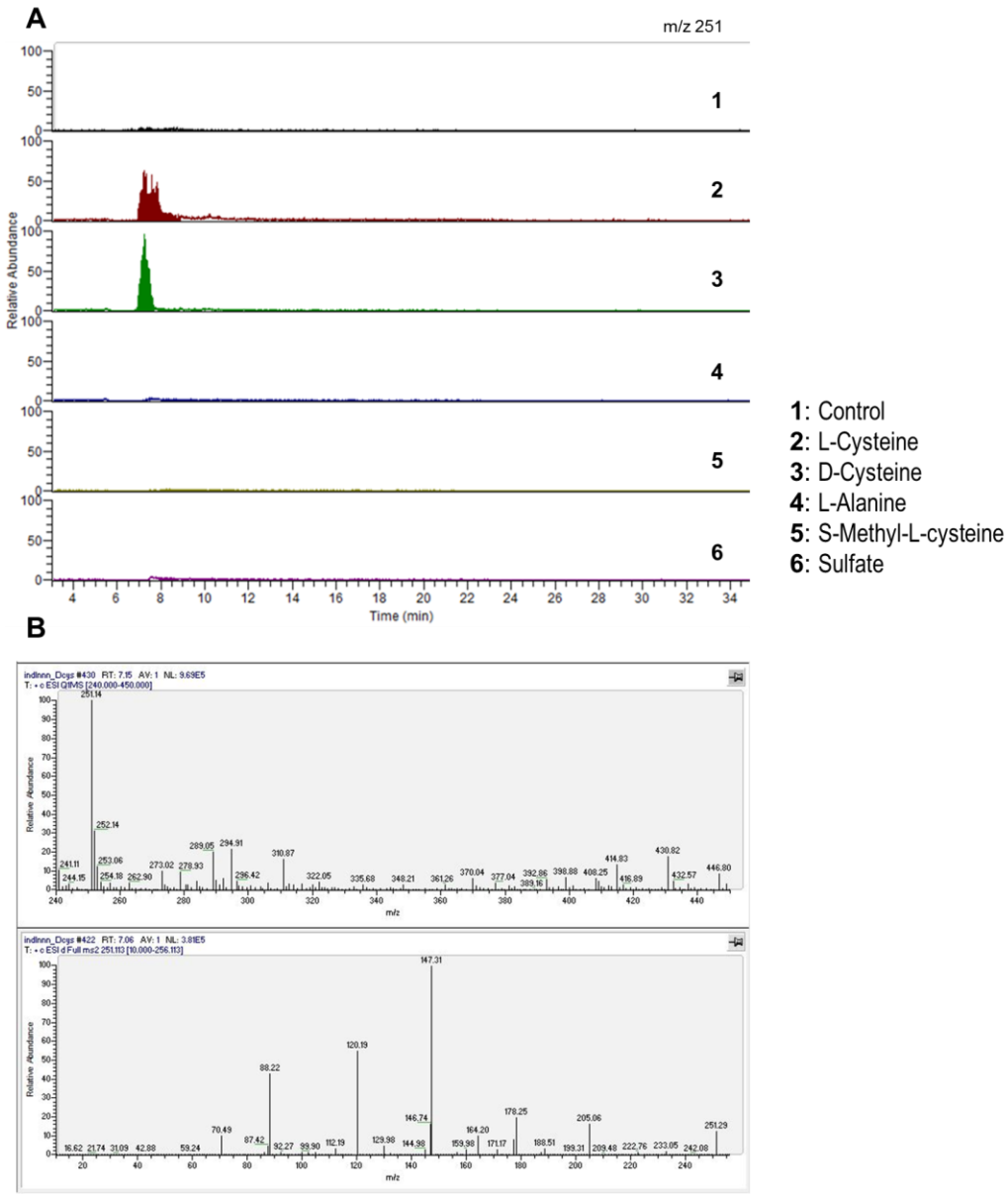

Figure S5.
(A) LC-MS chromatogram for the formation of **6a** with other supplementation including D-cysteine, L-alanine, S-methyl-L-cysteine, and sulfate. (B) A peak with the identical MS and MS/MS patterns with **6a** was generated when D-cysteine was supplemented to the MaFMO reaction, but not for the other compounds.

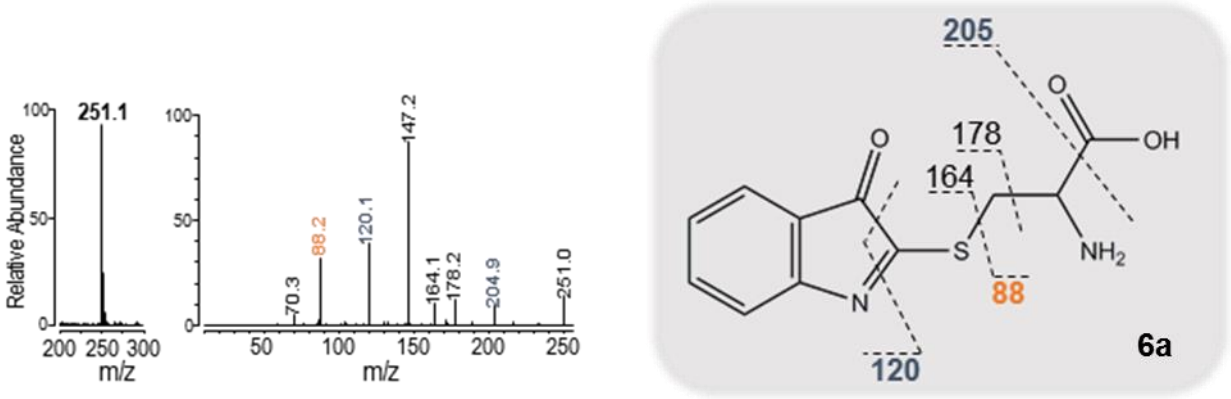

Figure S6.
MS and MS/MS analysis of 2-cysteinylinodoleninone. The labeling indicates the predicted mass fragments.

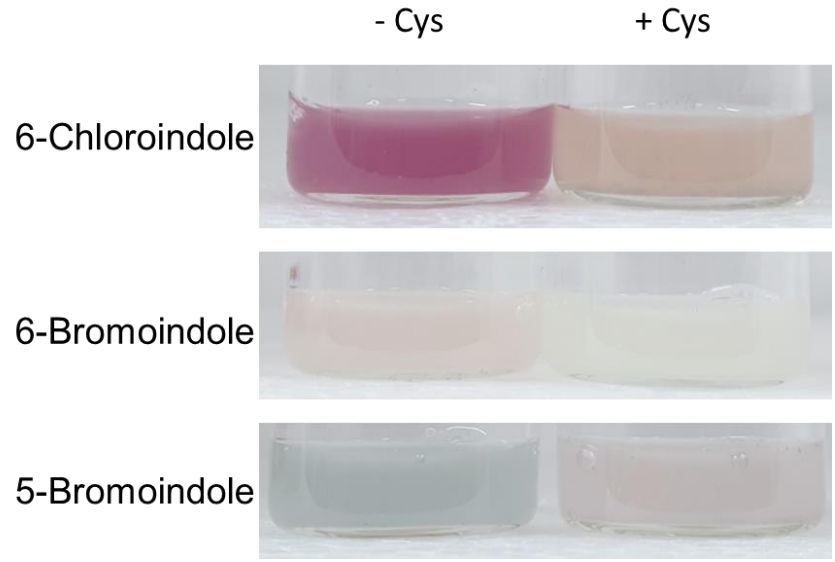

Figure S7.
Effect of L-cysteine supplementation for indole derivatives.
Oxidation of 6-chloroindole, 6-bromoindole, and 5-bromoindole by MaFMO results in production of indigo with color of red, purplish red and blue green, respectively. Addition of L-cysteine inhibited the production of according indigo.

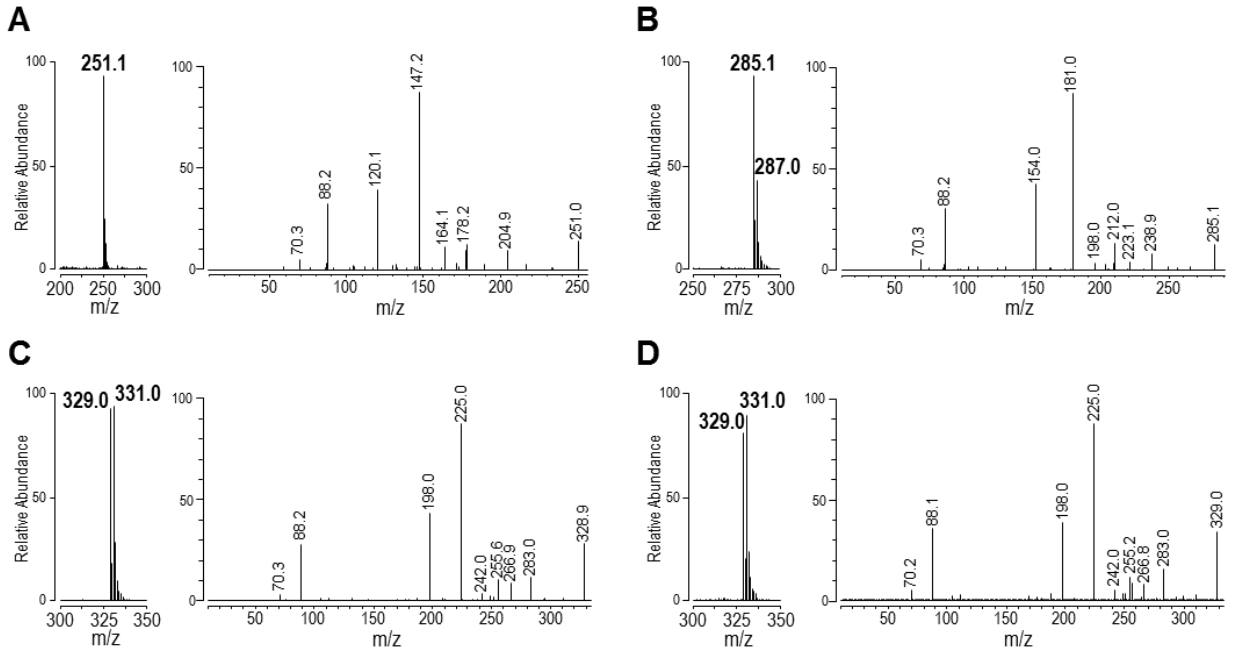

Figure S8.

MS and MS/MS analysis of 2-cysteinylindoleninone and its derivatives.

A: 2-cysteinylindoleninone; B: 6-chloro-2-cysteinylindoleninone; C:6-bromo-2-cyseinylindoleninone, D:5-bromo-2-cysteinylindoleninone. Isotope pattern of chlorine (^{35}Cl:^{37}Cl =3:1) was shown in B and that of bromine (^{79}Br:^{81}Br=1:1) was shown in C and D. The MS/MS patterns were generated by m/z 285.1, m/z 329.0, and m/z 329.0, respectively.

