

<https://doi.org/10.15388/vu.thesis.58>
<https://orcid.org/0000-0001-9072-6523>

VILNIUS UNIVERSITY

Mikas
SADAUSKAS

Indigo-producing enzymes: selection and application

DOCTORAL DISSERTATION

Natural sciences,
Biochemistry **N 004**

VILNIUS 2020

This dissertation was written between 2015 and 2019 at the Institute of Biochemistry, Life Sciences Center of Vilnius University. The research was supported by EU Horizon 2020 project INMARE: Industrial Applications of Marine Enzymes (BG-2014-2634486) and Research Council of Lithuania (doctoral studies were financed from the EU structural funds, and a scholarship was granted for academic accomplishments).

Academic supervisor – Prof. Dr. Rolandas Meškys (Natural sciences, Biochemistry – N 004).

This doctoral dissertation will be defended in a public meeting of the Dissertation Defence Panel:

Chairman – Prof. Dr. Daumantas Matulis (Vilnius University, Natural sciences, Biochemistry – N004).

Members:

Prof. Dr. Jolanta Sereikaitė (Vilnius Gediminas Technical University, Technological sciences, Chemical engineering – T005);

Prof. Dr. Saulius Serva (Vilnius University, Natural sciences, Biochemistry – N004);

Dr. Tomas Šinkūnas (Vilnius University, Natural sciences, Biochemistry – N004);

Prof. Dr. Dirk Tischler (Ruhr University Bochum, Natural Sciences, Biochemistry – N004).

The dissertation shall be defended at a public meeting of the Dissertation Defence Panel at 1 PM on 11th of September, 2019 in Room R102 of the Life Sciences Center, Vilnius University.

Address: Saulėtekio av. 7, Vilnius, LT-10257, Lithuania

Tel. +370 5 223 4420; e-mail: info@gmc.vu.lt

The text of this dissertation can be accessed at the Vilnius University library, as well as on the website of Vilnius University:

www.vu.lt/lt/naujienos/ivykiu-kalendorius

VILNIAUS UNIVERSITETAS

Mikas
SADAUSKAS

Indigoidinius junginius gaminančių fermentų paieška ir tyrimas

DAKTARO DISERTACIJA

Gamtos mokslai,
Biochemija **N 004**

VILNIUS 2020

Disertacija rengta 2015– 2019 metais Vilniaus universiteto Gyvybės mokslų centro Biochemijos instituto Molekulinės mikrobiologijos ir biotechnologijos skyriuje. Mokslinius tyrimus rėmė Europos sąjungos programa Horizontas 2020 (projektas INMARE (BG-2014-2634486)) ir Lietuvos mokslo taryba (doktorantūros studijos buvo finansuojamos Europos socialinio fondo lėšomis bei gauta stipendija už akademinis pasiekimus).

Mokslinis vadovas – prof. dr. Rolandas Meškys (Vilniaus universitetas, gamtos mokslai, biochemija – N 004).

Gynimo taryba:

Pirmininkas – **prof. dr. Daumantas Matulis** (Vilniaus universitetas, gamtos mokslai, biochemija – N004).

Nariai:

prof. dr. Jolanta Sereikaitė (Vilniaus Gedimino technikos universitetas, technologijos mokslai, chemijos inžinerija – T005);

prof. dr. Saulius Serva (Vilniaus universitetas, gamtos mokslai, biochemija – N004);

dr. Tomas Šinkūnas (Vilniaus universitetas, gamtos mokslai, biochemija – N004);

prof. dr. Dirk Tischler (Bochumo Rūro universitetas, gamtos mokslai, biochemija – N004).

Disertacija ginama viešame Gynimo tarybos posėdyje 2020 m. rugsėjo mėn. 11 d. 13 val. Vilniaus universiteto Gyvybės mokslų centro R102 auditorijoje. Adresas: Saulėtekio al. 7, Vilnius, Lietuva, LT-10257; Tel. +370 5 223 4420; el. paštas: info@gmc.vu.lt

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu: <https://www.vu.lt/naujienos/ivykiu-kalendorius>

TABLE OF CONTENT

LIST OF ABBREVIATIONS	6
LIST OF PUBLICATIONS	7
INTRODUCTION.....	8
RESULTS AND DISCUSSION	11
1. Application of indole-3-carboxylic acid monooxygenase.....	12
2. Hind8 oxygenase and synthesis of novel indigoids.....	14
3. Indole-oxidizing enzymes with important biological functions	16
3.1 Mechanism of indole biodegradation	16
3.2 Mechanism of indole-3-acetic acid biodegradation.....	18
4. Concluding remarks and future research	21
CONCLUSIONS.....	22
SUMMARY/SANTRAUKA	23
CONFERENCE PRESENTATIONS.....	27
CURRICULUM VITAE	28
REFERENCES.....	29
ACKNOWLEDGMENTS.....	33
COPIES OF PUBLICATIONS	35

LIST OF ABBREVIATIONS

DOAA	3-hydroxy-2-oxindole-3-acetic acid
IAA	indole-3-acetic acid
<i>iac</i>	indole-3-acetic acid gene cluster (<i>i</i> ndole-3- <i>a</i> cetic acid <i>c</i> atabolism)
IBA	indole-3-butyric acid
<i>iif</i>	indole biodegradation gene cluster (<i>i</i> ndole- <i>i</i> nducible <i>f</i> lavoprotein)
IPA	indole-3-propionic acid
NMR	nuclear magnetic resonance
Ox-IAA	2-oxindole-3-acetic acid

LIST OF PUBLICATIONS

Publication I. Casaite, V., Sadauskas, M., Vaitekunas, J., Gasparaviciute, R., Meskiene, R., Skikaite, I., Sakalauskas, M., Jakubovska, J., Tauraite, D., Meskys, R. Engineering of a chromogenic enzyme screening system based on an auxiliary indole-3-carboxylic acid monooxygenase. *Microbiology Open*, 2019, 8, e795.

Publication II. Sadauskas, M., Statkevičiūtė, R., Vaitekūnas, J., Petkevičius, V., Časaitė, V., Gasparavičiūtė, R., Meškys, R. Enzymatic synthesis of novel water-soluble indigoid compounds. *Dyes and Pigments*, 2020, 173, 107882.

Publication III. Sadauskas, M., Vaitekūnas, J., Gasparavičiūtė, R., Meškys, R. Indole Biodegradation in *Acinetobacter* sp. Strain O153: Genetic and Biochemical Characterization. *Applied and Environmental Microbiology*, 2017, 83(19):e01453-17.

Publication IV. Sadauskas, M., Statkevičiūtė, R., Vaitekūnas, J., Meškys, R. Bioconversion of biologically active indole derivatives with indole-3-acetic acid-degrading enzymes from *Caballeronia glathei* DSM50014. *Biomolecules*, 2020, 10(4):663.

INTRODUCTION

Indigo is a blue pigment that has been used for cotton dyeing for more than 6000 years [1]. It is composed of two indolinone aromatic ring systems, connected with a 2-2' double bond. Such cross-conjugated system, intramolecular formation of hydrogen bonds and the nature of the heteroatom are all believed to contribute to the deep blue color of indigo[2]. Substituent chemical groups can be introduced into the structure of indigo and the resulting compounds are regarded as indigoids, which often exhibit exclusive characteristics[3,4]. Although chemical synthesis can be used to produce indigoids even in asymmetric form[5], it employs environment-damaging conditions, such as cyanide, formaldehyde, aniline, strong bases and heating to 300°C[6]. Fairly recently, enzyme-catalyzed synthesis of indigoids emerged as an environmental-friendly alternative method offering additional benefits.

The first evidence for bacterial synthesis of indigo using the recombinant enzymes was described in 1983 by application of naphthalene dioxygenase[7]. Following this, different enzyme groups including cytochrome P450 monooxygenases, flavin-dependent monooxygenases and dioxygenases were reported to be capable of oxidizing indole to indigo. Some of these enzymes exhibit a substrate promiscuity and can oxidize indole derivatives to corresponding indigoids, becoming potential biocatalysts for industrial application. The challenging part for the industrial application of enzymatic indigo synthesis is the requirement of cofactors and cofactor regeneration systems. This can be circumvented by employing a whole-cell bioconversion strategy, since certain well-characterized bacterial hosts provide sufficient amount of the required cofactors[8].

Indole and its derivatives possess important biochemical functions. Indole itself has been recently recognized as a signaling molecule that regulates the interactions between organisms of different species[9]. Of special interest is the connection between indole production in the gut and the emotional and nutritional behavior of rats, possibly mediated through gut-microbiome-brain axis[10,11]. In addition, 3-substituted carboxyl derivatives of indole also carry different biological roles. While indole-3-acetic acid, indole-3-propionic acid and indole-3-butyric acid all perform as auxins and regulate the plant growth, indole-3-propionic acid has been demonstrated to possess an exceptional capability to inhibit the proliferation of *Mycobacterium tuberculosis*[12] through suppression of tryptophan biosynthesis[13]. These characteristics of

indole and its derivatives emphasize the importance of indole-converting enzymes, which can be employed to create enzymatic conversion systems for studying the roles of indole derivatives *in vivo*.

The demand for biocatalysts with novel properties still continues to rise. Although natural sources are believed to offer an unlimited genetic and biochemical variety of enzymes [14], the lack of efficient screening and selection systems limits the exploitation and application of biological catalysis. To demonstrate the versatility of indigo production-based enzyme screening systems, **the aim of this work** was to explore the potential of indigo-producing enzymes. In order to achieve this aim, the following tasks were formulated:

1. To construct a system based on indigo production for the selection of enzymes with different activities;
2. To construct a biosynthesis platform for the production of novel indigoid compounds;
3. To provide novel experimental data about the biological functions of different groups of indigo-producing enzymes.

Scientific novelty of the study

A novel enzyme, indole-3-carboxylic acid monooxygenase (Icm), which was able to convert indole-3-carboxylic acid into indigo, was identified from metagenome and characterized. Due to this activity, Icm enzyme was used as an auxiliary enzyme for the construction of enzyme-based screening system, where an enzyme of interest converts a substrate to indole-3-carboxylic acid, which is then oxidized to indigo by Icm, visually indicating the presence of the desired enzyme. The potential of this system was showcased by using indole-3-carboxaldehyde as a prochromogenic substrate for screening of the metagenomic libraries for aldehyde dehydrogenases. In addition to *Escherichia coli*, this system was also found to be applicable in a gram-positive host – *Rhodococcus erythropolis*. Also, Icm tolerated certain modifications in the indole ring of indole-3-carboxylic acid that could allow for the selection of enzymes with different specificity. Given these points, Icm-based selection system represents a feasible method for obtaining metagenome-encoded enzymes with desired activities.

Hind8 oxygenase, active towards indole, was selected as a promising tool for the conversion of a wide range of indole derivatives. By using *E. coli* cells with the recombinant Hind8 oxygenase as a bioconversion platform, a series

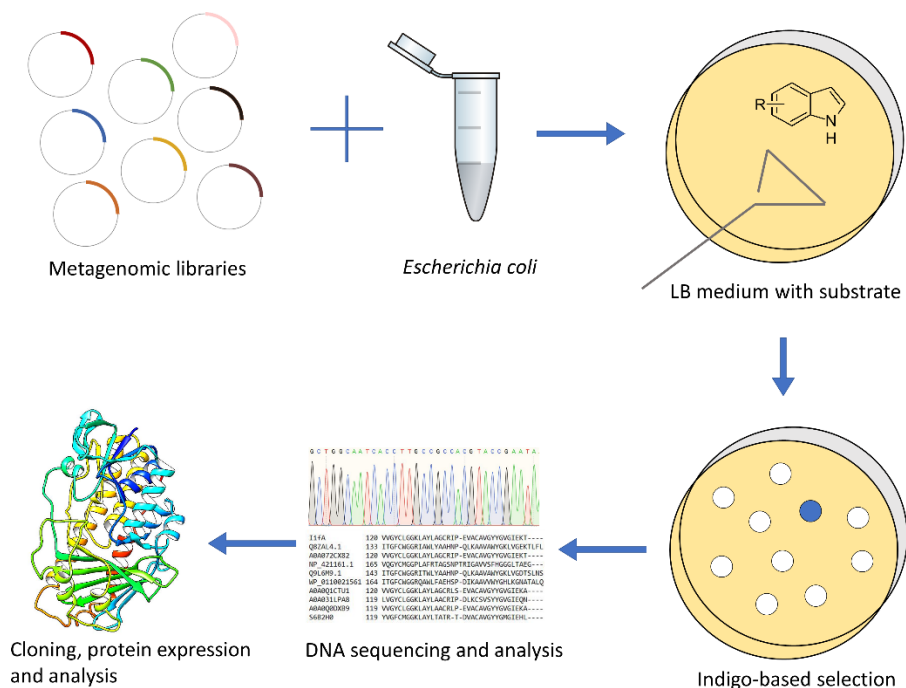
of novel indigoid compounds were produced for the first time. Among those, indigo and indirubin dicarboxylic acids showed increased solubility in water (from micromolar to millimolar range). Hence, the oxygenase Hind8 represents a novel tool with wide substrate specificity for the biosynthesis of indigoid compounds. On top of that, a novel approach can be suggested for increasing the water solubility of indigoids by the introduction of carboxyl groups.

Novel biological functions of indigo-producing and related enzymes are shown in this work. Although indole detoxification function has been suggested for the enzyme IifC, no functions for other Iif proteins have been assigned. By identifying that Iif proteins catalyze the initial steps of indole biodegradation, this study elucidates the long-searched genes and proteins responsible for indole catabolism, suggesting that indole can be detoxified by certain bacteria through Iif-mediated assimilation.

Following the elucidation of bacterial indole biodegradation, several important aspects of bacterial degradation of indole-3-acetic acid (IAA, a plant growth hormone) have been addressed. It has been known that Iac proteins are responsible for the initial steps of IAA biodegradation. This study clarifies that IacB protein does not participate in IAA biodegradation. Also, experimental evidence presented here suggest an H₂O-dependent formation of DOAA, an intermediate compound of IAA biodegradation. In addition, IacA and IacE enzymes were found to be capable of converting indole-3-propionic acid and indole-3-butyric acid into corresponding DOAA derivatives. As no biodegradation data has been available so far regarding these biologically active derivatives of indole, IacA and IacE enzymes could be a potential constituent of such hypothetical pathway.

RESULTS AND DISCUSSION

This thesis is based on four original publications listed as publications **I–IV**. Each of these publications deals with isolation and characterization of indigo-forming enzymes, following a general selection principle as shown in Scheme I. While sharing the ability to form indigo, all described enzymes belong to different groups of flavin-dependent oxygenases, and possess different activities and biological roles. Publication **I** describes the isolation and application of indole-3-carboxylic acid monooxygenase, which oxidizes indole-3-carboxylic acid (I3CA) to indigo. Based on this activity, an enzyme screening system was created that can be tuned for the search of the desired enzymatic activity. In publication **II**, the enzyme Hind8 is described that is able to synthesize a wide range of indigoid compounds, including water-soluble indigo dicarboxylic acids and indirubin dicarboxylic acid. Publications **III** and **IV** present novel pathways and mechanisms of microbial degradation of indole and IAA, respectively, where flavin-dependent oxygenases initiate the degradation cascades.



Scheme I. A general scheme for the selection of indigo-forming proteins.

1. Application of indole-3-carboxylic acid monooxygenase

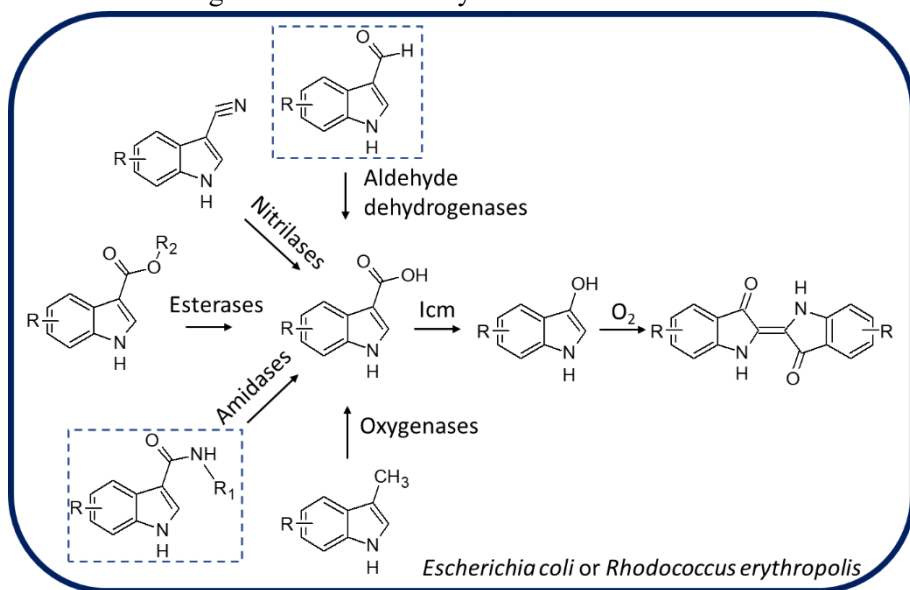
Aldehydes are very important intermediates in different metabolic pathways and biotechnological processes including the conversion of lignin into basic chemicals[15]. However, aldehydes are highly reactive compounds and are generally toxic to the living organisms. Different strategies can be used for the detoxification of these compounds. In living organisms, aldehyde dehydrogenases (ALDHs) are the main weapon responsible for the detoxification of aldehydes by oxidation to corresponding carboxylic acids. Also, ALDHs are important catalysts in numerous synthetic processes. Here, a new selection system for ALDHs was created by using indigo formation as a screening platform.

To create such system, an enzyme converting indole-3-carboxylic acid (I3CA) to indigo was isolated by applying a general metagenomic screening approach as depicted in Scheme I. This enzyme was annotated as a hypothetical flavin-dependent oxidoreductase, and monooxygenase-characteristic sequence motifs allowed the characterization of this protein as group A flavin-dependent monooxygenase. Only several dioxygenases have been shown to convert I3CA to indigo[16], but these enzymes were unrelated to the isolated monooxygenase. Based on sequence comparison and unique substrate scope (Publication I), this isolated enzyme was named as indole-3-carboxylate monooxygenase (Icm). The enzyme performs an oxidative decarboxylation of I3CA, resulting in indoxyl, which then dimerizes into indigo.

Although significant efforts were made to obtain a purified recombinant Icm, the activity of this protein disappeared as soon as the heterologous cells carrying the recombinant protein were disrupted. This could be explained by possible sensitivity of this protein to oxygen, requirement for binding to stabilizing molecules or other factors. However, *E. coli* cells with recombinant Icm protein were fully active and capable of converting I3CA into indigo, providing the basis for the construction of a double screening platform. The hypothesis was to use a substrate, which could be converted to I3CA, which, in turn, could be oxidized to indigo and the presence of I3CA-producing enzyme would become visible. To achieve this, *icm* gene was cloned to pACYC184 plasmid for constitutive expression in *E. coli*. Cells carrying pACYC184-*icm* plasmid were cotransformed with a compatible plasmid pUC19, which was used as a platform for creation of metagenomic libraries. By plating approximately 30000 transformants on indole-3-carboxaldehyde,

52 indigo-forming clones were selected. After removing the false-positives (mostly encoding Baeyer-Villiger monooxygenases, which were able to oxidize indole-3-carboxaldehyde to indoxyl with a concomitant formation of indigo) and recurrent hits, 20 different clones were selected, encoding hypothetical aldehyde dehydrogenases. The selected clones were demonstrated to possess different substrate scope towards aldehyde compounds (Publication I, Table 3), suggesting that the double screening system based on Icm and indigo formation was suitable for the selection of diverse groups of aldehyde dehydrogenases.

To expand the capabilities of this system, it was transferred to a different host, a gram-positive bacterium *Rhodococcus erythropolis* SQ1. It turned out that recombinant Icm was active in this host and was able to convert I3CA to indigo. Coexpression of Icm and a selected aldehyde dehydrogenase resulted in the oxidation of indole-3-carboxaldehyde into indigo, showing that the screening platform was active and could provide an additional tool for the selection of metagenome-encoded enzymes.



Scheme II. The principal scheme of the functional screening of enzymes based on an auxiliary Icm enzyme. The experimentally tested substrates are boxed. R = -H; 5-CH₃; 5-Br; 6-CH(CH₃)₂; [g]-benzene; [g]-tetrahydrocyclopentane, R₁ = any radical, R₂ = -CH₂(OH).

The versatility of this system lies in the possibility to select enzymes with desired enzymatic activity, which would be predetermined by the substrate being used (Scheme II). The only requirement is the ability of the enzyme to produce I3CA, and this could include esterases, amidases, nitrilases, oxidoreductases and other enzymes. Furthermore, Icm itself demonstrated substrate promiscuity, accepting substituted I3CA substrates (5-methyl, 5-bromo, benzo[g]indole and 1,6,7,8-tetrahydrocyclopenta[g]indole). The use of these substrates could allow for the selection of enzymes with different specific activity.

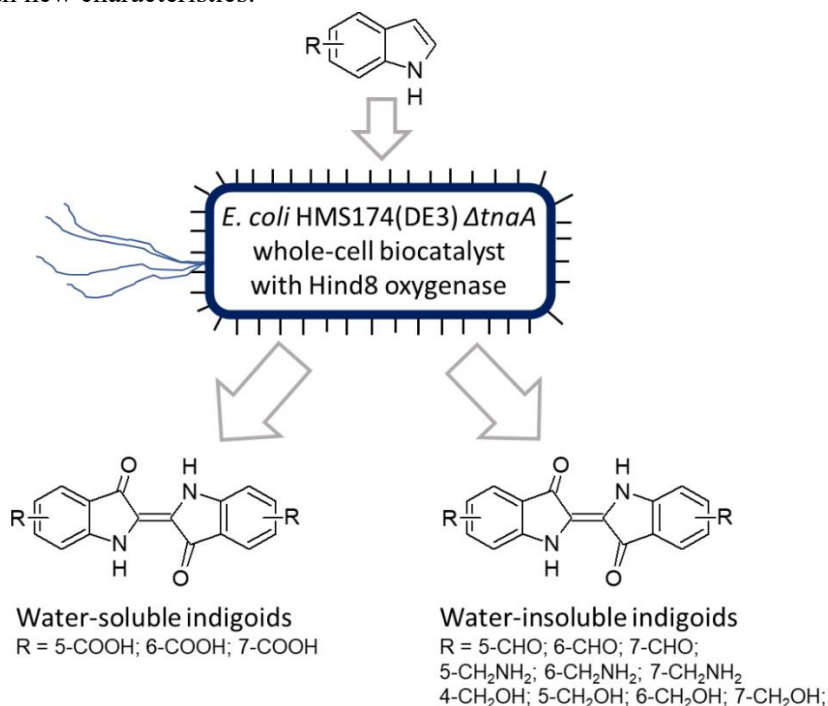
2. Hind8 oxygenase and synthesis of novel indigoids

To date, the enzymatic synthesis of indigoid compounds has been limited to synthesis of halogenated, methoxy-, cyano-, nitro-, and similar derivatives of indigo[17–19]. In order to expand the variety of indigoids by introducing new chemical groups and, presumably, novel characteristics, a metagenomic screening was performed using the formation of blue colonies on aminomethyl-, hydroxymethyl-, carboxaldehyde or carboxylic derivatives of indole as a screening platform. By using this approach, a metagenomic clone, which was capable of oxidizing a variety of indole compounds into corresponding indigoid pigments, was identified (Publication II, Fig. 1). A group B flavin-dependent monooxygenase, named Hind8, was found to be responsible for this activity.

Next, a biotransformation platform was created for the synthesis of indigoids carrying the abovementioned chemical groups by expressing Hind8 monooxygenase in the tryptophanase-deficient ($\Delta tnaA$) *E. coli* BL21(DE3) cells. This platform was successfully used for the biosynthesis of the following indigoids: 4,4'-di(hydroxymethyl)indigoid, 5,5'-di(hydroxymethyl)indigoid, 6,6'-di(hydroxymethyl)indigoid, 7,7'-di(hydroxymethyl)indigoid, 5,5'-di(aminomethyl)-indigoid, 6,6'-di(aminomethyl)-indigoid, 7,7'-di(aminomethyl)-indigoid, indigoid-4,4'-dicarboxaldehyde, indigoid-5,5'-di-carboxaldehyde, indigoid-6,6'-dicarboxaldehyde and indigoid-7,7'-di-carboxaldehyde (Scheme III). Importantly, the techniques used for the characterization of these pigments did not allow the separation between indigo, isoindigo and indirubin isoforms, therefore these compounds were regarded as indigoid compounds.

Indigo and indirubin are highly insoluble in water (3.7 μM and 4.8 μM , respectively). Several indigo derivatives have been created with increased

water solubility, the most notable being indigo carmine (indigo-5,5'-disulfonic acid, solubility in water ~20 mM). An alternative method for the synthesis of water-soluble indigoid compounds is presented herein. By using Hind8 monooxygenase and indole carboxylic acids (indole-5-carboxylic acid, indole-6-carboxylic acid and indole-7-carboxylic acid) as substrates, a range of indigo and indirubin dicarboxylic acids were produced, purified and analyzed. The structures of indigo-5,5'-dicarboxylic acid and indirubin-7,7'-dicarboxylic acid were confirmed by ^1H NMR, but the yield of indigoid-6,6'-dicarboxylic acid and indigoid-7,7'-dicarboxylic acid was insufficient for structure determination. However; based on absorbance spectrum and a clear blue color, these two indigoids were presumed to be indigo-6,6'-dicarboxylic acid and indigo-7,7'-dicarboxylic acid, respectively. It turned out that the solubility of all tested indigoid dicarboxylic acids in water (8–16 mM) was comparable to that of indigo carmine (publication II, Table 1). The increase in solubility by introduction of two carboxyl groups and a simple environmentally friendly method for the preparation of indigoids with such modification suggest a novel and feasible approach for synthesis of indigoids with new characteristics.



Scheme III. Biosynthesis of novel water-soluble and water insoluble-indigoids by using *E. coli* cells containing the recombinant Hind8 oxygenase.

Remarkably, an increase in water-solubility of indigoids often results in novel biological activities and applications. For example, due to reversible oxidation and reduction, indigo carmine has been used to create a redox flow cell at neutral pH[20]. Indirubin-3-oximes, which were highly water-soluble, also possess the ability to inhibit insulin-like growth factor receptor 1[21]. Water-soluble indigo and indirubin dicarboxylic acids described in this thesis were tested as electron acceptors during glucose dehydrogenase-mediated glucose oxidation. It turned out that among the four tested indigoids, only indigo-6,6'-dicarboxylic acid was reduced by glucose dehydrogenase, suggesting different properties even among these structurally similar indigoids. Overall, the introduction of carboxyl groups into indigoid structure could be a simple method for increasing the solubility in water and, in turn, the extent of biological activity.

3. Indole-oxidizing enzymes with important biological functions

Bioconversion of indole into indigo is not a process that often occurs in nature. In fact, compared to other pigments, indigo pigment does not participate in photosynthesis nor does it perform any other biological function. Therefore, the indigo-producing activities of enzymes described in publications I and II could be explained most likely by enzyme promiscuity as the physiological substrates of these enzymes are unknown. Recently, a physiological indole oxidation to indigo was reported in *Acinetobacter baumannii*[22]. This process was performed by a flavin-dependent monooxygenase with high efficiency and it would make sense for indole-exposed bacteria to detoxicate it by converting into non-toxic indigo. However, this was only a fraction of the mechanism how this reluctant bacterium is dealing with indole.

3.1 Mechanism of indole biodegradation

It was hypothesized that certain bacteria can not only oxidize indole, but fully assimilate it as carbon and energy source, therefore displaying no pigment formation as a dead-end reaction. To overcome this, 5-bromoindoline was used instead of indole as a substrate for the screening of bacteria that could oxidize it to a colored pigment. By using this approach, a single bacterial colony was isolated that was able to convert 5-bromoindoline to a deep purple

pigment yet displaying no color when growing on indole. This strain was identified as *Acinetobacter* sp. strain O153 and was indeed able to fully assimilate indole in the concentration ranging from 0.5 to 2 mM and use it as a sole source for nitrogen. Construction of genomic library of this strain allowed the identification of the gene cluster that was responsible for the oxidation of indole and 5-bromoindoline; *iif* operon. The main part of this cluster, was found to contain several genes, one of which shared a high sequence similarity to an indole inducible flavoprotein IifC, which has been shown to perform the detoxification of indole by converting it into indigo[22].

The roles of other Iif proteins were clarified by cloning each *iif* gene into compatible plasmids and co-expression of Iif proteins in *E. coli*. While IifC was a flavin-dependent indole monooxygenase, IifD was found to be a flavin reductase, which oxidized NAD(P)H and provided the reduced flavin cofactor for the IifC. However, IifD was not necessary for the oxidation of indole as *E. coli* cells carrying IifC successfully converted indole to indigo. This suggested that *E. coli* itself can complement the activity of flavin reduction, the activity, which has been well documented[23]. Next, when *E. coli* cells carrying IifC, IifD and IifB proteins were exposed to indole, no indigo formation was observed. Instead, the end product was identified as 3-hydroxyindolin-2-one. Since IifB was annotated as a short-chain dehydrogenase/reductase family protein, this protein is suggested to perform the oxidation of indoline-2,3-diol at C2 position, resulting in 3-hydroxyindolin-2-one and preventing the formation of indigo. Coexpression of four Iif proteins (IifC, IifD, IifB and IifA) resulted in the conversion of indole into 2-aminobenzenecarboxylic acid (anthranilic acid). Also, purified 3-hydroxyindolin-2-one was converted into anthranilic acid with the purified recombinant IifA protein. Altogether, these results demonstrated that IifB is responsible for the synthesis of 3-hydroxyindolin-2-one, which is then converted into anthranilic acid by IifA.