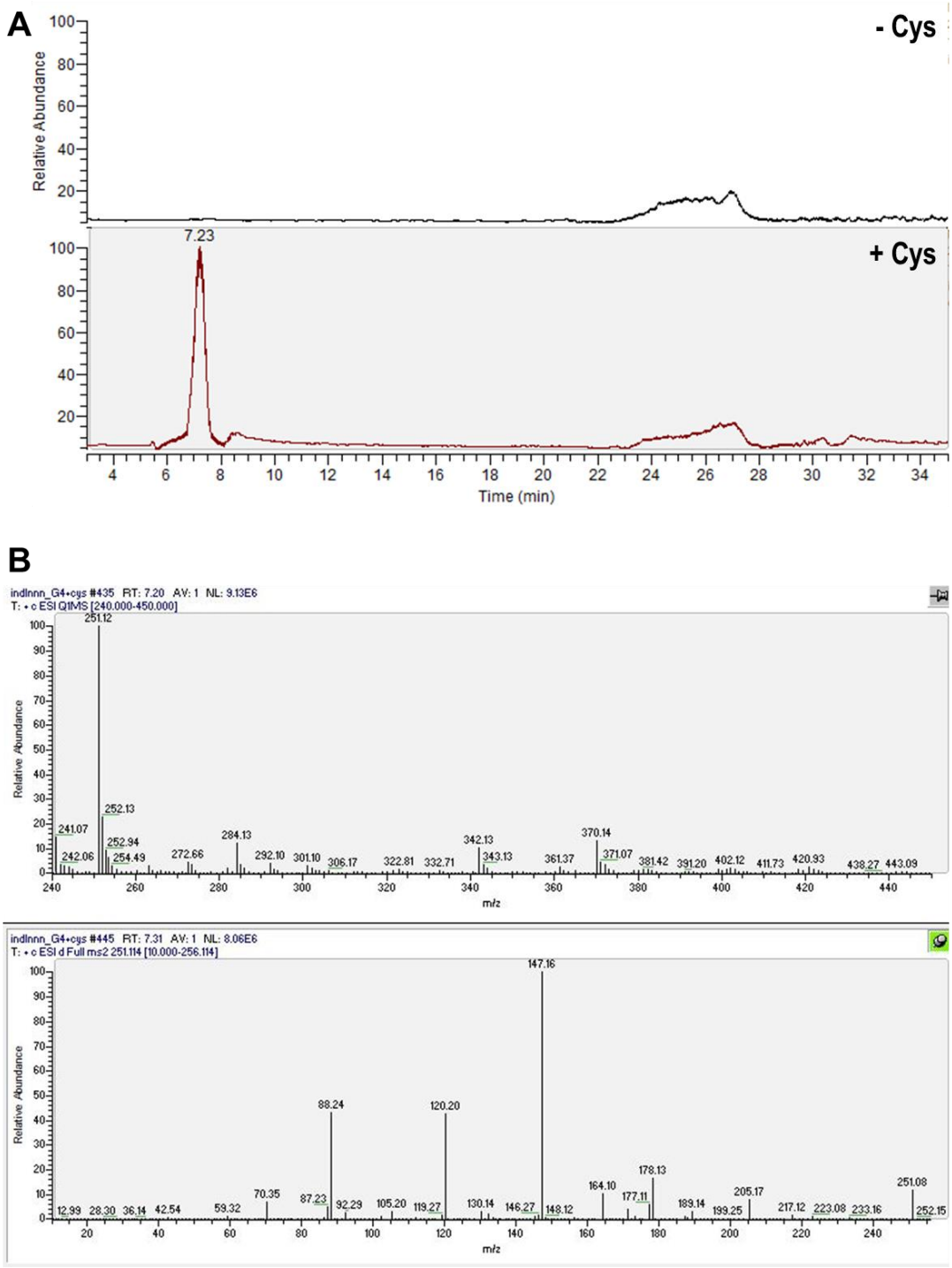

Figure S9.
(A) LC-MS chromatogram and (B) MS/MS analysis for the formation of **6a** using CYP102G4 with or without L-cysteine supplementation. Synthesis of **6a** was observed only when L-cysteine was supplemented in the CYP102G4 reaction system.

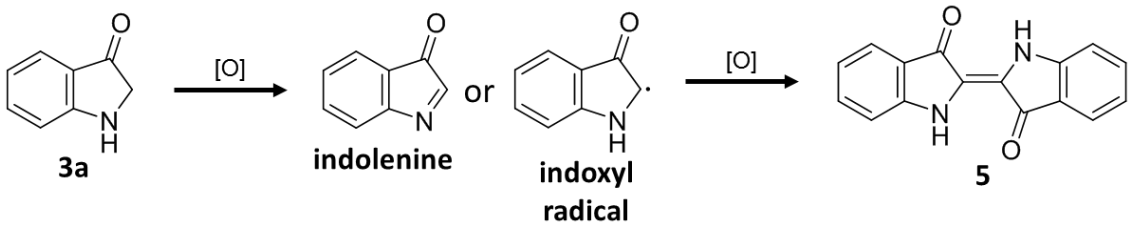

Figure S10.
Indolenine or indoxyl radical can be generated by autoxidation of indoxyl. They are reactive compounds, which react with indoxyl to form indigo.

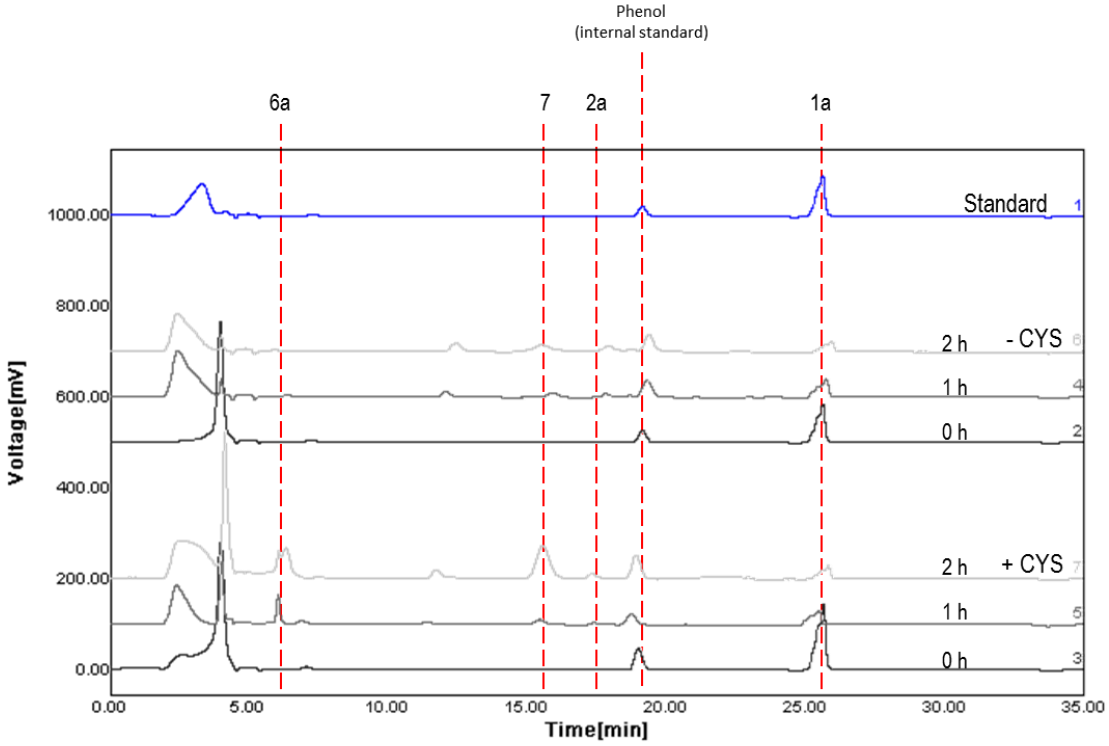

Figure S11.
Full LC spectra for *in vitro* MaFMO reaction using **1a** as a substrate. The condition with L-cysteine and without L-cysteine was compared at the time point of 0h, 1h, and 2h.

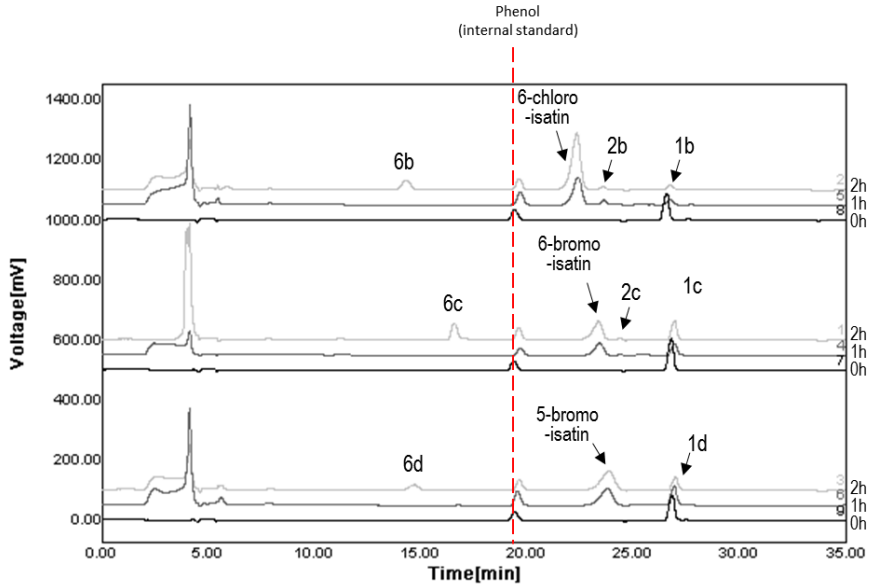

Figure S12.
Full LC spectra for *in vitro* MaFMO reaction using **1b, 1c,** and **1d** as substrates, respectively. Time points of 0h, 1hr, and 2hr were compared for each case.

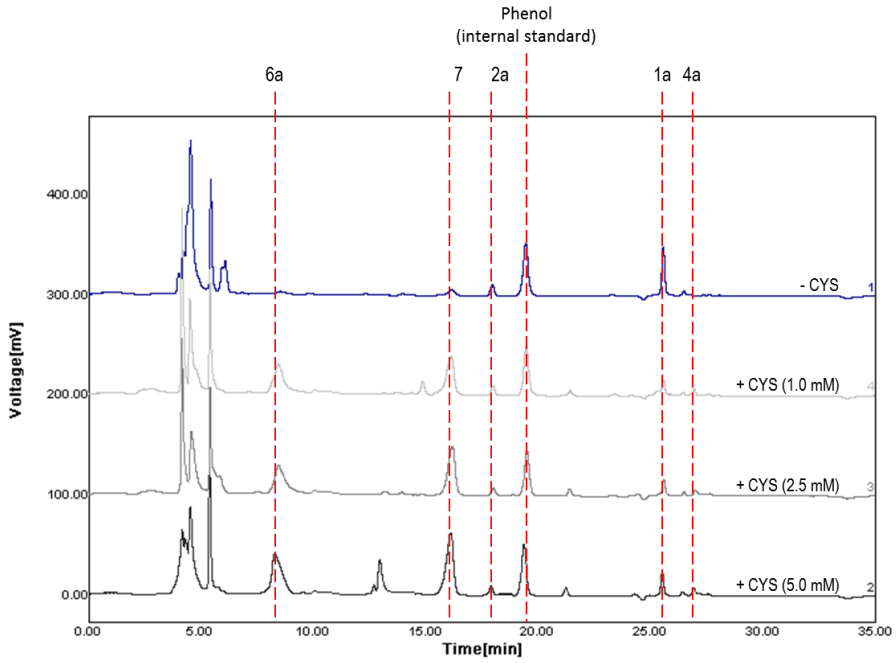

Figure S13.
Full LC spectra for the whole-cell reaction with *E.coli* expressing MaFMO using 1.0 mM of **1a** as a substrate. The conditions with variation in L-cysteine concentration were compared at the time point of 6h.