The only reaction in Iif-catalyzed indole degradation with an unclear mechanism is the conversion of 3-hydroxyindolin-2-one to anthranilic acid, performed by IifA. Functionally, this conversion with the purified recombinant IifA did not require any cofactors or metal ions, but was dependent on oxygen as it did not occur under anaerobic conditions. Structurally, IifA appears to be composed of two domains: the N-terminal dienelactone hydrolase (DLH)-like domain and the C-terminal SnoaL-like domain. The N-terminal domain possesses an α/β hydrolase fold with a canonical catalytic triad, composed of Cys124, Asp173, and His204. All these results demonstrate significant similarities between IifA and a cofactor-independent

dioxygenases, a distinct group of enzymes that catalyzes dioxygenolytic ring cleavage of 1*H*-3-hydroxy-4-oxoquinoline and 1*H*-3-hydroxy-4-oxoquinoline[24,25]. Recognition of these enzymes as cofactor-independent oxygenases and clarification of reactions mechanisms took more than a decade, suggesting a long road for IifA to be identified as member of this family. To achieve this, the role of both domains of IifA should be clarified, other reaction product containing carbon atom has to be identified and the mode of oxygen incorporation must be determined in order to accept this enzyme as a cofactor-independent oxygenase.

This work contributes significantly to the understanding of genes, proteins and reaction mechanisms that are involved in bacterial degradation of indole. Together with important inputs from other research groups[19,26], a comprehensive view of indole assimilation in certain bacteria can be offered. IifCD is a two-component flavin-dependent monooxygenase system that converts indole to an unstable indole-2,3-oxide[19], which then rapidly hydrolyzes to indoline-2,3-diol. This diol then breaks down to indoxyl, and, in the presence of oxygen and the absence of IifB, spontaneously dimerizes to indigo. When IifB is present, it oxidizes indoline-2,3-diol to 3-hydroxyindolin-2-one, a substrate for IifA, and the final product of Iif-mediated conversion is anthranilic acid. Further metabolism of anthranilic acid was showed to occur through CoA-thioester pathway, catalyzed by 2-aminobenzoyl-CoA gene cluster[26]. However, as *iif* and 2-aminobenzoyl-CoA gene clusters are rarely found together in bacterial genomes[19], alternatives mechanisms for catabolism of anthranilic acid are possible, including the synthesis of tryptophan performed by anthranilate phosphoribosyltransferase[27] or conventional salicylate, gentisate or catechol pathways. The latter is particularly likely as the operon encoding anthranilate dioxygenase, an enzyme, which converts anthranilic acid into catechol, is adjacent to the *iif* operon in some species related to *Acinetobacter* and *Pseudomonas* (Publication III, Fig. 3).

3.2 Mechanism of indole-3-acetic acid biodegradation

Having suggested novel proteins and mechanisms for indole biodegradation, the attention was focused on the biodegradation of indole derivative, IAA. It has been known that several bacterial species use Iac (indole-3-acetic acid catabolism) proteins for degradation of IAA. Surprisingly, the composition and structure of the *iac* operon resembles that of the

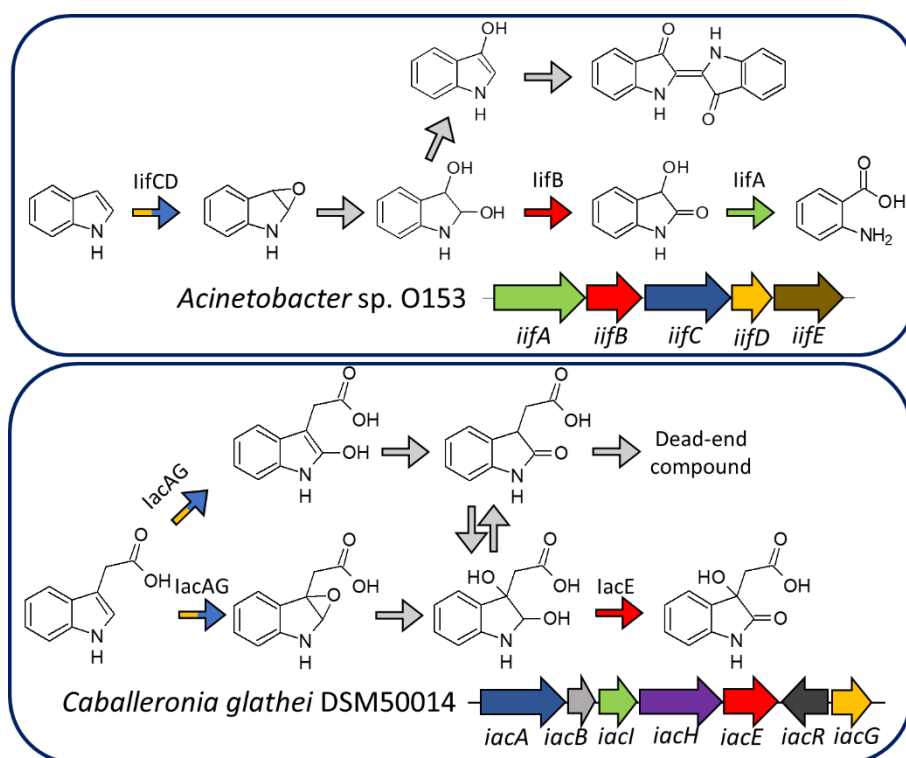
iif operon (Scheme IV). It has been demonstrated that a flavin-dependent oxygenase IacA, which can also oxidize indole to indigo, and a flavin reductase IacG perform the initial oxidation of IAA. Next, a short-chain dehydrogenase/reductase IacE produces DOAA, a stable intermediate of IAA biodegradation. IacB, a hypothetical protein with no sequence similarity to any other protein with known function has been speculated to be involved in the IacE-catalyzed reaction[28]. Also, the reaction product of IacA-catalyzed reaction has not been clarified and by analogy to indole oxidation, this product could be either a monohydroxylated, dihydroxylated or an epoxide form of IAA.

To clarify the roles of Iac proteins, *iacA*, *iacA* and *iacE* genes from *Caballeronia glathei* strain DSM50014 were cloned to compatible plasmids and co-expressed in *E. coli* in different combinations. It was found that *E. coli* cells with IacA protein converted IAA into Ox-IAA, and cells with IacAE combination synthesized DOAA from IAA. The addition of IacB did not result in any changes during these conversions, suggesting that IacB is not required in any of these reactions. Next, as the oxidation product of indole was highly unstable and formed a dead-end product (indigo), *E. coli* cells carrying IacE were tested for the ability to consume Ox-IAA, the oxidation product of IAA. Ox-IAA was fully consumed not only by *E. coli* cells with IacE protein, but also with wild-type IAA-induced *C. glathei* DSM50014 cells as well, suggesting that Ox-IAA is not a dead-end product.

To trace the origin of oxygen atoms incorporated into DOAA during IAA oxygenation, H₂¹⁸O was used. The molecular mass of Ox-IAA remained unchanged (191 Da) when *E. coli* cells with IacA were used for bioconversion of IAA in H₂¹⁸O environment. However, DOAA with increased molecular mass (209 Da compared to 207 Da) was clearly observed in two cases: i) when *E. coli* cells expressing IacAE were used for bioconversion of IAA in H₂¹⁸O environment and ii) with Ox-IAA and IacE-expressing *E. coli* cells in H₂¹⁸O environment, suggesting that the oxygen atom at C3 in DOAA is derived from water.

The results of these experiments allowed the clarification of the reaction mechanisms of IacA- and IacE-catalyzed reactions. For IacA-mediated oxidation of IAA, several alternative mechanisms can be proposed. In analogy with indole biodegradation and the mechanism of IifC, IacA could convert IAA into 2,3-epoxyIAA, which is then spontaneously hydrolyzed to 2,3-dihydroxyindoline-3-acetic acid and dehydrated to Ox-IAA. Alternatively, Ox-IAA could be formed by hydroxylation of IAA into 2-hydroxyIAA and

tautomerization into Ox-IAA. Depending on these two possible pathways, different mechanisms for the formation of DOAA could be proposed as well. If IacA forms and epoxide, IacE should catalyze the oxidation of 2,3-dihydroxyindoline-3-acetic acid to yield DOAA, since IacE is a dehydrogenase type enzyme. If, on the other hand, IacA hydroxylates IAA into 2-hydroxyIAA, another hydroxylation step at C3 atom has to occur to produce DOAA. This hydroxylation cannot be performed by IacE, and a more plausible mechanism would be a spontaneous addition of water to yield a diol derivative, which could then be oxidized to yield DOAA.



Scheme IV. Comparison of biodegradation of indole in *Acinetobacter* sp. strain O153 and biodegradation of IAA in *Caballeronia glathei* DSM50014. Grey reaction arrows indicate spontaneous reactions. Arrows with identical colors indicate genes and proteins with similar predicted functions.

In addition to bioconversion of IAA, *E. coli* bioconversion platform with IacAE proteins was found to be capable of converting homologs of IAA, including IPA and IBA, into corresponding derivatives of DOAA. These

compounds have important biological activities: both IPA and IBA can promote the plant growth[29], IPA shows important physiological activities in mammalian organisms[30–34] and even the antitubercular activity[12,13]. Since no biodegradation has been reported neither for IPA nor IBA, IacAE could be a central part of such hypothetical pathway. Another alternative could be the construction of artificial metabolic cascades by employing IacAE proteins. In such scenario, IacAE could convert 3-substituted carboxy derivatives of indole into the corresponding DOAA derivatives, and other enzymes would be required to further process DOAA compounds. For this, the IifA enzyme from the indole biodegradation pathway is a good starting point. Although it does not accept DOAA nor its derivatives as substrates, protein engineering might help to expand the substrate specificity of IifA and to demonstrate a ring-opening reaction of IAA and its homologs, which has not been reported to date.

4. Concluding remarks and future research

Indigo-based enzyme selection system presented in this thesis allows for a fast, convenient, cheap and easy-to-use method to obtain enzymes, which catalyze the important biological reactions and possess unique enzymatic properties. Although attempts were unsuccessful to obtain in a purified and active form some of the abovementioned heterologous indigo-producing enzymes, namely Hind8 and Icm, an *E. coli* host, which was used for the selection warrants the activity of these enzymes inside the cell, allowing the construction of whole-cell bioconversion systems. Several strategies may be employed to improve the current selection system. One way might be the use of thermophilic organisms as a host in order to select enzymes with higher thermostability. *In vitro* translation and indigo synthesis system might be also considered when aiming at enzymes active in a purified form.

CONCLUSIONS

1. An indole-3-carboxylic acid monooxygenase-based screening system is suitable for the selection of enzymes with desired activities.
2. Substituted indigoids with novel properties can be obtained by using Hind8-based bioconversion system.
3. If enzymes are responsible for the initial steps of indole biodegradation.
4. IacA and IacE enzymes possess the ability to convert biologically active indole derivatives.

SUMMARY/SANTRAUKA

Indigo pigmentas nuo neatmenamų laikų naudojamas tekstilinių audinių dažymui. Prieš beveik trisdešimt metų buvo parodyta, jog bakterijos su tam tikrais įterptais fermentais geba gaminti indigo pigmentą iš indolo, *N*-heterociklinio junginio. Nuo to laiko daugėjo žinių apie tokius fermentus – oksigenazes, gebančias oksiduoti indolą iki indoksilo, kuris spontaniškai dimerizuojasi į indigo. Nustatyta, jog šią reakciją gali katalizuoti tiek monooksigenazės, tiek dioksigenazės. Kai kurie iš šių fermentų pasižymi tokiu substratiniu savitumu, kuris leidžia katalizuoti biotechnologiškai svarbias reakcijas, o tai lemia jų pramoninį pritaikymą. Mokslininkai ėmėsi panaudoti indolą oksiduojančius fermentus modifikuotų indigo junginių gamybai, nes tokie junginiai pasižymi priešvėžiniu poveikiu bei yra ypač tinkami taikyti organinėje elektronikoje.

Daugėjo žinių ir apie indolą. Paaiškėjo, jog kai kurios bakterijos gamina jį ne kaip šalutinį produktą ir ne tik kaip tarpinį metabolizmo junginį, bet ir kaip signalinę molekulę. Vienoms bakterijoms indolas veikia kaip daugumos pojūčio signalas, tuo tarpu kitoms – kaip toksinis junginys, slopinantis gyvybinius procesus. Negana to, buvo nustatytas ir bakterijų gaminamų kai kurių indolo darinių biologinis aktyvumas, pvz., indolil-3-propioninės rūgšties slopina tuberkuliozę sukeliančių bakterijų augimą.

Iš aukščiau pateiktų duomenų matyti, jog indigo gaminantys fermentai turi daug perspektyvių panaudojimo sričių. Tiesa, nuo flavino priklausomos monooksigenazės liūdnai garsėja savo „kieta“ struktūra, kuri lengvai nepasiduoda, siekiant pagerinti fermento veikimą, pasitelkiant baltymų inžineriją. Todėl **šio darbo tikslas** buvo atrinkti ir ištirti skirtingus indigo pigmentą gaminančius baltymus, pasinaudojant indigo biosinteze paremta fermentų atrankos sistema. Šiam tikslui įgyvendinti buvo iškelti trys uždaviniai:

1. Sukurti indigo biosinteze paremtą fermentų atrankos sistemą, tinkančią tikslinei norimu aktyvumu pasižyminčių fermentų paieškai;
2. Sukurti naujų indigoidinių junginių biosintezės sistemą;
3. Ištirti indigo gaminančius fermentus, dalyvaujančius biologiniuose procesuose.

Šiame darbe buvo atliekama indigo pigmentą gaminančių fermentų paieška ir analizė. Fermentų paieškai buvo pasitelktos metagenominės DNR bibliotekos, sukonstruotos įterpiant metagenominės DNR fragmentus į pUC19 plazmidės. *E. coli* ląstelės buvo transformuojamos bibliotekomis ir

išsėjamos ant terpės su įvairiais substratais – indolo dariniais. Jei metagenominės DNR fragmente yra koduojamas baltymas, verčiantis substratą indigo pigmentu, ląstelių kolonija nusidažo mėlyna spalva ir tampa aiškiai atskiriama nuo bespalvių kolonijų (I schema). Tokiu būdu atrinkti fermentai buvo klonuojami į raiškos plazmides ir tiriami jų aktyvumai.

Substratu panaudojant indolil-3-karboksirūgštį, buvo atrinktas fermentas, verčiantis šį substratą indigo pigmentu. Paaiškėjo, jog tai yra A grupės nuo flavino priklausoma monooksigenazė, greičiausiai vykdomi indolil-3-karboksirūgšties oksidacinį dekarboksilinimą iki indoksilo, kuris dimerizuojasi iki indigo. Išgryninta indolil-3-karboksirūgšties monooksigenazė (*Icm*) nebuvo aktyvi, tačiau fermentas veikė būdamas *E. coli* ląstelėse, sukurdamas sąlygas dvigubos fermentų atrankos sistemos konstravimui. Konstruojant šią sistemą, *icm* genas buvo perkeltas į pACYC184 plazmidę, suderinamą su pUC19 plazmide, kuri buvo panaudota metagenominės DNR bibliotekų kūrimui. Išsėjant *E. coli* ląsteles su metagenomine biblioteka ir pACYC184-*icm* plazmide ant kietos terpės su indolil-3-karboksialdehidu, pavyko atrinkti 20 skirtingų fermentų aldehido dehidrogenazių, katalizuojančių indolil-3-karboksialdehido oksidaciją iki indolil-3-karboksirūgšties. Negana to, šie fermentai pasižymėjo gebėjimu oksiduoti įvairius aldehidus iki atitinkamų karboksirūgščių. Taip pat parodyta, jog *Icm* fermentu paremta dviguba fermentų atrankos sistema veikia ne tik gram-neigiamose *E. coli* ląstelėse, bet ir gram-teigiamose bakterijose *R. erythropolis*, taip praplečiant dvigubos fermentų atrankos sistemos galimybes. Galiausiai, pasirenkant atitinkamą substratą, kuris būtų verčiamas indolil-3-karboksirūgštimi, ši sistema gali būti panaudota skirtingų fermentų grupių tikslinei atrankai.