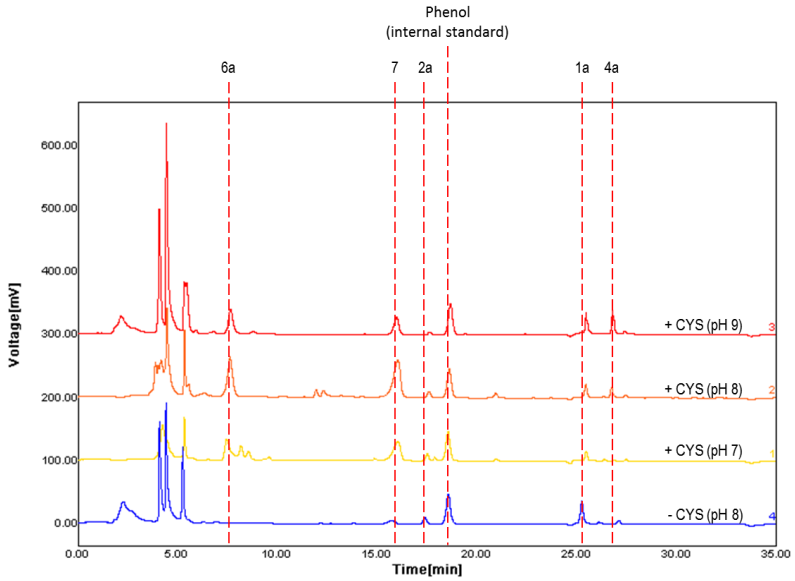

Figure S14.
Full LC spectra for the whole-cell reaction with *E.coli* expressing MaFMO using 1.0 mM of **1a** as a substrate. The conditions with pH variation were compared at the time point of 6h.

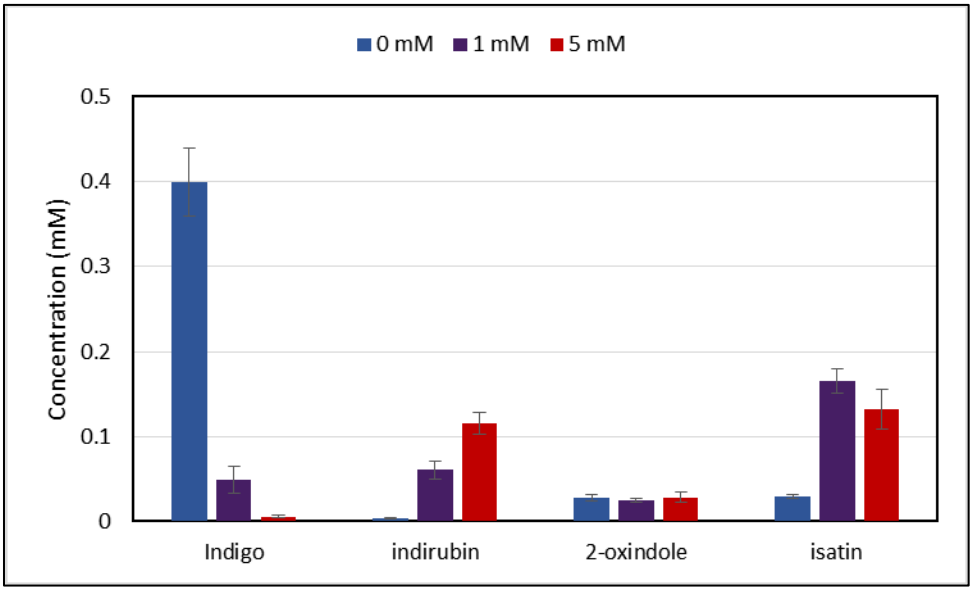

Figure S15.
Oxidized product profile for *in vivo* MaFMO reaction when 0mM, 1mM, and 5mM of L-cysteine were supplemented, respectively.

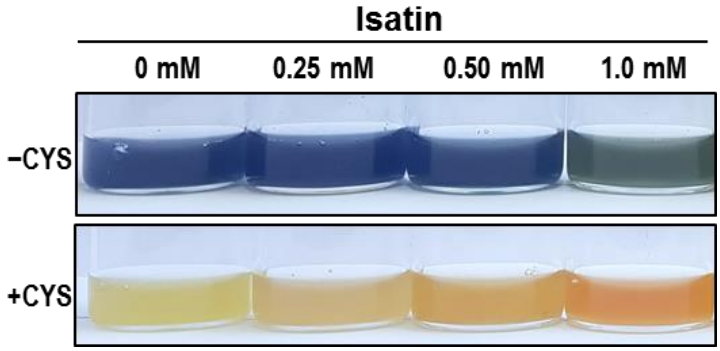

Figure S16.
Effect of isatin supplementation in L-cysteine-assisted biosynthesis of indirubin.
Isatin showed increase in red color, production of indirubin.

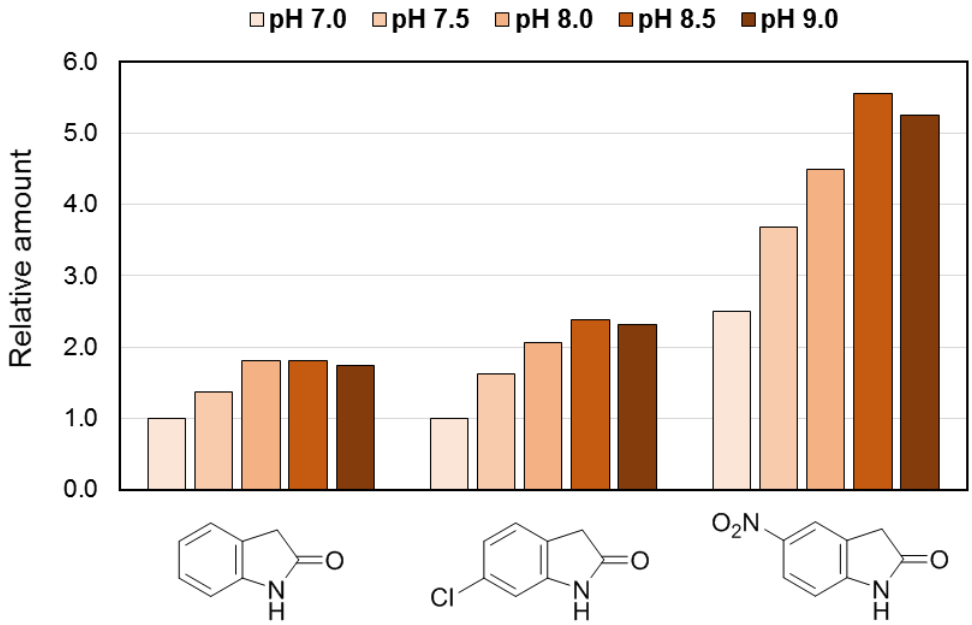

Figure S17.
pH dependency of *in vitro* indirubin synthesis from **2** (shown below figure) and 2-cysteinylindoleninone.

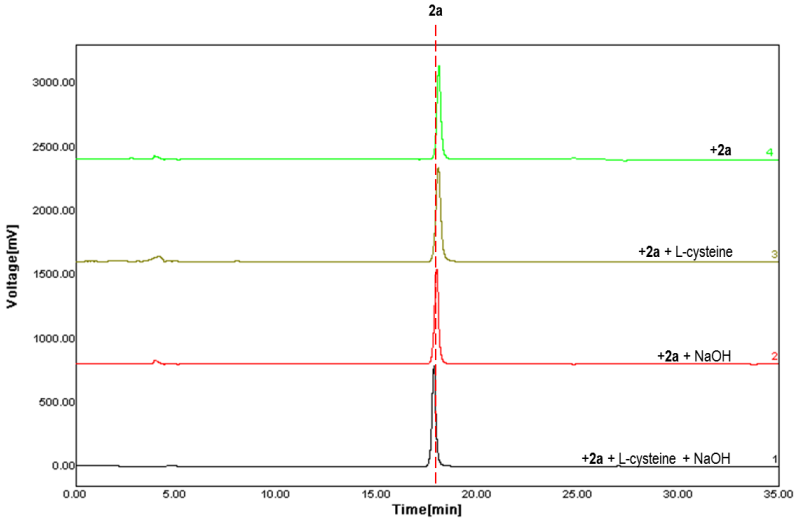

Figure S18.
Overnight reactions by mixing 1 mM 2-oxindole, 1 mM L-cysteine and 50 mM NaOH, individually or together, were carried out and no reactions were identified. Contrast to indoxyl, 2-oxindole did not react with L-cysteine.

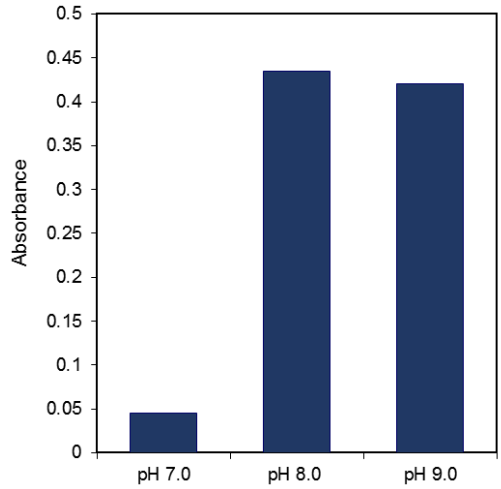

Figure S19.
pH dependency of *in vitro* indirubin synthesis from isatin and 2-cysteinylindoleninone.

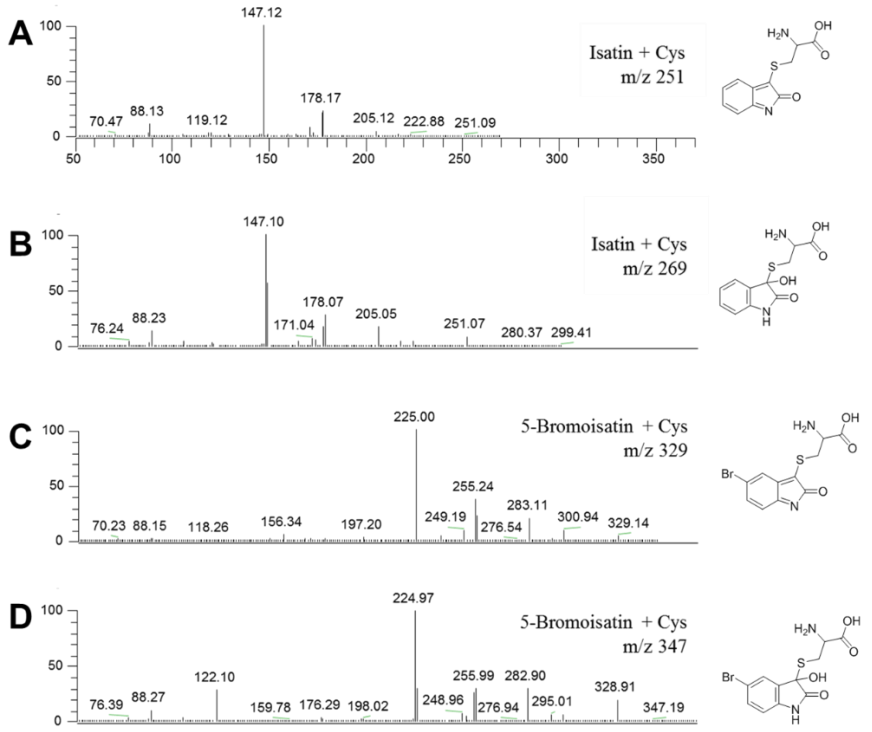

Figure S20.
An intermediate molecule detected in the mixture of isatin and L-cysteine. In the mixture of isatin and L-cysteine, the peaks with (A) m/z 251 and (B) m/z 269 were detected. In the mixture of 5-bromoisatin and L-cysteine, the peaks with (C) m/z 329 and (D) m/z 347 were detected. The following chemical structures are proposed structure of the according compounds.

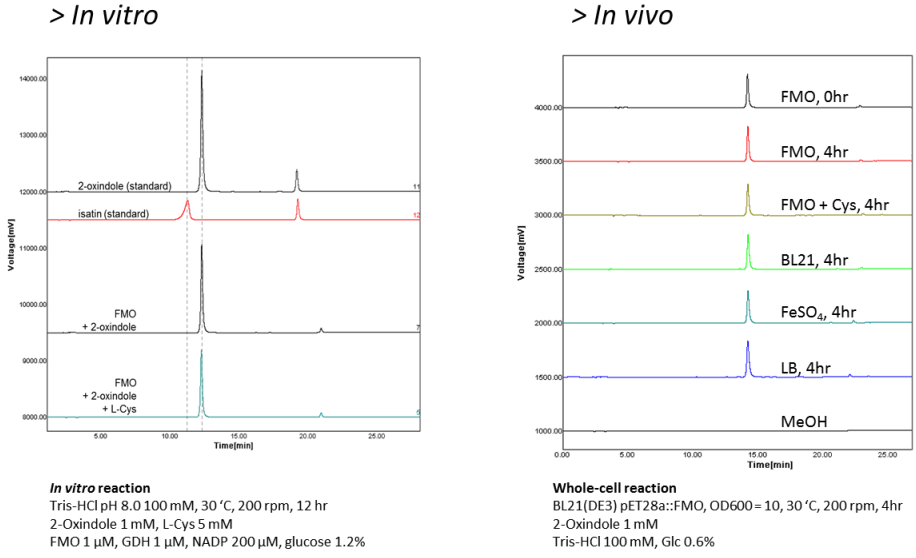

Figure S21.
Conditions to evaluate the origin of isatin from 2-oxindole.

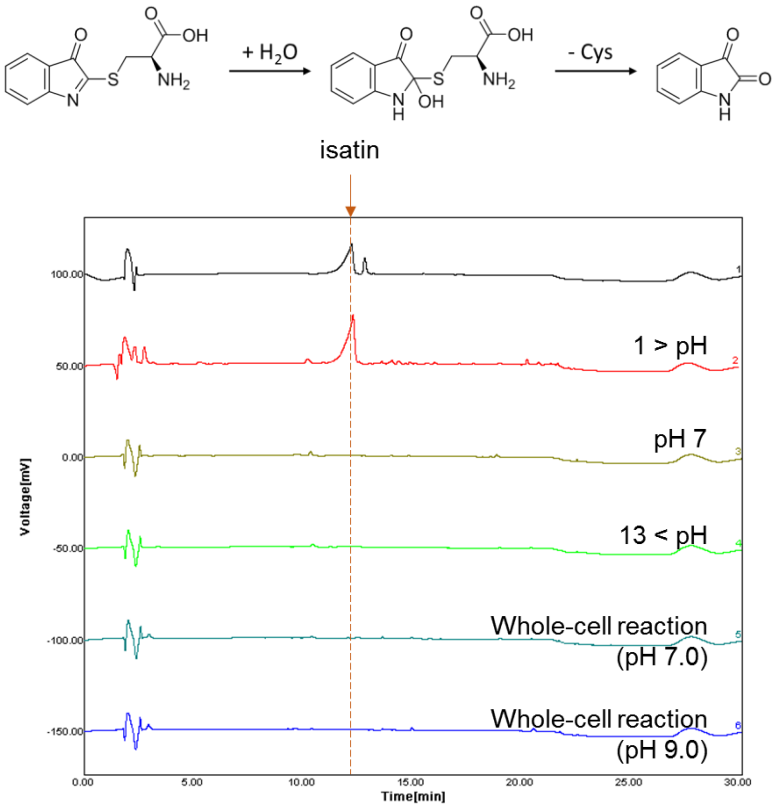

Figure S22.
Conditions to evaluate the origin of isatin from 2-cysteinylindoleninone.

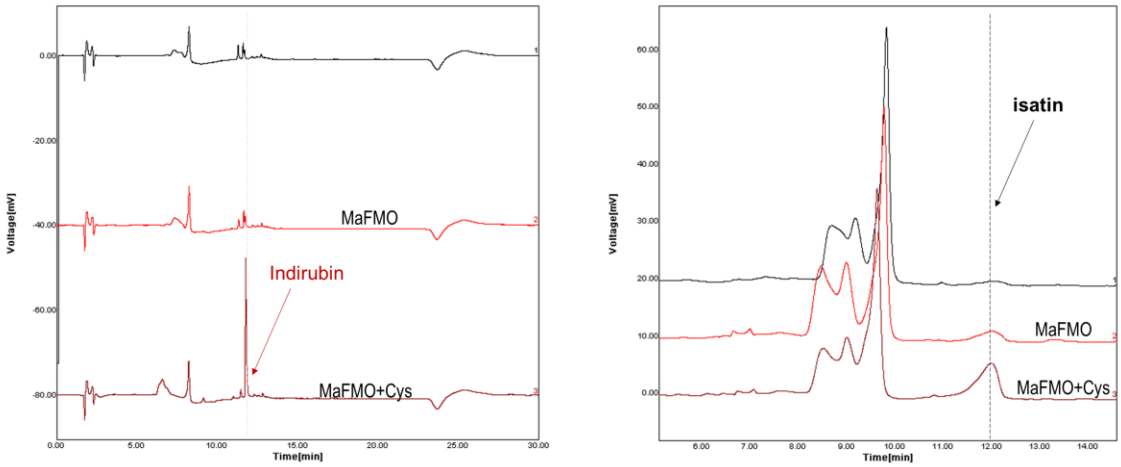

Figure S23.
Production of isatin and indirubin from indican using MaFMO according to L-cysteine supplementation.

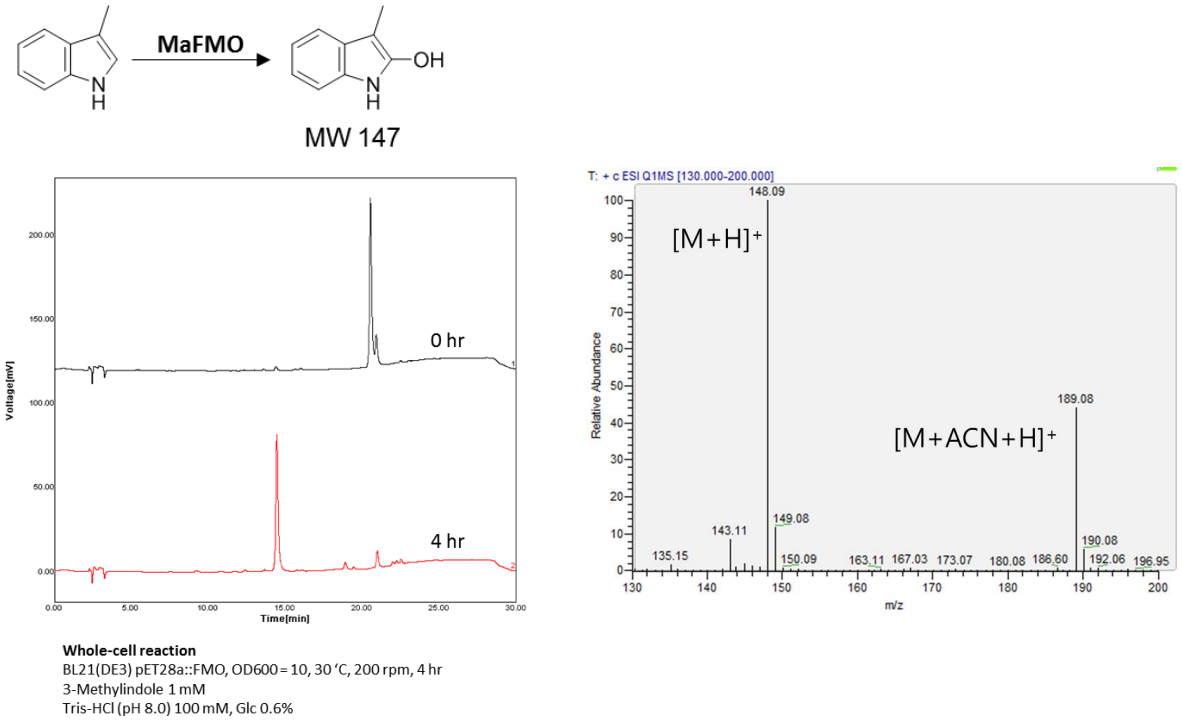

Figure S24.
HPLC and ESI-MS result of skatole oxidation.

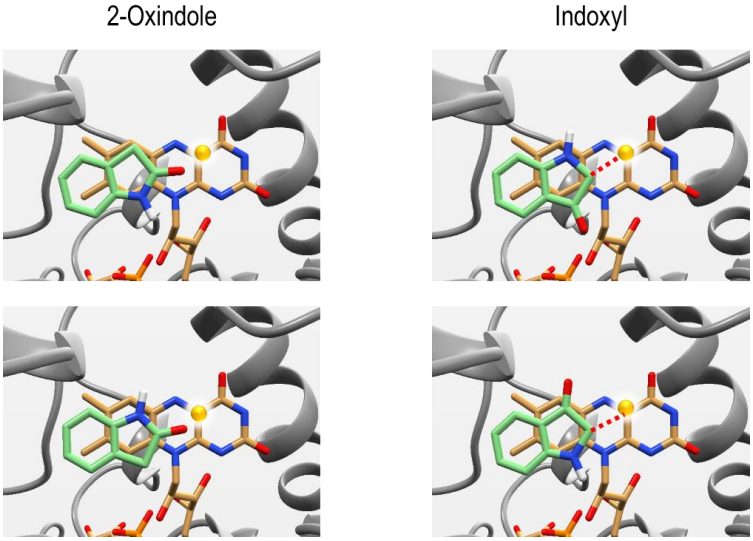

Figure S25.
Rigid docking simulation of 2-oxindole and indoxyl in MaFMO (2xvh.pdb).