Siekiant praplėsti žinomų indigoidinių junginių įvairovę, buvo ieškoma fermentų, pasižyminčių plačiu substratiniu savitumu ir gebėjimu oksiduoti įvairius indolo junginius iki atitinkamų indigo junginių. Rastas vienas toks fermentas – Hind8. Panaudojant *E. coli* ląsteles su Hind8 fermentu kaip biokonversijos platformą, susintetinti šie indigoidiniai junginiai: 5-oje, 5'-oje, 6-oje, 6'-oje, 7-oje ir 7'-oje padėtyse turintys -CH₂OH, -COOH, -CH₂NH₂ arba -CHO grupes. Išskirtinėmis savybėmis pasižymėjo susintetinti indigoidiniai junginiai su karboksigrupėmis: indigo-5,5'-dikarboksirūgštis, indigo-6,6'-dikarboksirūgštis, indigo-7,7'-dikarboksirūgštis ir indirubino-7,7'-dikarboksirūgštis. Šių junginių tirpumas vandenyje buvo trimis eilėmis didesnis nei indigo arba indirubino tirpumas vandenyje. Kadangi indigoidinių junginių tirpumo vandenyje padidinimas dažnai suteikia šiems junginiams

naujas savybes, pvz., tam tikrų receptorių slopinimą, karboksigrupės įvedimas į indigoidinius junginius gali būti naujas būdas ženkliai padidinti indigoidinių junginių tirpumą vandenyje bei suteikti šiems junginiams naujas savybes.

Neseniai buvo parodyta, jog bakterijos, kurioms indolas yra toksiškas, geba apsisaugoti nuo toksinio poveikio oksiduodamos indolą iki mažiau toksiško indigo pigmento. Šiame darbe pavyko parodyti, jog *Acinetobacter* genties bakterijos geba ne tik nukenksminti indolą, bet ir suskaidyti jį bei panaudoti kaip azoto šaltinį augimui. Nustatyta, jog už šį procesą atsakingi Iif baltymai. Fermentas IifC, E grupės nuo flavino priklausoma oksigenazė, oksiduoja indolą, o oksidacijos produktas yra 2,3-dihidroksiindolinas. Šį junginį fermentas IifB, trumpų grandinių dehidrogenazė, oksiduoja iki 3-hidroksiindolin-2-ono. Pastarasis junginys fermento IifA yra verčiamas iki antranilo rūgšties – indolo skaidymo pagrindinio tarpinio junginio, toliau skaidomo kitų fermentų. Verta paminėti, jog fermentas IifA pagal savo savybes – deguonies naudojimą reakcijoje be jokių kofaktorių ar metalo jonų ir α/β hidrolazės tipo struktūrą – yra panašus į išskirtinę neseniai aprašytų fermentų grupę – nuo kofaktorių nepriklausomas dioksigenazes. Tikslus IifA fermento reakcijos mechanizmas nustatymas bei galimas priskyrimas nuo kofaktorių nepriklausomoms dioksigenazėms neabejotinai reikalauja papildomų tyrimų.

Išaiškinus bakterijų vykdomo indolo skaidymo pradinių reakcijų mechanizmus, dėmesys buvo atkreiptas į indolil-3-acto rūgšties skaidymą. Buvo žinoma, jog šiame procese dalyvauja panašūs baltymai, kaip ir nustatytame indolo skaidyme, tačiau tikslus skaidymo mechanizmas nebuvo aiškus. Tiriant indolil-3-acto rūgštį – augalų augimą skatinantį hormoną – skaidančias *Caballeronia* genties bakterijas, pavyko parodyti, jog fermentas IacA, B grupės nuo flavino priklausoma oksigenazė, katalizuoja indolil-3-acto rūgšties oksidaciją iki 2-oksindolil-3-acto rūgšties, o fermentas IacE, trumpų grandinių dehidrogenazė, verčia šį junginį 3-hidroksi-2-oksindolil-3-acto rūgštimi, ir šiame procese nedalyvauja fermentas IacB. Taip pat paaiškėjo, kad ties C3 atomu esančios hidroksigrupės deguonies atomas atkeliauja iš vandens molekulės. Siūlomi du indolil-3-acto rūgšties skaidymo mechanizmai: (1) IacA vykdomos oksidacijos metu susidaro epoksidai, kuris yra hidrolizuojamas iki 2,3-dihidroksiindolin-3-acto rūgšties ir fermento IacE toliau oksiduojamas iki 3-hidroksi-2-oksindolil-3-acto rūgšties, arba (2) IacA vykdomos reakcijos produktas yra 2-hidroksiindolil-3-acto rūgštis, prie kurios spontaniškai prisijungia vanduo ir susidaro 2,3-dihidroksiindolin-3-acto rūgštis, kurią oksiduoja IacE. Taip pat nustatyta, jog IacA ir IacE geba

konvertuoti ir kitus biologiškai aktyvius indolo darinius – indolil-3-propiono rūgštį ir indolil-3-butano rūgštį. Šių junginių katabolizmo keliai kol kas nėra aprašyti, todėl IacA ir IacE galėtų būti panaudoti dirbtinių skaidymo kelių kūrimui.

Apibendrinant galima teigti, jog indigo biosinteze paremta atrankos sistema leidžia greitai, nesudėtingai, sąlyginai pigiai ir tikslingai rasti fermentus, tiek atliekančius svarbias biologines funkcijas, tiek ir pasižyminčius įdomiomis konversijos savybėmis. Nors dalies šiame darbe aprašytų aktyvių rekombinantinių baltymų gauti nepavyko (Hind8, Icm), *E. coli* ląstelėse vykdoma tiesioginė atranka užtikrina, jog atrenkami tik ląstelėse aktyvūs fermentai, todėl šie fermentai gali būti panaudojami biokonversijos platformų kūrimui. Siekiant pagerinti naudotą sistemą, ją galima perkelti į kitus mikroorganizmus, pvz., termofilines bakterijas, tokiu būdu tikintis atrinkti termostabilius indigo gaminančius fermentus, arba taikyti transliacijos ir indigo sintezės *in vitro* atrankos sistemą, siekiant atrinkti *in vitro* stabilius indigo gaminančius fermentus.

Darbo išvados:

1. Icm paremta dviguba fermentų atrankos sistema gali būti sėkmingai naudojama norimos grupės fermentų paieškai ir atrankai.
2. Naudojant biokonversijos sistemą su Hind8 fermentu, galima pagaminti modifikuotus indigoidinius junginius, pasižyminčius netipiškomis savybėmis.
3. Iif fermentai yra atsakingi už pradines bakterijų vykdomo indolo skaidymo reakcijas.
4. IacA ir IacE fermentai geba konvertuoti biologiškai aktyvius indolo darinius.

CONFERENCE PRESENTATIONS

Oral presentations:

- **Sadauskas M**, Vaitekūnas J, Gasparavičiūtė R, Meškys R. Similarities and peculiarities between bacterial degradation of indole and indole-3-acetic acid. VAAM 2019, March 16-20, Mainz, Germany.

Poster presentations:

- **Sadauskas M**, Vaitekūnas J, Gasparavičiūtė R, Meškys R. Biochemical characterization of indole degradation in *Acinetobacter* sp. O153. 3rd Congress of Baltic Microbiologists, October 18-21, 2016, Vilnius, Lithuania.

- **Sadauskas M**, Časaitė V, Vaitekūnas J, Gasparavičiūtė R, Meškys R. Isolation, expression and stability improvement of indole-3-carboxylic acid oxygenase. 12th Conference on Protein Stabilization. May 16-18, 2018, Vilnius, Lithuania.

- Statkevičiūtė R., **Sadauskas M.**, Vaitekūnas J., Gasparavičiūtė R., Meškys R. Biocatalytic synthesis of indigo dimethanols and indigo dicarboxaldehydes. TheCoins 2019, January 26-28, Vilnius, Lithuania.

- Statkevičiūtė R., **Sadauskas M.**, Vaitekūnas J., Gasparavičiūtė R., Meškys R. Enzymatic synthesis of novel indigoid pigments. Open Readings 2019, March 19-22, Vilnius, Lithuania.

- **Sadauskas M**, Statkevičiūtė R, Vaitekūnas J, Petkevičius V, Meškienė R, Gasparavičiūtė R, Meškys R. Biocatalytic synthesis of indigoid pigments. EuropaCat2019, August 18-23, Aachen, Germany.

CURRICULUM VITAE

Name, Last name Mikas Sadauskas

Education/Competence:

2007 – 2009	Zarasai „Ažuolas“ Gymnasium
2009 – 2013	Vilnius University, Faculty of Natural Sciences, Bachelor in molecular biology
2013 – 2015	Vilnius University, Faculty of Natural Sciences, Master in microbiology and biotechnology
2015 – 2019	PhD studies. Vilnius University, Life Sciences Center, Institute of Biochemistry.

Record of employment:

2017 02 – 2017 12	Research assistant, Vilnius University,
2019 10 - present	Institute of Biochemistry

REFERENCES

1. Splitstoser, J.C.; Dillehay, T.D.; Wouters, J.; Claro, A. Early pre-Hispanic use of indigo blue in Peru. *Science Advances* **2016**, *2*, e1501623, doi:10.1126/sciadv.1501623.
2. Hunger, K. *Industrial Dyes: Chemistry, Properties, Applications*; John Wiley & Sons, **2007**; ISBN 978-3-527-60606-1.
3. Irimia-Vladu, M.; Głowacki, E.D.; Troshin, P.A.; Schwabegger, G.; Leonat, L.; Susarova, D.K.; Krystal, O.; Ullah, M.; Kanbur, Y.; Bodea, M.A.; et al. Indigo – From Jeans to Semiconductors: Indigo - A Natural Pigment for High Performance Ambipolar Organic Field Effect Transistors and Circuits (Adv. Mater. 3/2012). *Advanced Materials* **2012**, *24*, 321–321, doi:10.1002/adma.201290011.
4. Głowacki, E.D.; Voss, G.; Leonat, L.; Irimia-Vladu, M.; Bauer, S.; Sariciftci, N.S. Indigo and Tyrian Purple – From Ancient Natural Dyes to Modern Organic Semiconductors. *Israel Journal of Chemistry* **2012**, *52*, 540–551, doi:10.1002/ijch.201100130.
5. Klein, L.L.; Petukhova, V.; Wan, B.; Wang, Y.; Santasiero, B.D.; Lankin, D.C.; Pauli, G.F.; Franzblau, S.G. A novel indigoid anti-tuberculosis agent. *Bioorganic & Medicinal Chemistry Letters* **2014**, *24*, 268–270, doi:10.1016/j.bmcl.2013.11.024.
6. Steingruber, E. Indigo and Indigo Colorants. In *Ullmann's Encyclopedia of Industrial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, **2004** ISBN 978-3-527-30673-2.
7. Ensley, B.D.; Ratzkin, B.J.; Osslund, T.D.; Simon, M.J.; Wackett, L.P.; Gibson, D.T. Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science* **1983**, *222*, 167–169, doi:10.1126/science.6353574.
8. Fabara, A.N.; Fraaije, M.W. An overview of microbial indigo-forming enzymes. *Appl Microbiol Biotechnol* **2020**, *104*, 925–933, doi:10.1007/s00253-019-10292-5.
9. Lee, J.-H.; Wood, T.K.; Lee, J. Roles of Indole as an Interspecies and Interkingdom Signaling Molecule. *Trends in Microbiology* **2015**, *23*, 707–718, doi:10.1016/j.tim.2015.08.001.
10. Jaglin, M.; Rhimi, M.; Philippe, C.; Pons, N.; Bruneau, A.; Goustard, B.; Daugé, V.; Maguin, E.; Naudon, L.; Rabot, S. Indole, a Signaling Molecule Produced by the Gut Microbiota, Negatively Impacts Emotional Behaviors in Rats. *Front Neurosci* **2018**, *12*, doi:10.3389/fnins.2018.00216.
11. Osadchiy, V.; Labus, J.S.; Gupta, A.; Jacobs, J.; Ashe-McNalley, C.; Hsiao, E.Y.; Mayer, E.A. Correlation of tryptophan metabolites with connectivity of extended central reward network in healthy subjects. *PLOS ONE* **2018**, *13*, e0201772, doi:10.1371/journal.pone.0201772.

12. Negatu, D.A.; Liu, J.J.J.; Zimmerman, M.; Kaya, F.; Dartois, V.; Aldrich, C.C.; Gengenbacher, M.; Dick, T. Whole-Cell Screen of Fragment Library Identifies Gut Microbiota Metabolite Indole Propionic Acid as Antitubercular. *Antimicrob. Agents Chemother.* **2018**, *62*, doi:10.1128/AAC.01571-17.
13. Negatu, D.A.; Yamada, Y.; Xi, Y.; Go, M.L.; Zimmerman, M.; Ganapathy, U.; Dartois, V.; Gengenbacher, M.; Dick, T. Gut Microbiota Metabolite Indole Propionic Acid Targets Tryptophan Biosynthesis in *Mycobacterium tuberculosis*. *mBio* **2019**, *10*, doi:10.1128/mBio.02781-18.
14. Smith, J.E. *Biotechnology*; Cambridge University Press, **2009**; ISBN 978-1-139-47680-5.
15. Kamimura, N.; Goto, T.; Takahashi, K.; Kasai, D.; Otsuka, Y.; Nakamura, M.; Katayama, Y.; Fukuda, M.; Masai, E. A bacterial aromatic aldehyde dehydrogenase critical for the efficient catabolism of syringaldehyde. *Sci Rep* **2017**, *7*, doi:10.1038/srep44422.
16. Eaton, R.W.; Chapman, P.J. 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* **1995**, *177*, 6983–6988.
17. Namgung, S.; Park, H.A.; Kim, J.; Lee, P.-G.; Kim, B.-G.; Yang, Y.-H.; Choi, K.-Y. Ecofriendly one-pot biosynthesis of indigo derivative dyes using CYP102G4 and PrnA halogenase. *Dyes and Pigments* **2019**, *162*, 80–88, doi:10.1016/j.dyepig.2018.10.009.
18. Wu, Z.-L.; Podust, L.M.; Guengerich, F.P. Expansion of substrate specificity of cytochrome p450 2a6 by random and site-directed mutagenesis. *J. Biol. Chem.* **2005**, *280*, 41090–41100, doi:10.1074/jbc.M508182200.
19. Heine, T.; Großmann, C.; Hofmann, S.; Tischler, D. Indigoid dyes by group E monooxygenases: mechanism and biocatalysis. *Biological Chemistry* **2019**, *400*, 939–950, doi:10.1515/hsz-2019-0109.
20. Carretero-González, J.; Castillo-Martínez, E.; Armand, M. Highly water-soluble three-redox state organic dyes as bifunctional analytes. *Energy & Environmental Science* **2016**, *9*, 3521–3530, doi:10.1039/C6EE01883A.
21. Cheng, X.; Merz, K.-H.; Vatter, S.; Zeller, J.; Muehlbeyer, S.; Thommet, A.; Christ, J.; Wölfl, S.; Eisenbrand, G. Identification of a Water-Soluble Indirubin Derivative as Potent Inhibitor of Insulin-like Growth Factor 1 Receptor through Structural Modification of the Parent Natural Molecule. *Journal of Medicinal Chemistry* **2017**, *60*, 4949–4962, doi:10.1021/acs.jmedchem.7b00324.
22. Lin, G.-H.; Chen, H.-P.; Shu, H.-Y. Detoxification of Indole by an Indole-Induced Flavoprotein Oxygenase from *Acinetobacter baumannii*. *PLOS ONE* **2015**, *10*, e0138798, doi:10.1371/journal.pone.0138798.