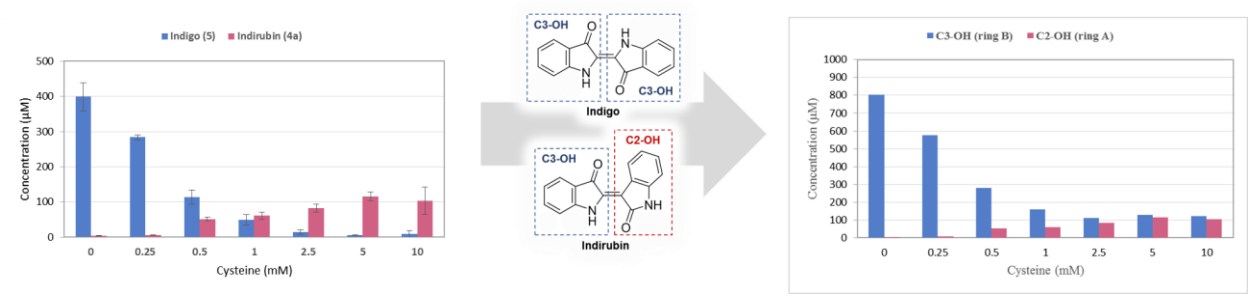

Figure S26.
Comparison of the changes in oxidized indole components according to L-cysteine supplementation. The amount of C2-OH and C3-OH indole were determined based on the distribution of indigoid produced. Indigo possesses two C3-OH indole rings while indirubin possesses one C3-OH indole and one C2-OH indole. The moiety of the ring A component (C2-OH) for indirubin synthesis significantly increases by the addition of L-cysteine.

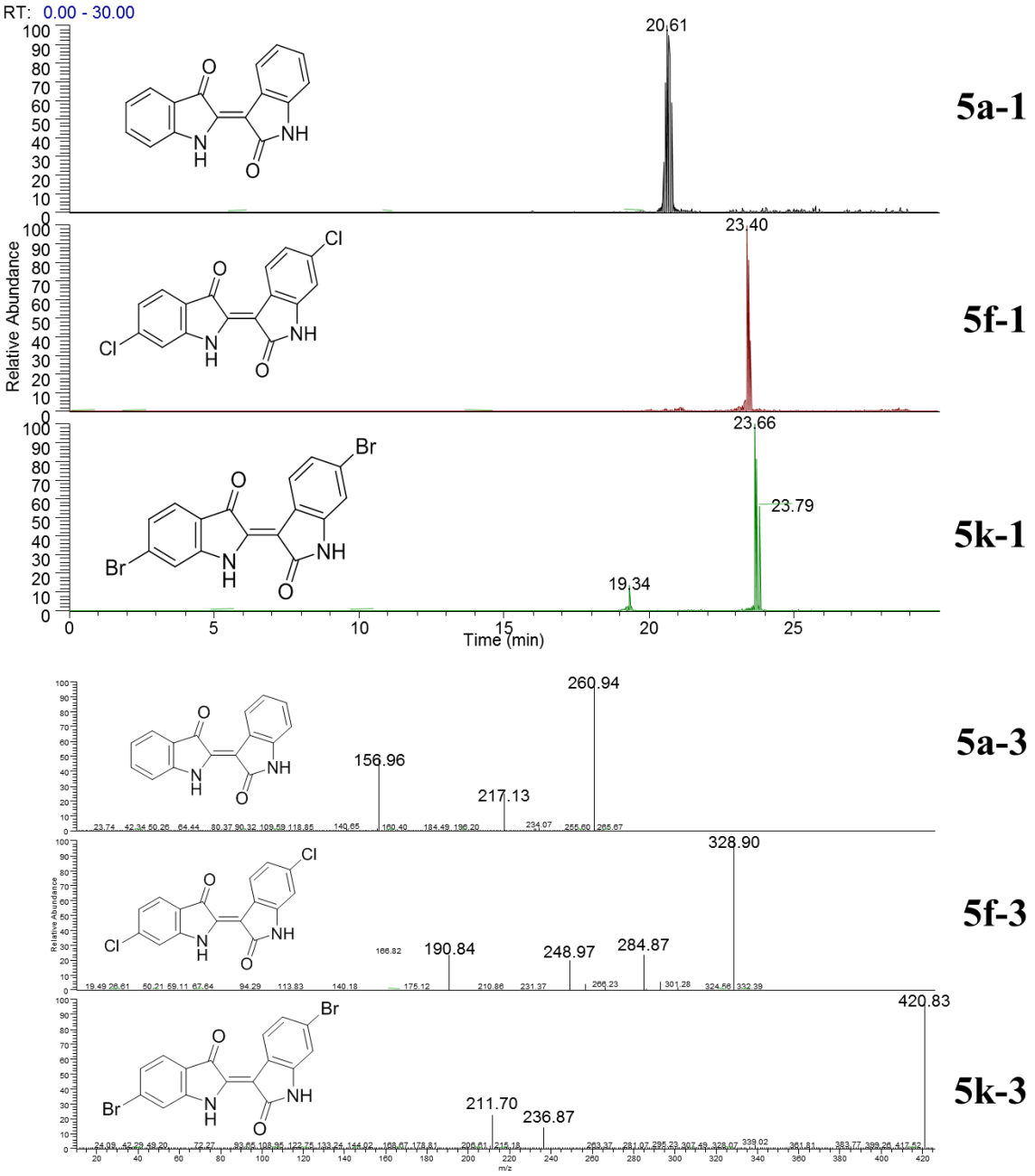

Figure S27.
LC (1) and MS/MS analysis (3) of indirubin (5a), 6,6'-dichloroindi-rubin (5f), and 6,6'-dibromoindirubin (5k).

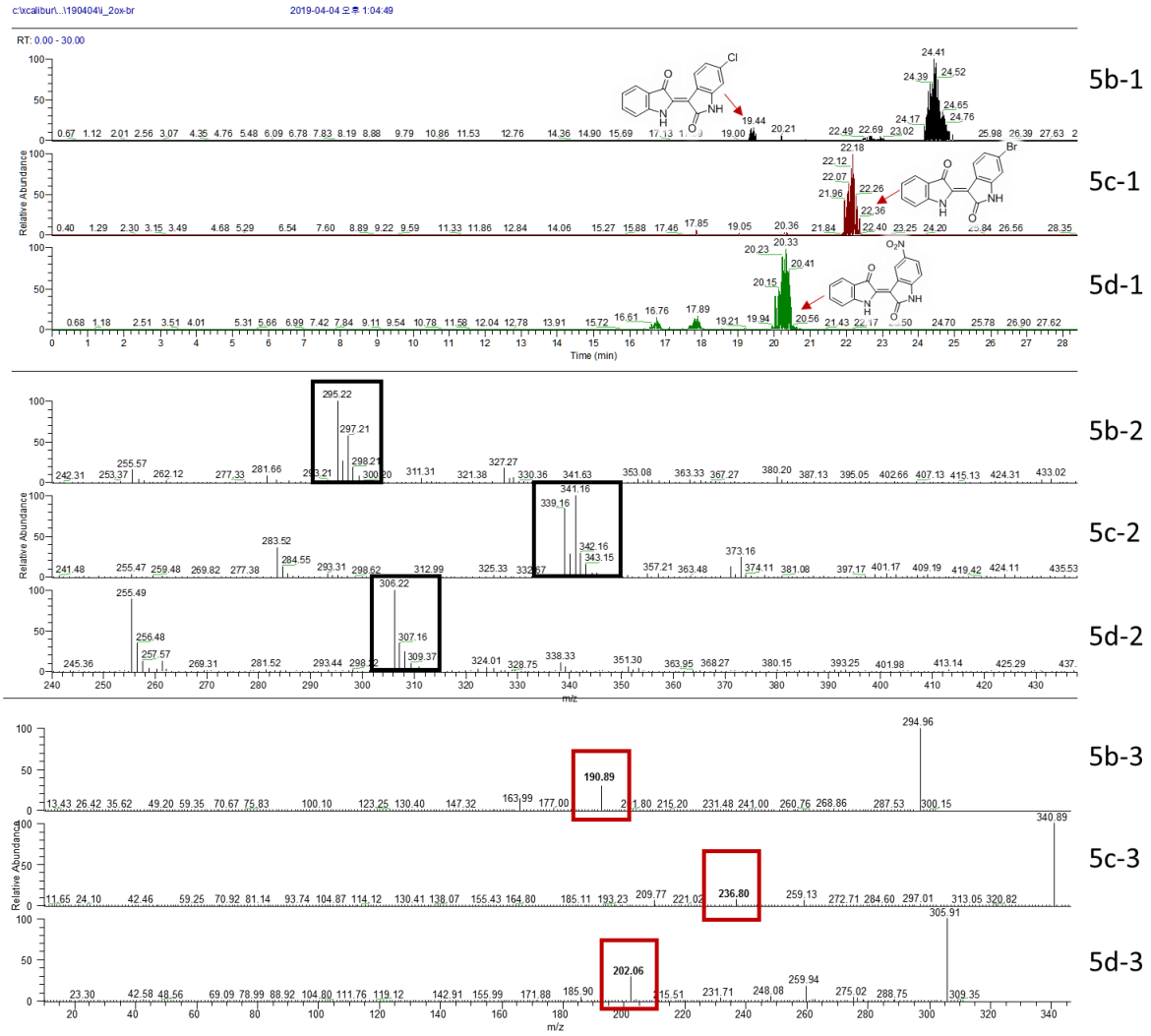

Figure S28.
LC (1), MS (2) and MS/MS analysis (3) of indirubin derivatives synthesized from indole. 6-chloroindirubin (5b), 6-bromoindirubin (5c), and 5-nitroindirubin (5d).

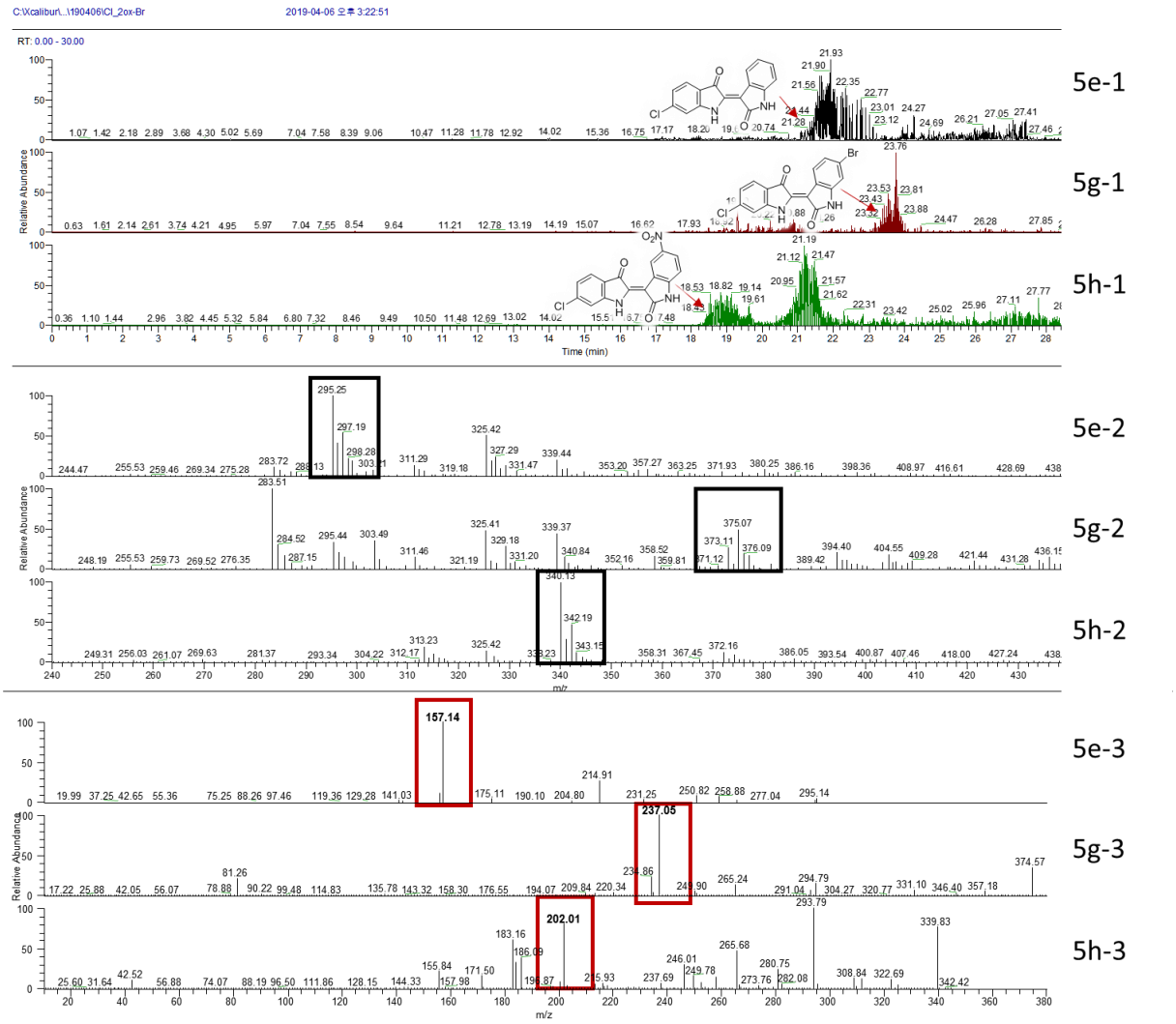

Figure S29.
LC (1), MS (2) and MS/MS analysis (3) of indirubin derivatives synthesized from 6-chloroindole. 6'-chloroindirubin (5e), 6-bromo-6'-chloro-indirubin (5g), and 6-chloro-5-nitroindirubin (5h).

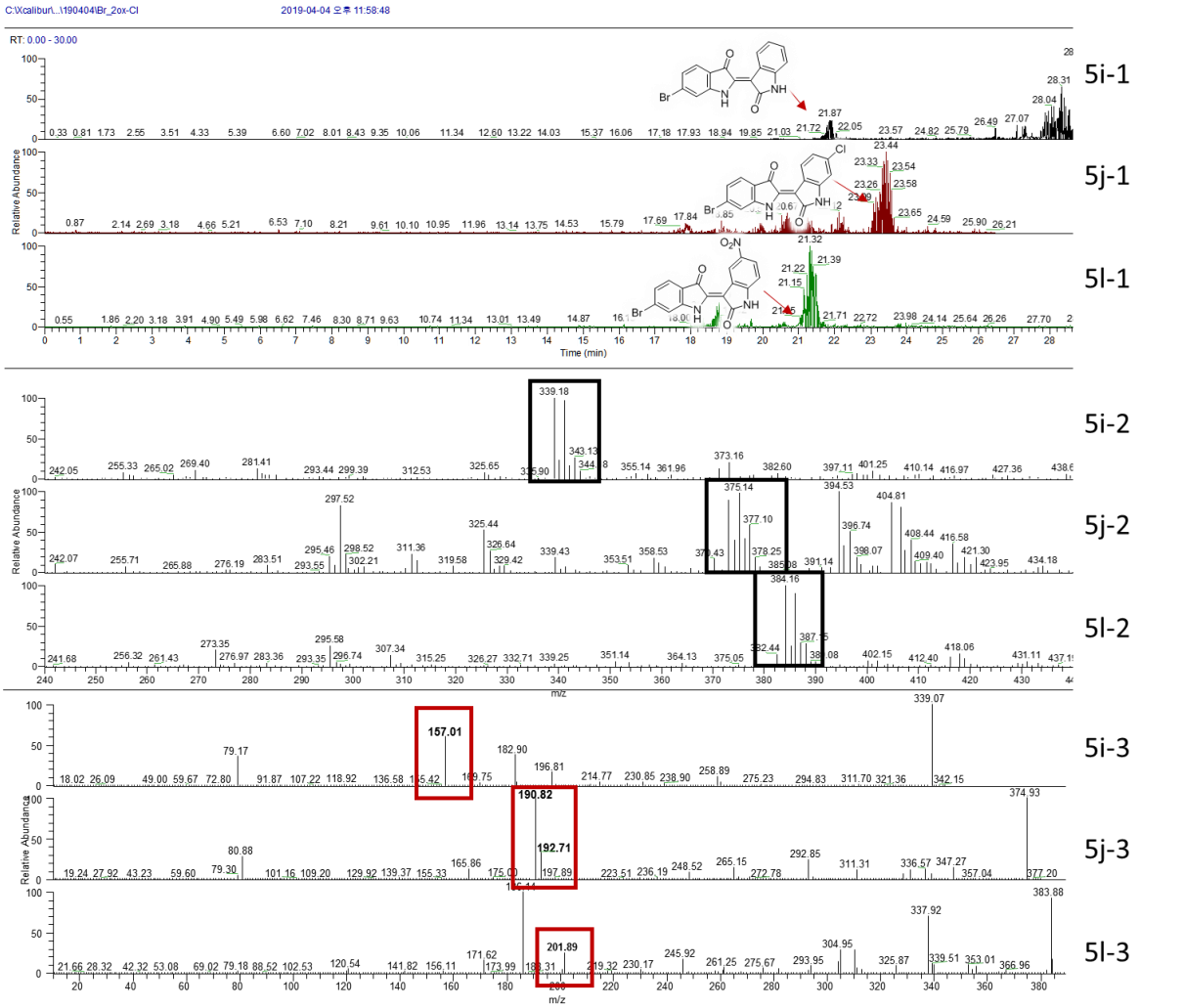

Figure S30.
LC (1), MS (2) and MS/MS analysis (3) of indirubin derivatives synthesized from 6-bromoindole. 6'-bromoindirubin (5i), 6'-bromo-6-chloro-indirubin (5j), and 6'-bromo-5-nitroindirubin (5l).

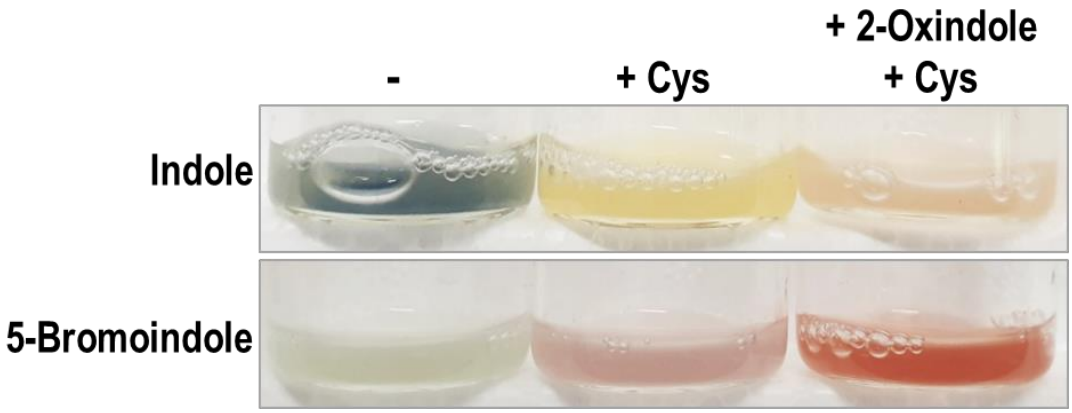

Figure S31.
The effect of supplementation of L-cysteine and 2-oxindole (**2a**) on the CYP102G4 reaction system.

When indole (**1a**) was used as a substrate, indigo (**5**) was produced. When L-cysteine was supplemented, production of **5** was inhibited but synthesis of indirubin (**4a**) was not observed. Only when **2a** was additionally supplemented, the synthesis of **4a** was observed through LC-MS and MS/MS. We previously showed that CYP102G4 produces only **5**, but not **4a**. The statement indicates that CYP102G4 is capable of indoxyl (**3a**) production, but difficult to synthesize ring A precursors for 4a. When L-cysteine was added, **3a** generated by CYP102G4 was converted into **6a** (Figure S9), but did not progress to **4a** synthesis due to the absence of ring A precursors (Figure S32). However, 4a was only synthesized when **2a** was supplemented as a ring A precursor. Therefore, the result is consistent with the previous and the present studies.

When 5-bromoindole was applied as a substrate, 5,5'-dibromoindigo (blue dye) was produced. When L-cysteine was supplemented, production of 5,5'-dibromoindirubin was observed through LC-MS and MS/MS (Figure S33). This indicates that CYP102G4 possess different regiospecificity for the oxidation of 5-bromoindole compared to that of indole. When **2a** was further supplemented, synthesis of a mixture of 5,5'-dibromoindirubin and 5'-bromoindirubin was observed and confirmed by LC-MS and MS/MS.

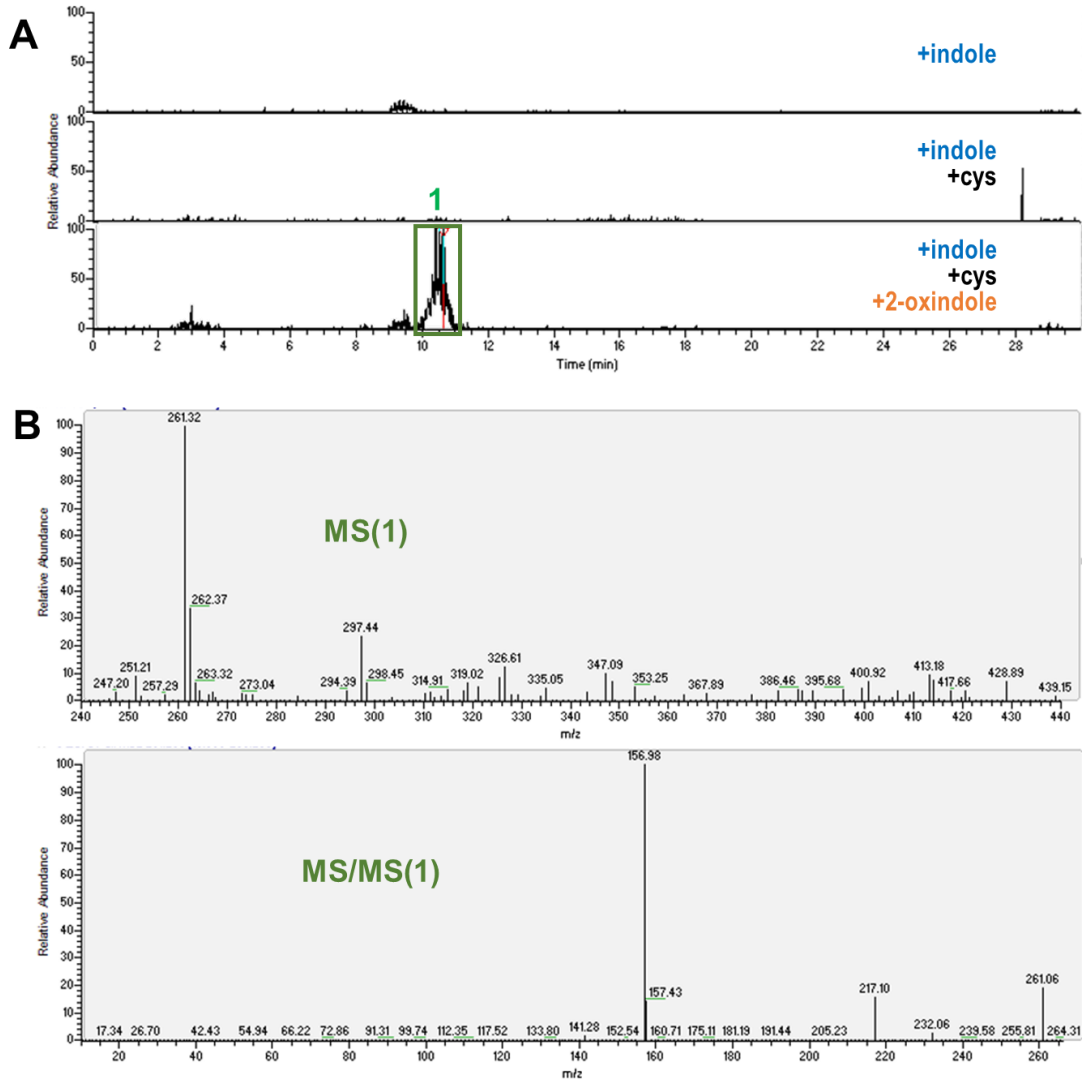

Figure S32.
(A) LC-MS chromatogram obtained from the various CYP102G4 reaction systems with indole as a substrate. (B) MS/MS analysis of the chromatogram peak at box 1 in Figure S32(A), which is revealed as indirubin. Synthesis of indirubin was observed when both L-cysteine and 2-oxindole were supplemented.