23. Campbell, Z.T.; Baldwin, T.O. Fre Is the Major Flavin Reductase Supporting Bioluminescence from *Vibrio harveyi* Luciferase in *Escherichia coli*. *J Biol Chem* **2009**, *284*, 8322–8328, doi:10.1074/jbc.M808977200.
24. Frerichs-Deeken, U.; Fetzner, S. Dioxygenases Without Requirement for Cofactors: Identification of Amino Acid Residues Involved in Substrate Binding and Catalysis, and Testing for Rate-Limiting Steps in the Reaction of 1H-3-Hydroxy-4-oxoquinoline 2,4-dioxygenase. *Curr Microbiol* **2005**, *51*, 344–352, doi:10.1007/s00284-005-0065-3.
25. Steiner, R.A.; Janssen, H.J.; Roversi, P.; Oakley, A.J.; Fetzner, S. Structural basis for cofactor-independent dioxygenation of N-heteroaromatic compounds at the α/β -hydrolase fold. *Proc Natl Acad Sci U S A* **2010**, *107*, 657–662, doi:10.1073/pnas.0909033107.
26. Qu, Y.; Ma, Q.; Liu, Z.; Wang, W.; Tang, H.; Zhou, J.; Xu, P. Unveiling the biotransformation mechanism of indole in a *Cupriavidus* sp. strain. *Molecular Microbiology* **2017**, *106*, 905–918, doi:10.1111/mmi.13852.
27. Cohn, W.; Crawford, I.P. Regulation of enzyme synthesis in the tryptophan pathway of *Acinetobacter calcoaceticus*. *Journal of Bacteriology* **1976**, *127*, 367–379, doi:10.1128/JB.127.1.367-379.1976.
28. Donoso, R.; Leiva-Novoa, P.; Zúñiga, A.; Timmermann, T.; Recabarren-Gajardo, G.; González, B. Biochemical and Genetic Bases of Indole-3-Acetic Acid (Auxin Phytohormone) Degradation by the Plant-Growth-Promoting Rhizobacterium *Paraburkholderia phytofirmans* PsJN. *Appl Environ Microbiol* **2016**, *83*, doi:10.1128/AEM.01991-16.
29. Enders, T.A.; Strader, L.C. Auxin activity: Past, present, and future. *American Journal of Botany* **2015**, *102*, 180–196, doi:10.3732/ajb.1400285.
30. Tuomainen, M.; Lindström, J.; Lehtonen, M.; Auriola, S.; Pihlajamäki, J.; Peltonen, M.; Tuomilehto, J.; Uusitupa, M.; Mello, V.D. de; Hanhineva, K. Associations of serum indolepropionic acid, a gut microbiota metabolite, with type 2 diabetes and low-grade inflammation in high-risk individuals. *Nutr & Diabetes* **2018**, *8*, 1–5, doi:10.1038/s41387-018-0046-9.
31. de Mello, V.D.; Paananen, J.; Lindström, J.; Lankinen, M.A.; Shi, L.; Kuusisto, J.; Pihlajamäki, J.; Auriola, S.; Lehtonen, M.; Rolandsson, O.; et al. Indolepropionic acid and novel lipid metabolites are associated with a lower risk of type 2 diabetes in the Finnish Diabetes Prevention Study. *Sci Rep* **2017**, *7*, 46337, doi:10.1038/srep46337.
32. Zhao, Z.-H.; Xin, F.-Z.; Xue, Y.; Hu, Z.; Han, Y.; Ma, F.; Zhou, D.; Liu, X.-L.; Cui, A.; Liu, Z.; et al. Indole-3-propionic acid inhibits gut dysbiosis and endotoxin leakage to attenuate steatohepatitis in rats. *Exp Mol Med* **2019**, *51*, 1–14, doi:10.1038/s12276-019-0304-5.
33. Sun, C.-Y.; Lin, C.-J.; Pan, H.-C.; Lee, C.-C.; Lu, S.-C.; Hsieh, Y.-T.; Huang, S.-Y.; Huang, H.-Y. Clinical association between the metabolite

of healthy gut microbiota, 3-indolepropionic acid and chronic kidney disease. *Clinical Nutrition* **2019**, 38, 2945–2948, doi:10.1016/j.clnu.2018.11.029.

34. Konopelski, P.; Konop, M.; Gawrys-Kopczynska, M.; Podsadni, P.; Szczepanska, A.; Ufnal, M. Indole-3-Propionic Acid, a Tryptophan-Derived Bacterial Metabolite, Reduces Weight Gain in Rats. *Nutrients* **2019**, 11, 591, doi:10.3390/nu11030591.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to dr. Rolandas Meškys for guiding my research interests, thorough lessons about hyphens and dashes and life in general.

None of this work could have been accomplished without the significant contributions from my laboratory colleagues and publication co-authors, which are deeply acknowledged as well. Of special mention is dr. Renata Gasparavičiūtė and dr. Rita Meškienė for performing a grueling process of constructing metagenomic libraries and Justas Vaitekūnas for assistance whenever needed.

I acknowledge the financial support from Research Council of Lithuania.

Finally, I would like to thank my friends and family for continuous support and encouragement through all the process of studying and working on this thesis.

LIST OF PUBLICATIONS

Publication I. Casaite, V., **Sadauskas, M.**, Vaitekunas, J., Gasparaviciute, R., Meskiene, R., Skikaite, I., Sakalauskas, M., Jakubovska, J., Tauraite, D., Meskys, R. Engineering of a chromogenic enzyme screening system based on an auxiliary indole-3-carboxylic acid monooxygenase. *Microbiology Open*, 2019, 8, e795.

Publication II. **Sadauskas, M.**, Statkevičiūtė, R., Vaitekūnas, J., Petkevičius, V., Časaitė, V., Gasparavičiūtė, R., Meškys, R. Enzymatic synthesis of novel water-soluble indigoid compounds. *Dyes and Pigments*, 2020, 173, 107882.

Publication III. **Sadauskas, M.**, Vaitekūnas, J., Gasparavičiūtė, R., Meškys, R. Indole Biodegradation in *Acinetobacter* sp. Strain O153: Genetic and Biochemical Characterization. *Applied and Environmental Microbiology*, 2017, 83(19):e01453-17.



Publication IV. **Sadauskas, M.**, Statkevičiūtė, R., Vaitekūnas, J., Meškys, R. Bioconversion of biologically active indole derivatives with indole-3-acetic acid-degrading enzymes from *Caballeronia glathei* DSM50014. *Biomolecules*, 2020, 10(4):663.

COPIES OF PUBLICATIONS

Publication I

Casaite, V., **Sadauskas, M.**, Vaitekunas, J., Gasparaviciute, R., Meskiene, R., Skikaite, I., Sakalauskas, M., Jakubovska, J., Tauraite, D., Meskys, R. Engineering of a chromogenic enzyme screening system based on an auxiliary indole-3-carboxylic acid monooxygenase. *Microbiology Open*, 2019, 8, e795
DOI: 10.1002/mbo3.795

Engineering of a chromogenic enzyme screening system based on an auxiliary indole-3-carboxylic acid monooxygenase

Vida Časaitė*  | Mikas Sadauskas* | Justas Vaitekūnas* | Renata Gasparavičiūtė | Rita Meškienė | Izabelė Skikaitė | Mantas Sakalauskas | Jevgenija Jakubovska | Daiva Tauraitė | Rolandas Meškys 

Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Life Sciences Center, Vilnius University, Vilnius, Lithuania

Correspondence

Vida Časaitė, Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Life Sciences Center, Vilnius University, Vilnius, Lithuania.
Email: vida.casaitė@bchi.vu.lt

Funding information

This work was supported by the European Union's Horizon 2020 research and innovation program (BlueGrowth: Unlocking the potential of Seas and Oceans) under grant agreement no. 634486 (project acronym INMARE).

Abstract

Here, we present a proof-of-principle for a new high-throughput functional screening of metagenomic libraries for the selection of enzymes with different activities, pre-determined by the substrate being used. By this approach, a total of 21 enzyme-coding genes were selected, including members of xanthine dehydrogenase, aldehyde dehydrogenase (ALDH), and amidohydrolase families. The screening system is based on a pro-chromogenic substrate, which is transformed by the target enzyme to indole-3-carboxylic acid. The later compound is converted to indoxyl by a newly identified indole-3-carboxylate monooxygenase (Icm). Due to the spontaneous oxidation of indoxyl to indigo, the target enzyme-producing colonies turn blue. Two types of pro-chromogenic substrates have been tested. Indole-3-carboxaldehydes and the amides of indole-3-carboxylic acid have been applied as substrates for screening of the ALDHs and amidohydrolases, respectively. Both plate assays described here are rapid, convenient, easy to perform, and adaptable for the screening of a large number of samples both in *Escherichia coli* and *Rhodococcus* sp. In addition, the fine-tuning of the pro-chromogenic substrate allows screening enzymes with the desired substrate specificity.

KEYWORDS

aldehyde dehydrogenase, amidohydrolase, functional screening, indole-3-carboxylic acid monooxygenase, metagenomics

1 | INTRODUCTION

In light of the growing importance of biocatalysis, strategies that provide improvements in screening of novel enzymes are of considerable interest. Among other enzymes, aldehyde dehydrogenases (ALDHs), especially exhibiting a broad substrate spectrum, are

potential biocatalysts for biotechnology and are applicable in the detoxification of aldehydes, generated during metabolism of different natural and xenobiotic compounds (Kotchoni, Kuhns, Ditzer, Kirch, & Bartels, 2006; Lyu et al., 2017; Singh et al., 2014).

Metagenomics, which helps to circumvent the cultivation of bacteria and select genes directly from the environment, has

*These authors contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *MicrobiologyOpen* published by John Wiley & Sons Ltd.

become a powerful tool in search of new enzymes and metabolic pathways for the industrial biotechnology over the past decades (Allen, Moe, Rodbumer, Gaarder, & Handelsman, 2009; Maruthamuthu, Jiménez, Stevens, & Elsas, 2016; Suenaga, Ohnuki, & Miyazaki, 2007; Varaljay et al., 2016). Many studies show that the function-based screening or selection approaches permits an effective identification of different biocatalysts, such as lipases/esterases (Reyes-Duarte, Ferrer, & García-Arellano, 2012), cellulases (Maruthamuthu et al., 2016), and oxygenases (Nagayama et al., 2015), from diverse environmental sources and microbial habitats. However, the common problem in the search for new enzymes is the absence of an appropriate screening system. Usually, the functional screening of desired activities is based on chromogenic approach including the formation of blue indigo pigment, fluorogenic substrates, and/or sensors (Kennedy et al., 2011; Rüther, 1980; Seok et al., 2018; Shang, Chan, Wong, & Liao, 2018; Ye, Peng, Niu, Luo, & Zhang, 2018). Notwithstanding that several chromogenic substrates such as indole, indole carboxylic acids, and indole-3-carboxaldehyde applicable for plate and other high-throughput (HTP) assays have been developed and applied for screening various dioxygenases and broad substrate range monooxygenases (Celik, Speight, & Turner, 2005; Choi et al., 2003; Eaton & Chapman, 1995; Ensley et al., 1983; Furuya, Takahashi, Ishii, Kino, & Kirimura, 2004; McClay, Boss, Keresztes, & Steffan, 2005; O'Connor, Dobson, & Hartmans, 1997; Shi et al., 2013; Willetts, Joint, Gilbert, Trimble, & Mühlhng, 2012), a limited number of HTP methods for detection of other oxidoreductases, for example, aldehyde dehydrogenases, have been elaborated (Chen et al., 2014; Oyobiki et al., 2014; Reisinger et al., 2006; Seok et al., 2018; Wexler, Bond, Richardson, & Johnston, 2005). Moreover, those approaches are too restricted for a special substrate, cannot be used in a plate format or require sophisticated equipment.

The aim of this study was to develop a novel platform for the functional screening of the enzymes, particularly ALDHs. First, we searched for indole-3-carboxylic acid (I3CA)-degrading microorganisms and corresponding genes in metagenomes to determine whether any could transform I3CA to indigo. We have successfully identified *Icm* encoding gene, which was used for the creation of the screening method. By using the developed approach, we succeeded in a screening of diverse ALDHs with a broad substrate specificity. Furthermore, the auxiliary *Icm* enzyme was applied for screening of amidohydrolases using the amide of indole-3-carboxylic acid as a substrate. The *Icm* was active both in Gram-negative and Gram-positive bacteria, and hence, the enzyme was suitable for a functional screening of enzymes in different hosts.

2 | MATERIALS AND METHODS

2.1 | Chemicals

Chemicals used in this study are listed in Table A1. Gel resins were purchased from GE Healthcare (Little Chalfont, UK). Restriction endonucleases and DNA polymerases were from Thermo Fisher

Scientific (Vilnius, Lithuania). All reagents used in this study were of analytical grade.

2.2 | Bacterial strains, plasmids and media

The bacterial strains and plasmids used in this study are listed in Table 1. *Escherichia coli* and *Rhodococcus erythropolis* SQ1 cells were routinely grown in Luria-Bertani (LB) medium at 16–37°C. The following reagents were added to media as needed: IPTG, 40 µg/ml; ampicillin (Ap), 50 µg/ml; chloramphenicol (Cm), 20 µg/ml; kanamycin (Km) 50 µg/ml; tetracycline (Tc), 20 µg/ml; derivatives of I3CA, 1 mM.

2.3 | General DNA manipulation

Plasmid preparation, restriction endonuclease digestion, DNA ligation, agarose gel electrophoresis, and other standard recombinant DNA techniques were carried out by standard methods (Sambrook, Fritsch, & Maniatis, 1989). DNA sequencing and primer synthesis were performed commercially at the Macrogen (the Netherlands). DNA sequences were analyzed with a BLAST program available at the National Center for Biotechnology Information web site (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Evolutionary analyzes were conducted in MEGA7 (Kumar, Stecher, & Tamura, 2016).