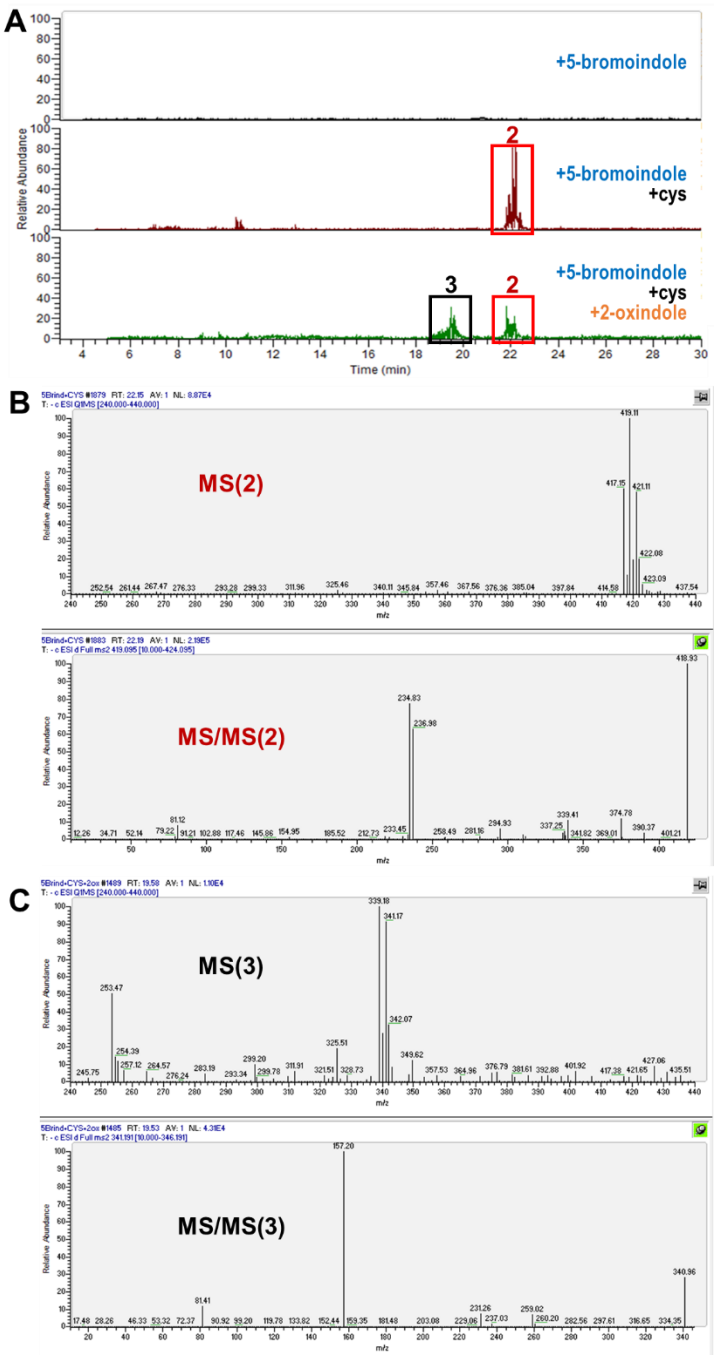

Figure S33.
(A) LC-MS chromatogram obtained from the various CYP102G4 reaction systems with 5-bromoindole as a substrate. (B) MS/MS analysis of the chromatogram peak at box 2 in Figure S33(A), which is revealed as 5,5'-dibromoindirubin. (C) MS/MS analysis of the chromatogram peak at box 3 in Figure Figure S33(A), which is revealed as 5'-bromoindirubin. When cysteine was added, 5,5'-dibromoindirubin was synthesized. When **2a** was additionally supplemented, a mixture of 5,5'-dibromoindirubin and 5'-bromoindirubin was produced.

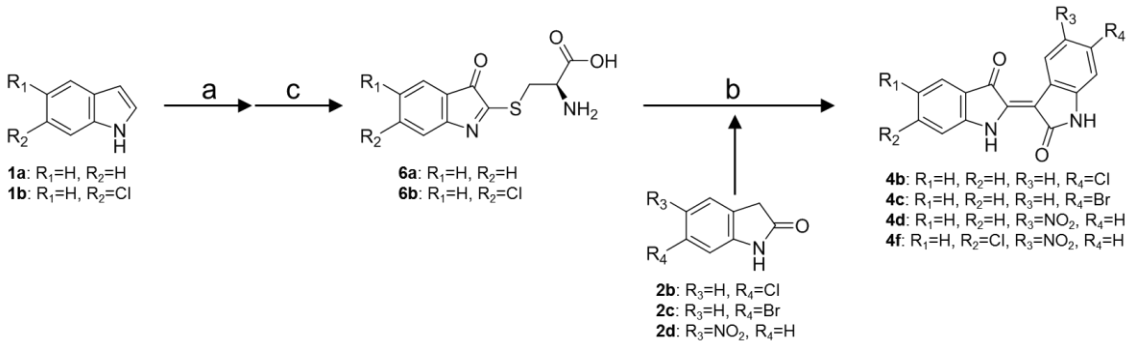

Figure S34.
Selective synthesis of indirubin. a: MaFMO, 200 rpm, 30 ºC, 100 mM Tris-HCl pH 7. b: pH adjustment to 9 and addition of 2. c: extraction using chloroform and ethyl acetate

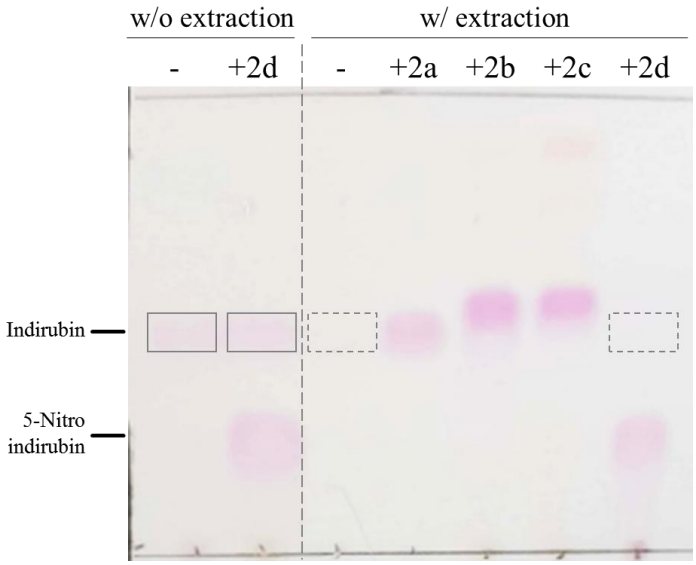

Figure S35.
TLC result showing the effect of adding extraction step in selective indirubin production. Boxes with solid lines and dotted lines indicate the existence and the absence of indirubin, respectively.

Table S1.
Result of selective synthesis of indirubin including 4b, 4c, 4d, and 4h.

Substrate		Product					Conversion (%)
			R_1	R_2	R_3	R_4	
1a	2b	4b	H	H	H	Cl	22
	2c	4c	H	H	H	Br	34
	2d	4d	H	H	NO_2	H	68
1b	2d	4h	H	Cl	NO_2	H	17