2.4 | Screening of soil samples and gene cloning

About 1 g of soil samples were suspended in 1 ml 0.9% w/v NaCl solution, and 50 µl aliquots were spread on the agar plates supplemented with 1 mM I3CA. The plates were incubated at 30°C for 48 hr and were subsequently visually inspected for colonies producing the blue indigo pigment. Chromosomal DNA was isolated from the blue pigment producing bacteria, digested with the PstI restriction endonuclease and ligated in the pUC19 vector. *Escherichia coli* DH5α was used for screening of blue colonies on the plates supplemented with 1 mM I3CA.

For the screening assay, the pKVIABam8 encoding the *icm* gene was digested with BamHI and PstI and subcloned to the BamHI and PstI restriction sites of pACYC184 vector and resulted plasmid was designated pACYC-KVIA. For construction of expression vectors, *icm* gene was PCR-amplified with primers *KviaEcoR* and *KviaNde2F* (Table 1) and pKviaBam8 as a DNA template. All PCR amplifications were performed using Phusion High-Fidelity PCR Master Mix. PCR product was digested with NdeI/XhoI restriction endonucleases and ligated into pET-21a(+) previously digested with the corresponding enzymes to obtain pET21-KVIA. N-terminal His₆-tag was added by subcloning of the *icm* gene into pET28c(+), resulting in pET28-KVIA. For expression in *R. erythropolis* SQ1, the digested PCR fragment was ligated into pNitQC1 resulting in plasmid pNit-KVIA. To obtain N-terminal fusion of *Icm* with maltose-binding protein (MBP), *malE* was amplified with primers MBP_F and MBP_R_Nco, digested with XbaI/NcoI, and ligated into pET28-KVIA resulting in pET28-MBP-KVIA. To obtain N-terminal fusion with Strep-Tag, *Icm* encoding gene

TABLE 1 Strains, plasmids, and primers used in this study

	Relevant characteristics	Source
Strain		
<i>Escherichia coli</i> DH5 α	endA1, gyrA96, hsdR17, recA1, relA1, supE44, thi-1, Δ (lacZYA-argF)U169, Φ d80lacZ Δ M15	Sambrook et al. (1989)
<i>E. coli</i> BL21(DE3)	F ⁻ , ompT, hsdSB (rB ⁻ , mB ⁻), dcm, gal, λ (DE3)	Novagen
<i>Rhodococcus erythropolis</i> SQ1	Wild-type strain	Quan and Dabbs (1993)
<i>Bosea</i> sp. KVIA	Soil isolate	This study
Plasmid		
pET-21c (+)	Expression vector Ap ^r , f1, pBR322 ori	Novagen
pET-28c (+)	Expression vector Km ^r , f1, pBR322 ori	Novagen
pASK-IBA3	Ap ^r expression vector, Strep-tag®, f1 origin	IBA lifesciences
pUC19	Ap ^r , cloning vector pMB1ori	Thermo Fisher
pACYC184	Cm ^r , Tc ^r cloning vector p15A ori	Thermo Fisher
pNitQC1	Ap ^r (<i>E. coli</i>), Cm ^r (<i>R. erythropolis</i> SQ1)	Nakashima and Tamura (2004a)
pNitRT1	Ap ^r (<i>E. coli</i>), Tc ^r (<i>R. erythropolis</i> SQ1)	
pKVIAbam8	pUC19 cloned lcm gene, Ap ^r	This study
pACYC-KVIA	pACYC184 cloned lcm gene, Cm ^r	This study
pET21-KVIA	pET21a cloned lcm gene, Ap ^r	This study
pET28-KVIA	N-terminal His ₆ -tagged lcm, Km ^r	This study
pET28-MBP-KVIA	N-terminal MBP-tagged lcm, Km ^r	This study
pASK-IBA3-KVIA	N-terminal Strep-tagged lcm, Ap ^r	This study
pET21-MO13	C-terminal His ₆ tag Amidohydrolase gene, Ap ^r	This study
pNitQC1-KVIA	lcm gene for expression in <i>Rhodococcus</i> sp.	This study
pNitRT1-Vmix	Vmix ALDH gene for expression in <i>Rhodococcus</i> sp	This study
Primer, 5'-3'		
MBP_R_Nco	TTCCATGGGCCCTGGAACAG	This study
MBP_F	GTGAGCGGATAACAATTCC	This study
KviaEcoR	GAGAATTCGCCATAGATCAGGACC	This study
KviaNde2F	TACATATGAAGTCATCATCGTAG	This study
Kvia-IBA3-F	GAGCGGGTCTCGAATGAAGGTCATCATC	This study
Kvia-IBA3-R	CTGCGAGGTCTCAGCGCTGGACCGCCGCGC	This study
VmixNdeF	TACATATGAGTGC GAACGATATTTAAAC	This study
VmixHindR	CCAAGCTTCAGAACGGAAACCCGC	This study
am13F	GCCATATGGAAAAATCATCATAC	This study
am13R2	ACTCGACCTGGGATTAATAG	This study

was amplified with primers Kvia-IBA3-F and Kvia-IBA3-R, digested with Eco311 and ligated into Eco311-digested pASK-IBA3, resulting in pASK-IBA3-KVIA. For cloning of aldehyde dehydrogenase Vmix gene, it was PCR-amplified with primers VmixHindR and VmixNdeF (Table 1) and DNA from the metagenome clone Vmix as template. PCR product was digested with NdeI/HindIII restriction endonucleases and ligated into pNitRT1 previously digested with the corresponding enzymes to obtain pNitRT-Vmix. For construction of C-terminal His₆-tagged amidohydrolase, MO13 gene was PCR-amplified with primers am13F and am13R2 (Table 1) and pMO13 as DNA matrix. PCR product was digested with NdeI/XhoI restriction endonucleases and ligated into

pET-21a(+) previously digested with the corresponding enzymes to obtain pET21-MO13. Electrocompetent cells were prepared as described previously (Nakashima & Tamura, 2004b; Stanislauskienė et al., 2012) and used for transformation.

2.5 | Construction of the metagenomic library and screening for enzymes

For the construction of environmental DNA libraries, surface soils (0–15 cm) from a different fields in district Vilnius (Lithuania) were collected. The environmental DNA was isolated from samples using

ZR Soil Microbe DNA Kit (Zymo Research), partially digested with the endonucleases PstI or HindIII and ligated in the pUC19 vector. To analyze the number of clones in the library, quality of the library (a ratio of white/blue colonies), and the average insert length, *E. coli* DH5 α cells were transformed with ligation mixtures and spread on LB agar plates supplemented with ampicillin, 1 mM IPTG, and 1 mM X-gal. Eight white colonies-forming clones from each library were chosen for plasmid DNA isolation and analysis of the length of the insert. For functional screening, *E. coli* DH5 α cells harboring pACYC-KVIA were transformed with the metagenomic libraries and plated on LB agar plates containing Ap, Cm, as needed and 1 mM solution of derivative I3CA. The plates were incubated at 37°C for 2 days and were subsequently screened for colonies that were able to produce the blue pigment indigo by visual detection. The positive clones were subjected for DNA sequencing. The sequences obtained in the present study were deposited to the GenBank database under the accession numbers MG770119–MG770138, MG786188, MG786189, MG775032, MK284926, and MH476458. The full list is given in Table A2.

2.6 | Expression and purification of the recombinant proteins

For gene expression, *E. coli* BL21 (DE3) were transformed with pET21-KVIA, pET28-KVIA, pET28-MBP-KVIA, pASK-IBA3-KVIA, and pET21-MO13. The cells were grown at 30°C with rotary shaking until OD₆₀₀ reached 0.8, and gene expression was induced with 0.05–0.5 mM IPTG for pET plasmids and 200 μ g/L anhydrotetracycline for pASK-IBA3 plasmid. The cells were incubated at 16–30°C for either 4 hr or overnight, collected by centrifugation, suspended in lysis buffer (50 mM Tris-HCl, pH 8.0, containing 150 mM of NaCl), disrupted by sonication and the lysates were used as total protein sample, while centrifugation-clarified lysates (16,000 g for 10 min) were treated as a soluble fraction. The recombinant proteins were analyzed with 12% denaturing SDS–PAGE.

Purification through His₆-tag was carried out with nickel HisTrap™ HP column according to the manufacturer's instructions. Strep-tagged protein was purified with Strep-Tactin XT Starter Kit according to manufacturer's protocol. MBP-fused protein was purified by MBP-starch affinity chromatography using commercial grade cationic starch-packed column essentially as described in (Duong-Ly & Gabelli, 2015). Protein quantification was performed by densitometry with GelAnalyzer software (Pavel & Vasile, 2012) using different concentrations (100, 250, and 500 μ g/ml) of bovine serum albumin (ThermoFisher Scientific) as standard.

2.7 | Bioconversion of aldehydes or carboxylic acids by whole cells

The *E. coli* or *R. erythropolis* SQ1 cells transformed with the appropriate plasmids were grown aerobically in LB containing appropriated antibiotic at 30°C until optical density reached 0.8 (A_{600}), then 0.5 mM of IPTG was added and cells were grown aerobically at

30°C for 12 hr. Cells were harvested by centrifugation, washed with 50 mM potassium phosphate buffer (pH 7.2), suspended in the same buffer and used as the whole cells. Then, 1 mM solutions of substrates were added, and bioconversion reactions were carried out at 30°C with shaking at 180 rpm for 1–24 hr. The conversion was followed by changes in UV absorption spectrum in 200–400 nm range or by HPLC/MS analysis, as described previously (Stankevičiūtė et al., 2016).

2.8 | Monoxygenase activity assay

The monoxygenase activity was evaluated from the decrease of the absorbance at 340 nm due to oxidation of NADH or NADPH ($\epsilon_{340} = 6,220$ M/cm), using spectrophotometer and was performed at room temperature. Simultaneously, reaction mixtures were incubated overnight at 30°C and inspected for the formation of blue precipitate. A total reaction volume of 1 ml contained 50 mM Tris-HCl, pH 7.5, 1 mM I3CA, different amounts (1–20 mM) of NADH or NADPH and 50 μ M of flavin (FAD, FMN or riboflavin). Reactions were initiated by adding 2.5 μ g of the purified enzyme or 20 μ l of the soluble fraction (approx. 10 μ g of total protein).

2.9 | Aldehyde dehydrogenase activity assay

For colorimetric assay, the cells were disrupted by sonication and the cell-free extracts were used to analyze the ALDH activity as described in (Bianchi et al., 2017). In brief, the obtained supernatants were mixed with NAD⁺ (200 μ M) and NADP⁺ (200 μ M), nitroblue tetrazolium chloride (NBT, 200 μ M), phenazine methosulfate (PMS, 20 μ M), and an appropriate aldehyde (200 μ M) in 50 mM Tris-HCl buffer, pH 8.0, at 30°C. A total reaction volume of 200 μ l contained 50 μ l of cell lysates (approx. 20 μ g of total protein), and the reaction was followed spectrophotometrically ($\lambda = 580$ nm) in 96-well microtiter plates by monitoring the production of formazan dye after 1 and 3 hr.

2.10 | Amidohydrolase activity assay

A total reaction volume of 0.5 ml contained 50 mM Tris-HCl, pH 8.5, and 1 mM of appropriate substrate. Reactions were initiated by adding 2.5 μ g of the purified enzyme. The progress of the reaction was followed by changes in UV absorption spectrum in 200–600 nm range or by HPLC/MS analysis, as described previously (Stankevičiūtė et al., 2016).

2.11 | Synthesis of N-(3-hydroxypropyl)-indole-3-carboxamide

A solution of indole-3-carboxylate (100 mg, 0.62 mmol) and *N,N,N',N'*-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU, 235.3 mg, 0.62 mmol) in dimethylformamide (1.24 ml) was vigorously stirred for 30 min at room temperature. Then, 3-amino-1-propanol (46.6 mg, 0.62 mmol) and triethylamine

(86.5 μ l, 0.62 mmol) were added to the reaction mixture and continued stirring for additional 12 hr at the same temperature. The reaction mixture was diluted with water (10 ml) and extracted with ethyl acetate (3 \times 15 ml). The organic phase was dried (Na_2SO_4) and the solvent evaporated under reduced pressure. The residue was purified by column chromatography (silica gel, chloroform/methanol mixture). Yield 65 mg (48%). Synthesized derivative was characterized by NMR spectroscopy and HPLC/MS analysis. NMR spectra were recorded in $\text{DMSO}-d_6$ on a Bruker Ascend 400: ^1H NMR-400 MHz, ^{13}C NMR-100 MHz. Chemical shifts (δ) are reported in ppm relative to the solvent resonance signal as an internal standard. MS (ESI⁺): m/z 219 [M+H]⁺, 217 [M-H]⁻.

^1H NMR ($\text{DMSO}-d_6$): δ = 1.64–1.74 (m, 2H, CH_2), 3.32 (dd, 2H, J = 12.8, 6.7 Hz, CH_2), 3.48 (dd, 2H, J = 12.7, 6.4 Hz, CH_2), 4.52 (bs, 1H, OH), 7.06–7.18 (m, 2H, CH), 7.42 (d, 1H, J = 7.8 Hz, CH),

7.87 (t, 1H, J = 5.5 Hz, NH), 7.99 (d, 1H, J = 2.9 Hz, CH), 8.13 (d, 1H, J = 7.7 Hz, CH), 11.52 (s, 1H, NH). ^{13}C NMR ($\text{DMSO}-d_6$): δ = 33.31, 46.23, 59.15, 111.22, 112.22, 120.68, 121.42, 122.22, 126.52, 128.00, 136.57, 165.19.