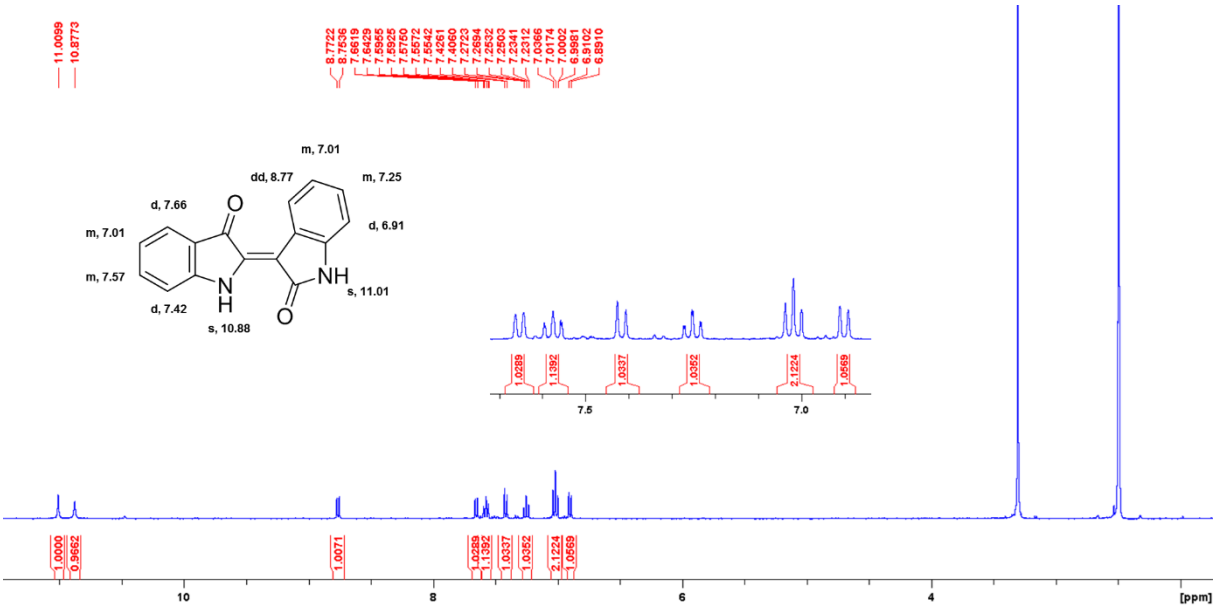

Figure S36.
H-NMR of indirubin

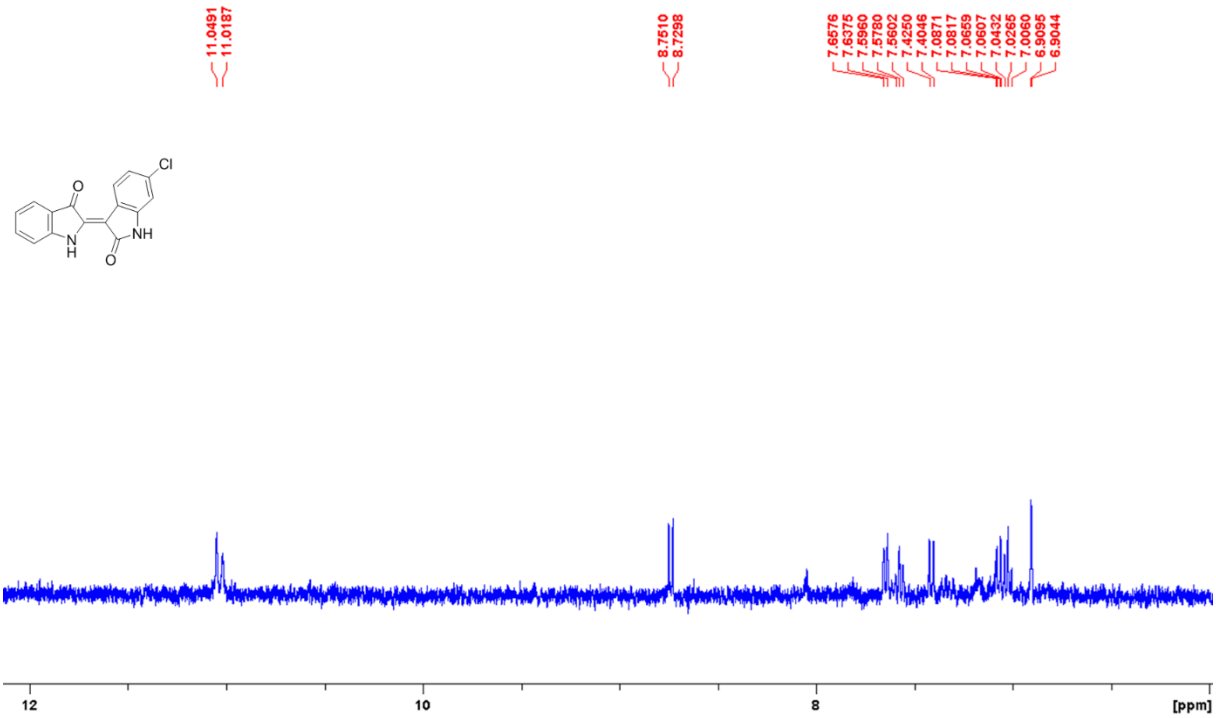

Figure S37.
H-NMR of 6-chloroindirubin

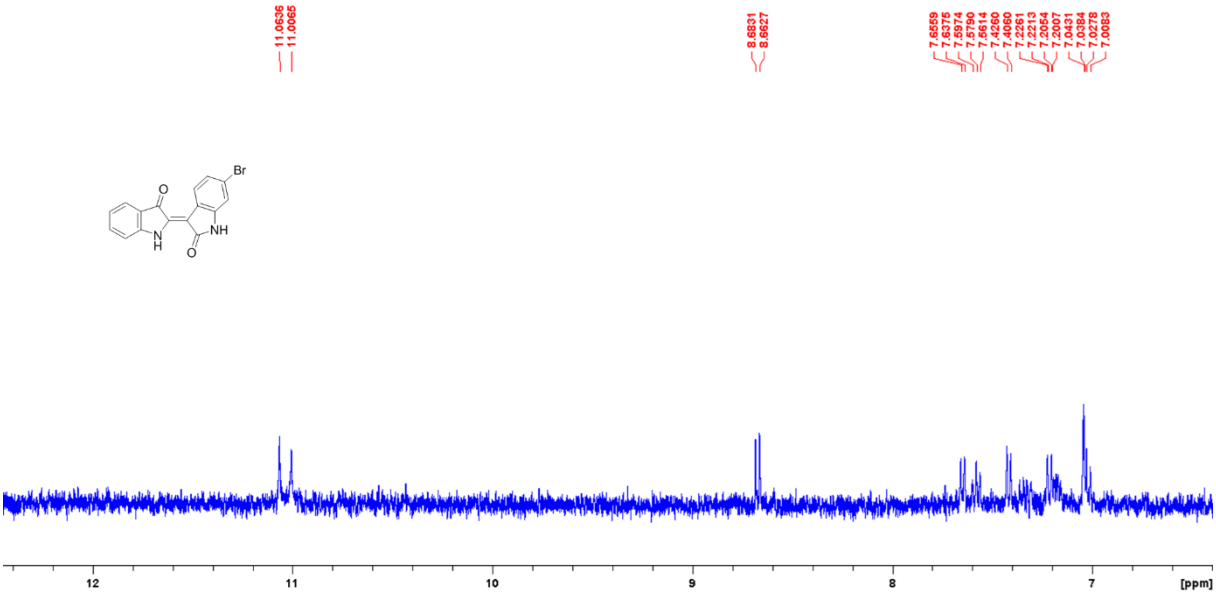

Figure S38.
H-NMR of 6-bromoindirubin

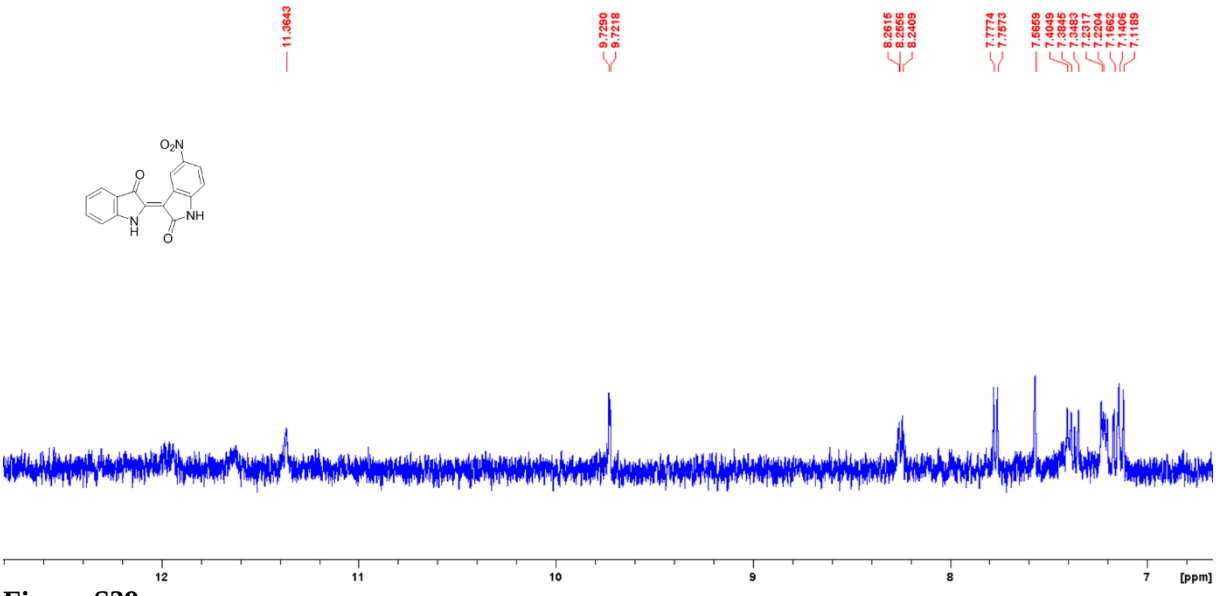

Figure S39.
H-NMR of 5-nitroindirubin

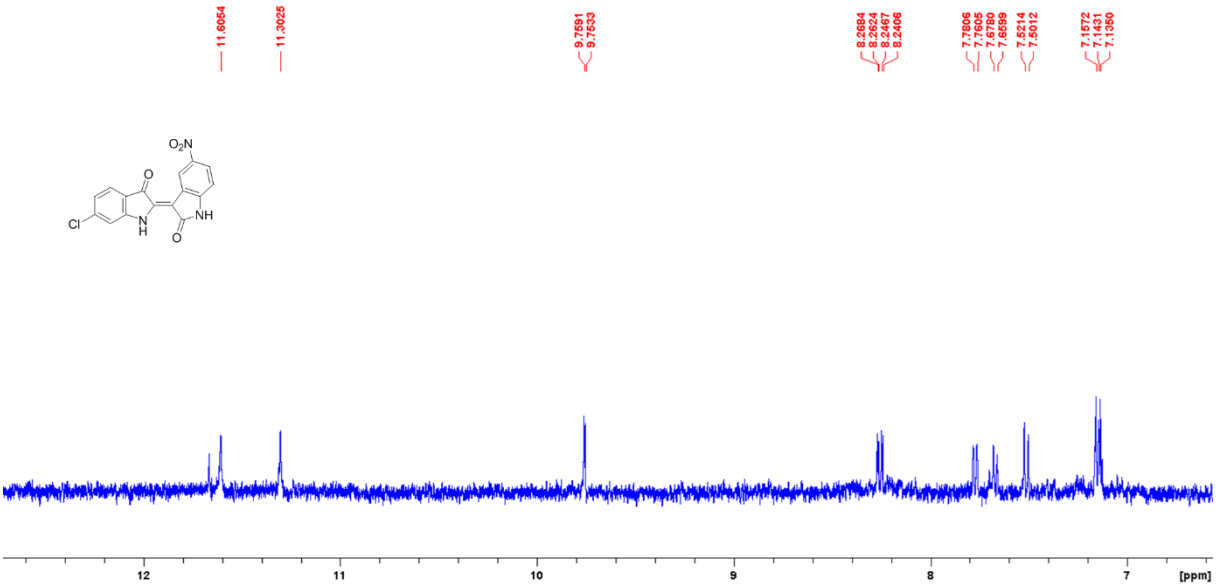

Figure S40.
H-NMR of 6'-chloro-5-nitroindirubin

References

1. Han, G. H.; Gim, G. H.; Kim, W.; Seo, S. I.; Kim, S. W., Enhanced Indirubin Production in Recombinant *Escherichia coli* Harboring a Flavin-Containing Monooxygenase Gene by Cysteine Supplementation, *J. Biotechnol.* **2012,** 164, 179-187.
2. Jensen, C. N.; Ali, S. T.; Allen, M. J.; Grogan, G., Mutations of an NAD(P)H-Dependent Flavoprotein Monooxygenase that Influence Cofactor Promiscuity and Enantioselectivity, *FEBS. Open Bio.* **2013,** 3, 473-478.
3. Park, A.-R.; Kim, H.-J.; Lee, J.-K.; Oh, D.-K., Hydrolysis and Transglycosylation Activity of a Thermostable Recombinant β-Glycosidase from *Sulfolobus acidocaldarius*, *Applied Biochem. Biotech.* **2010,** 160, 2236-2247.
4. Kim, J.; Lee, P. G.; Jung, E. O.; Kim, B. G., *In Vitro* Characterization of CYP102G4 from *Streptomyces cattleya*: A Self-Sufficient P450 Naturally Producing Indigo, *Biochim. Biophys. Acta Proteins Proteom.* **2018,** 1866, 60-67.