3 | RESULTS AND DISCUSSION

3.1 | Cloning and identification of indole-3-carboxylate monoxygenase

To screen enzymes displaying an indigo-forming activity in the presence of I3CA, two approaches were used. Initially, several blue colonies-forming bacteria were screened using soil samples and the agar plates supplemented with I3CA. One of these isolates, KVIA, was chosen for further studies. The analysis of the 16S rRNA

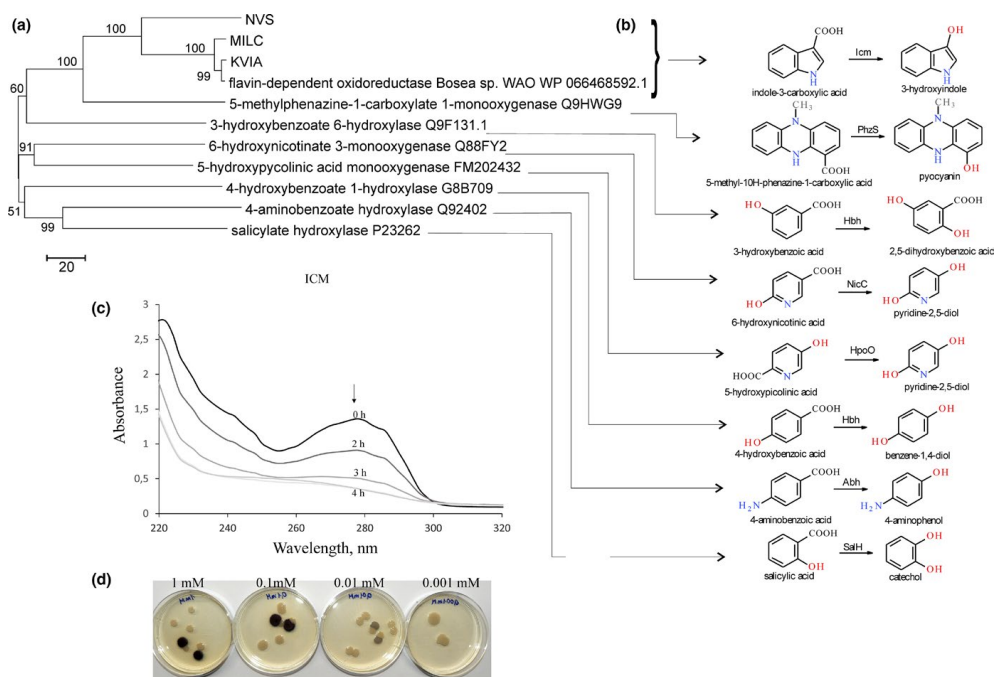


FIGURE 1 Characterization of Icm-KVIA. (a) Evolutionary relationship of decarboxylating flavin-dependent oxidoreductases. The evolutionary history was inferred using the Neighbor-Joining method (Saitou & Nei, 1987), the evolutionary distances were computed using the number of differences method (Nei & Kumar, 2000) and are in the units of the number of amino acid differences per sequence. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test are shown next to the branches (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. (b) Hydroxylation reactions performed by Icm-related enzymes. SalH, salicylate-1-hydroxylase (EC 1.14.13.1); Hbh, 4-hydroxybenzoate hydroxylase (EC 1.14.13.64); Abh, 4-aminobenzoate 1-monoxygenase (EC 1.14.13.27); PhzS, 5-methylphenazine-1-carboxylate 1-monoxygenase (EC 1.14.13.218); NicC, 6-hydroxynicotinate 3-monoxygenase (EC 1.14.13.114); HpoO, 5-hydroxypicolinate monoxygenase; Icm-KVIA, I3CA monoxygenase. (c) Time-course of consumption of I3CA by Icm producing *Escherichia coli* cells. Primary spectrum is black; spectra after two, three, and four hours are depicted in brightening gray. (d) Colonies of *E. coli* DH5 α on the plates supplemented with varied concentration of I3CA, blue colonies contain pACYC-KVIA plasmid, white colonies—an empty pACYC184 vector

gene sequence (GenBank accession No. MG775032) revealed that the bacteria belonged to the *Bosea* genus. The genomic library of *Bosea* sp. KVIA was constructed, and the positive clone harboring the plasmid pKVIABam8 was identified based on the ability to form blue colonies on the plates supplemented with I3CA. The nucleotide sequence analysis showed one 1,242 bp long ORF in the insert. The ORF encoded a 414 aa long protein, which was 98% identical to the hypothetical flavin-dependent oxidoreductase from *Bosea* sp. WAO (GenBank accession No. WP_066468592). Two additional blue colonies-forming clones were selected from the metagenomic libraries on I3CA agar plates. Both hits, named MILC and NVS, encoded the proteins, which shared 95.7% and 62.7% identity to the protein encoded by the pKVIABam8 plasmid, respectively. According to the sequence analysis, all three screened proteins (KVIA, MILC and NVS clones) belonged to the group A of flavin monooxygenases, which depend on NAD(P)H as external electron donor and contain a glutathione reductase (GR-2) type Rossmann fold (GXGXG) for FAD binding. Moreover, several conserved motifs such as DGX₂R, and GDAX₁₀GX₄DX₃L characteristic for monooxygenases were identified (Huijbers, Montersino, Westphal, Tischler, & Berkel, 2014). Some dioxygenases such as cumate and *m*-toluate dioxygenases convert indole-2-carboxylic acid and I3CA to indigo. The dioxygenases incorporate two atoms of molecular oxygen, leading to the formation of 2,3-dihydroxyindoline-3-carboxylate. Subsequent reactions are spontaneous and lead to the mixture of indigo, isatin, and indirubin.

Moreover, those enzymes are also active toward indole (Eaton & Chapman, 1995). In contrast, the enzymes encoded by the KVIA, MILC, and NVS clones were unrelated to any known dioxygenase and showed the highest sequence similarity to the experimentally characterized monooxygenases such as 5-methylphenazine-1-carboxylate 1-monooxygenase from *Pseudomonas aeruginosa* PAO1 and 3-hydroxybenzoate-6-hydroxylase from *Pseudomonas alcaligenes* or *Klebsiella oxytoca*. Moreover, the identified enzymes were not active toward indole since the clones did not form colored colonies in the presence of this substrate. In addition, no substrate consumption was observed (HPLC-MS analysis) when nicotinic, 2- and 4-picolinic, 5-hydroxypiperazine-2-carboxylic, salicylic acid, indoline-2-carboxylic, indole-2-carboxylic, indole-4-carboxylic, indole-5-carboxylic, indole-6-carboxylic, and indole-7-carboxylic were used as substrates for indole-3-carboxylate monooxygenase (Icm). We also tested this enzyme with 5-nitroindole-3-carboxylic, 7-methylindole-3-carboxylic, 1-methylindole-3-carboxylic and as well as indole-3-carboxaldehyde, indole-3-carbonitrile or methyl ester of indole-3-carboxylic acid for formation of the blue colonies on plates. No color changes were observed using these derivatives of indole-3-carboxylic acid. Based on sequence analysis and substrate specificity, we designated the identified enzyme as an indole-3-carboxylate monooxygenase (Icm). We proposed that Icm performed an oxidative decarboxylation reaction like other known flavin-dependent monooxygenases that catalyze the decarboxylative hydroxylation of aromatic

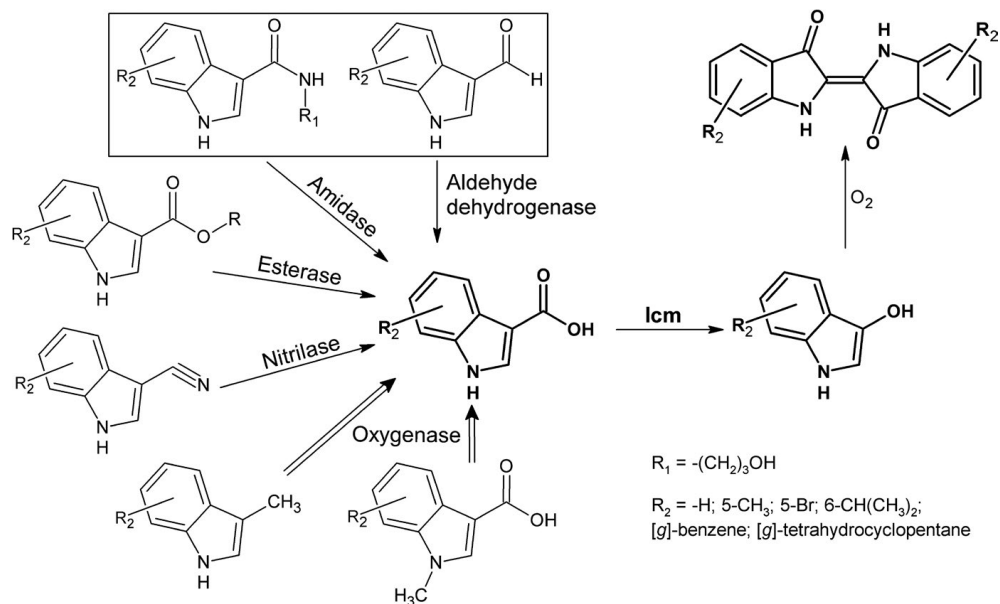


FIGURE 2 The principal scheme of the functional screening of enzymes based on an auxiliary Icm enzyme. The experimentally tested substrates are boxed. R: any radical

carboxylic acids (Figure 1b) such as the salicylate monooxygenase from *Pseudomonas putida* (Uemura et al., 2016), 6-hydroxynicotinic acid 3-monooxygenases NicC from *P. putida* and *Bordetella bronchiseptica* (Hicks et al., 2016), 5-methyl phenazine-1-carboxylate-1-monooxygenase PhzS from *P. aeruginosa* (Mavrodi et al., 2001), 4-hydroxybenzoate 1-hydroxylase from *Candida parapsilosis* (Van Berkel, Eppink, Middelhoven, Vervort, & Rietjens, 1994), 4-aminobenzoate monooxygenase from *Agaricus bisporus* (Tsuji, Ogawa,

Bando, & Sasaoka, 1986). The relationship between similar enzymes is shown in the phylogenetic tree (Figure 1a).

3.2 | Expression, protein purification, and characterization of the lcm

To characterize lcm in more detail, the gene encoding lcm was cloned to several expression plasmids, fusing it to His₆-Tag,

TABLE 2 Functional annotation of clones with aldehyde dehydrogenase activity

Clone	Protein length, aa	The nearest homolog, accession no	Identity, %
DON4	482	Salicylaldehyde dehydrogenase <i>Betaproteobacteria bacterium</i> OGA51247.1	77
JU61	507	Salicylaldehyde dehydrogenase <i>Hydrogenophaga</i> sp. Root209 WP_056264373	97
pALD442	483	Salicylaldehyde dehydrogenase <i>Cupriavidus</i> sp. BIS7 WP_019448853	89
pALD458	484	Phenylacetaldehyde dehydrogenase <i>Alcaligenes faecalis</i> WP_060185347	99
pALDBS21	515	Phenylacetaldehyde dehydrogenase <i>Alcaligenes faecalis</i> WP_045929579	96
pALDBSal	436	NAD(P)-dependent benzaldehyde dehydrogenase <i>Pseudomonas putida</i> WP_016501743	99
pALDGA1	483	Salicylaldehyde dehydrogenase <i>Afipia massiliensis</i> WP_046830129	91
pALDJU6	488	Phenylacetaldehyde dehydrogenase <i>Pseudomonas</i> sp. MIACH WP_053136087	96
pALDMO9	485	Aldehyde dehydrogenase <i>Bacillus thermoamylovorans</i> WP_041902008	82
pALDMO11	487	Benzaldehyde dehydrogenase <i>Stenotrophomonas</i> sp. LM091 WP_070425978	98
pALDR177	768	Xanthine dehydrogenase family protein molybdopterin-binding sub unit <i>Rhizobium</i> sp. Root564 WP_062426820	99
	329	Xanthine dehydrogenase family protein subunit M <i>Rhizobium</i> sp. Leaf155 WP_062597871	98
	182	(2Fe-2S)-binding protein <i>Rhizobium</i> sp. WP_062442533	97
pALDSV3	485	Aldehyde dehydrogenase <i>Pseudomonas</i> sp. A214 WP_076384861	72
pEGA1	504	Aldehyde dehydrogenase <i>Microbacterium pygmaeum</i> WP_091486269	73
pEMMO	484	Benzaldehyde dehydrogenase <i>Acinetobacter</i> sp. ANC 3832 WP_086192356	85
pER2AH	491	Aldehyde dehydrogenase <i>Arthrobacter</i> sp. Leaf69 WP_056430460	94
pER2AH2	490	Salicylaldehyde dehydrogenase <i>Arthrobacter</i> sp. P2b WP_079598892	98
pRG1	501	Aldehyde dehydrogenase family protein <i>Bacillus</i> sp. WP_057215027.1	91
pRG2	490	Aldehyde dehydrogenase <i>Pseudomonas fulva</i> WP_013791146	90
URAGR	472	Benzaldehyde dehydrogenase <i>Agrobacterium</i> sp. SCN 61-19 ODS51427	85
Vmix	490	Phenylacetaldehyde dehydrogenase <i>Verrucomicrobia</i> sp. OHE78850	73

Strep-Tag, maltose-binding protein (MBP) or glutathione S-transferase (GST) or without any tag for protein expression. Also, the plasmid (pNitQC1-KVIA) for protein expression in *R. erythropolis* SQ1 cells was created. Only the N-terminal fusion of Icm with MBP (His₆-MBP-His₆-Icm) resulted in partially soluble protein (Table A3). Conventional optimization strategies (variation of temperature, inductor concentration, cell density, expression host, buffer composition, etc.) did not result in significant improvement of protein solubility. Once outside the cell, the activity of Icm diminished. No in vitro activity was detected with the purified His₆-MBP-His₆-Icm by using different flavin cofactors and following the oxidation of either NADH or NADPH. Similarly, neither substrate consumption nor any intermediate products were detected by HPLC/MS, and no blue precipitate was formed in these in vitro reactions.

Since the active purified protein could not be obtained, further work was carried out using the whole cells of recombinant *E. coli* or *R. erythropolis* SQ1 bacteria. It was found that I3CA was consumed by all Icm derivatives at a similar rate (Figure A1). The amount of a blue precipitate formed during the bioconversion of I3CA corresponded to the consumption of this substrate. Meanwhile, no pigment appeared in the control reactions, in which the cells transformed with blank vectors were used. This indicates that Icm is active inside the cell and is involved in the conversion of I3CA to indigo blue.

3.3 | Application of Icm as an auxiliary enzyme for functional screening of aldehyde dehydrogenases

Despite the fact that Icm activity was not detected in vitro, *E. coli* cells harboring the *icm* gene readily produced a blue indigo dye on the agar plates supplemented with I3CA. This property was further exploited to create a system for a functional screening of metagenomic libraries. The idea was to use the appropriate substrate, for example indole-3-carboxaldehyde, which would be converted to I3CA by the target enzyme, in this case ALDH. Then, Icm as an auxiliary enzyme would oxidize I3CA into indigo; hence, the colored *E. coli* colonies would indicate the presence of the active ALDH (Figure 2). To test such screening platform, the *icm* gene was subcloned into the pACYC184 vector, compatible with the pUC19, which was used for creation of metagenomic DNA libraries. The *E. coli* DH5 α cells transformed with pACYC-KVIA produced blue colonies on the agar plates supplemented with I3CA (0.01 mM of I3CA in the medium was sufficient for the formation of blue pigment (Figure 1d), but only white colonies were observed when indole-3-carboxaldehyde was used as a substrate. Therefore, this strain was further used for screening of metagenomic libraries.

Twenty-one metagenomic libraries were created using the pUC19 plasmid and DNA isolated from soil. Each library contained clones with inserts of ~3–15 kb average size, yielding approximately 0.5 Gb of total cloned genomic DNA per library. In order to screen

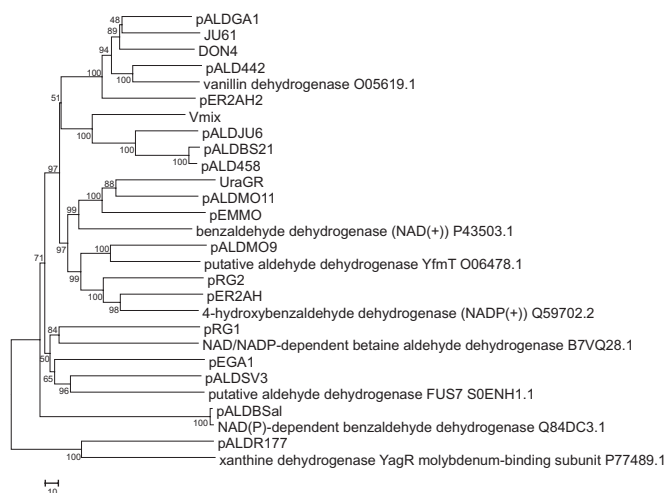


FIGURE 3 Sequence diversity of metagenomic aldehyde dehydrogenases. The evolutionary history was inferred using the Neighbor-Joining method (Saitou & Nei, 1987), the evolutionary distances were computed using the number of differences method (Nei & Kumar, 2000) and are in the units of the number of amino acid differences per sequence. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test are shown next to the branches (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The optimal tree with the sum of branch length = 2,362.75195313 is shown. All positions containing gaps and missing data were eliminated. A total of 351 positions in the final dataset were used for data analysis

for ALDH activity, about 30,000 clones per library were spread on LB agar supplemented with indole-3-carboxaldehyde. In this way, 52 indigo-forming clones were identified. The clones producing indigo without the presence of Icm (the false positives, i.e., most of such clones encoded Baeyer–Villiger monooxygenases, data not shown) as well as redundant clones were omitted resulting in 20 unique hits harboring the distinct genomic fragments. The sequence analysis of the screened ALDH-positive clones revealed the presence of genes encoding the proteins that were 73%–99% identical to the known sequences in the NCBI databank and homologous to ALDHs (19 clones), and molybdopterin xanthine dehydrogenase (one clone; see Table A2). Thus, the proposed functional screening approach was suitable for identification of hits expressing ALDHs (Table 2). To gain insight into the phylogenetic relationship of all selected enzymes, the phylogenetic tree was constructed (Figure 3). As revealed by comparison between UniProtKB/SwissProt sequences, nine ALDHs, that is, pDON4, pALDGA1, JU61, pALD442, pER2AH2, Vmix, pALDJU6, pALDBS21, and pALD458 were closest to vanillin dehydrogenase, pALDMO9 was related to *B. subtilis* vanillin dehydrogenase. pEMMO, pALDMO11, and UraGR were related to NAD(+)-dependent benzaldehyde dehydrogenase and pALDBSal to NAD(P)-dependent benzaldehyde dehydrogenase. The sequences of clones pRG1, pEGA1, and pALDSV3 were closest to betaine aldehyde dehydrogenase.

Also, two 4-hydroxybenzaldehyde dehydrogenase-like enzymes were selected (pER2AH, pRG2).

To analyze a substrate specificity of the screened enzymes, the bioconversion of substrates by whole cells was monitored by UV-Vis spectrophotometer and products of the reaction were confirmed by HPLC-MS analysis (Table 3). For some substrates, the colorimetric assay based on the formation of formazan by the cell-free extracts was applied (Table 4). Thirteen derivatives of indole-3-carboxaldehyde were tested. The most preferred substrates among the tested ones were 5-bromindole-3-carboxaldehyde, 6-benzoyloxyindole-3-carboxaldehyde, and 1H-benzo[g]indole-3-carboxaldehyde, which were oxidized by 18 ALDHs (Table 3). Only one strain (pALDR177) could oxidize 2-phenylindole-3-carboxaldehyde. The whole cells with an empty vector (*E. coli* DH5α/pUC19) did not show any activity on the tested substrates, confirming that the ALDHs were encoded by the metagenomic inserts. Even though among aldehydes without indole ring, the favorable substrate was 3-hydroxybenzaldehyde, which was oxidized by 19 clones, the hits showed very different substrate specificity (Table 4), and hence, the offered screening platform allowed the identification of ALDHs both of different structures and catalytic properties.

To test whether the screening of ALDHs could be carried out in another bacterial host, one ALDH gene was subcloned to the

TABLE 3 Activity of aldehyde dehydrogenases toward derivatives of indole-3-carboxaldehyde

	5BR3C	6B3C	1HB3C	5M3C	T3C	6I3C	BT3C	1M3C	4N3C	5B3C	1,2MHC	4B3C	2P3C
pALDR177	+	+	+	+	+	+	+	+	+	+	+	+	+
pALDBS21	+	+	+	+	+	+	+	+	+	+	+	-	-
JU61	+	+	+	+	+	+	+	+	+	+	-	-	-
pALD442	+	+	+	+	+	+	+	+	+	+	-	-	-
pALDGA1	+	+	+	+	+	+	+	+	+	+	-	-	-
pALDBSal	+	+	+	+	+	+	+	+	+	+	-	-	-
pRG1	+	+	+	+	+	+	+	+	+	-	-	-	-
PER2AH2	-	+	+	+	+	+	+	+	+	+	-	-	-
pALDMO11	+	+	+	-	+	+	+	+	+	-	-	-	-
pEGA1	-	+	+	+	-	-	+	+	-	+	-	+	-
pALDMO9	+	+	+	+	+	+	-	-	+	-	-	-	-
pER2AH	+	+	-	-	+	+	+	+	+	-	-	-	-
Vmix	+	+	+	+	+	-	-	+	-	-	-	-	-
pRG2	+	+	+	+	-	+	-	-	-	-	-	-	-
pALDSV3	+	-	+	+	+	+	-	-	-	-	-	-	-
DON4	+	+	+	+	+	-	-	-	-	-	-	-	-
pALDJU6	+	+	+	+	+	-	-	-	-	-	-	-	-
pALD458	+	+	+	+	-	-	-	-	-	-	-	-	-
pEMMO	+	+	-	-	-	+	+	-	-	-	-	-	-
URAGR	+	-	+	-	-	-	-	-	-	-	-	-	-

Note. 1,2MHC: 1,2-dimethyl-indole-3-carboxaldehyde; 1HB3C: benzo[g]indole-3-carboxaldehyde; 1M3C: 1-methylindole-3-carboxaldehyde; 2P3C: 2-phenylindole-3-carboxaldehyde; 4B3C: 4-benzoyloxyindole-3-carboxaldehyde; 4N3C: 4-nitroindole-3-carboxaldehyde; 5B3C: 5-benzoyloxyindole-3-carboxaldehyde; 5BR3C: 5-bromoindole-3-carboxaldehyde; 5M3C: 5-methylindole-3-carboxaldehyde; 6B3C: 6-benzoyloxyindole-3-carboxaldehyde; 6I3C: 6-isopropylindole-3-carboxaldehyde; BT3C: benzo[b]thiophene-3-carboxaldehyde; T3C: 1,6,7,8-tetrahydrocyclopenta(g)indole-3-carboxaldehyde; "+": the reaction product was observed by the HPLC-MS analysis and/or by the UV-VIS spectrum; "-": no conversion.

TABLE 4 Substrate specificity of aldehyde dehydrogenases

	3HBA	VAN	MFU	FU	DHBA	BA	SA	NA	TCA	PYCA	DMBA	3CHCA	CHCA	FAA	2PHPA
pALDBSal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
PER2AH	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
pALDR177	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+
pALDGA1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
pALD442	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
URAGR	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
PER2AH2	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
pRG2	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-
JU61	+	+	+	+	-	+	+	+	+	+	+	+	+	+	-
pEGA1	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-
pALDMO11	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-
pEMMO	+	-	+	+	+	+	+	+	+	+	+	+	+	-	-
pALDSV3	+	+	+	+	-	+	+	-	+	+	-	+	+	+	-
pRG1	+	+	+	+	+	+	+	+	-	+	+	-	-	-	-
DON4	+	-	+	+	+	+	+	+	+	-	-	+	-	+	-
Vmix	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
pALDMO9	+	+	-	+	+	-	-	+	-	-	+	-	-	-	-
pALD458	+	+	+	-	+	-	-	+	-	-	-	-	-	-	-
pALDBS21	+	+	-	+	+	-	-	+	-	-	-	-	-	-	-
pALDJU6	-	+	+	+	-	+	-	-	-	-	-	-	-	-	-

Note: 2PHPA: 2-phenylpropanaldehyde; 3CHCA: 3-cyclohexene-1-carboxaldehyde; 3HBA: 3-hydroxybenzaldehyde; BA: benzaldehyde; CHCA: cyclohexanecarboxaldehyde; DHBA: 3,4-dihydroxybenzaldehyde; DMBA: 4-(dimethylamino)benzaldehyde; FAA: phenylacetaldehyde; FU: furfural; MFU: 5-methylfurfural; NA: 1-naphthaldehyde; PYCA: pyrrole-2-carboxaldehyde; SA: salicylaldehyde; TCA: trans-cinnamaldehyde; VAN: vanillin; "+": The reaction product was observed by the colorimetric assay in the cell-free extracts of recombinant *Escherichia coli*; "-": the concentration of the resulting formazan dye was not different from the control.

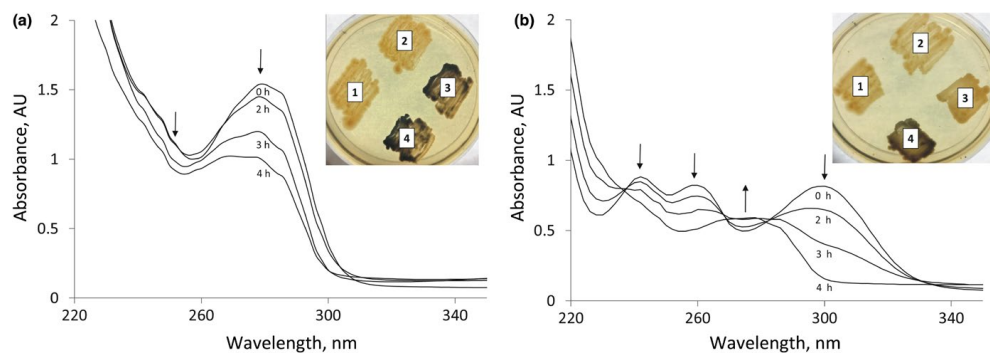


FIGURE 4 Icm and aldehyde dehydrogenase Vmix activity in the *Rhodococcus erythropolis* SQ1 cells. (a) Consumption of I3CA by cells carrying pNitQC1-Icm plasmid, (b) bioconversion of I3CA by cells carrying pNitQC1-KVIA and pNitRT1-Vmix plasmids. *R. erythropolis* SQ1 cells with different plasmid combinations were plated on I3CA (a) and indole-3-carboxaldehyde (b). 1—pNitQC1 + pNitRT1 (two empty vectors), 2—pNitQC1 + pNitRT1-Vmix (one empty vector and the other one carrying *aldh* gene), 3—pNitQC1-Icm + pNitRT1 (one empty vector and the other one carrying *icm* gene), 4—pNitQC1-Icm + pNitRT1-Vmix (*icm* and *aldh* genes)

pNitRT1 plasmid for expression in *R. erythropolis* SQ1. It turned out, that the cells transformed with pNitRT1-Vmix and pNitQC1-KVIA could produce indigo dye on the plates supplemented with indole-3-carboxaldehyde (Figure 4). Considering the fact that not all enzymes encoded in the metagenome can be active in *E. coli* cells, the Gram-positive host such as *Rhodococcus* sp. would be a good additional alternative for a functional screening of ALDHs, thereby expanding the variety of the selectable enzymes.

To test further the substrate specificity of Icm and to enlarge the list of compounds applicable for the screening purposes, we have chosen *E. coli* cells transformed with pACYC-KVIA and pALDR177 plasmids. According to the activity tests, the ALDR177 clone was able to oxidize the widest spectrum of derivatives of indole-3-carboxaldehydes to the corresponding carboxylic acids. Transformants were spread on the agar plates supplemented with various indole ring containing aldehydes and incubated at 30°C for 48 hr. Colonies remained uncolored on 4-nitroindole-3-carboxaldehyde, 4-benzyloxyindole-3-carboxaldehyde, 5-benzyloxyindole-3-carboxaldehyde, 6-benzyloxyindole-3-carboxaldehyde, benzo[b]thiophene-3-carboxaldehyde, however, pigmented colonies appeared on media supplemented with 5-methylindole-3-carboxaldehyde, benzo[g]indole-3-carboxaldehyde, 1,6,7,8-tetrahydrocyclopenta[g]indole-3-carboxaldehyde, 5-bromoindole-3-carboxaldehyde (Figure A2), indicating that the corresponding carboxylic acids served as substrates for Icm. The consumption of aldehydes was confirmed by HPLC-MS. It could be concluded that those aldehydes might be applicable for a more selective screening of ALDHs.

3.4 | Screening of amidohydroxylases

Encouraged with the successful screening of ALDHs, we tested whether *icm* gene-based approach could be extended for the

functional screening of other enzymes. *N*-(3-hydroxypropyl)-indole-3-carboxamide was synthesized and used as a substrate for amidohydroxylases. One positive clone forming a blue colony was identified after testing two metagenomic DNA libraries (approx. 20,000 clones). The plasmid pMO13 isolated from this hit contained a DNA fragment encoding a 489 aa long protein, which was 90% identical to hypothetical amidase (WP_010677135) and shared 41% identity to indoleacetamide hydrolase (WP_011083078). Subsequently, MO13 amidase was cloned into pET-21a(+) vector, heterologously expressed in *E. coli* BL21(DE3), and purified as the C-His₆-tagged recombinant protein. The analysis of the substrate specificity of MO13 amidohydroxylase showed that in addition to *N*-(3-hydroxypropyl)-indole-3-carboxamide, the enzyme could hydrolyze indole-5-carboxamide, nicotinamide, hippuric acid, glycyl-L-leucine, and L-valyl-L-valine to corresponding carboxylic acids. MO13 was also active toward L-leucin-*p*-nitroanilide, 4-nitroacetanilide, and 4-nitrobenzanilide. Moreover, this amidase was able to regioselectively deprotect lysine in *N*_ε position when *N*_α,*N*_ε-di-*Z*-L-lysine or *N*_ε-Boc-*N*_α-*Z*-L-lysine was used as substrates.

4 | CONCLUSIONS

In this study, we have successfully identified a monooxygenase (Icm) active toward indole-3-carboxylic acid. The indigo formation due to activity of Icm allowed the development of a simple system for functional screening of enzymes from the metagenomic libraries. We showed that different enzymes, for example, ALDHs or amidohydroxylases could be identified depending on the used substrate. Moreover, the system might be easily extended for screening other activities as shown in Figure 2. The only requirement is that the product of enzymatic reaction would be indole-3-carboxylic acid (with or without substituents in the indole ring), which could

be a substrate for an auxiliary enzyme lcm. It should be noted that lcm was active not only in *E. coli* but also in *R. erythropolis* SQ1 cells that could open additional possibilities to use the different bacterial hosts for the functional screening.

ACKNOWLEDGEMENT

We are grateful to Kristė Šalkauskienė for technical assistance.

CONFLICT OF INTEREST

The authors declare no conflict of interests.

AUTHORS CONTRIBUTION

VČ, MS, JV, and RoM designed the experiments. VČ, MS, JV, RG, RiM, IS, MS, JJ, and DT performed the experiments. All authors analyzed the data. VČ, MS, and JV wrote the manuscript. All authors read the final manuscript.

ETHICS STATEMENT

None required.

DATA ACCESSIBILITY

All DNA sequences are submitted to GenBank. The sequences obtained in the present study were deposited to the GenBank database under the accession numbers MG770119–MG770138, MG786188, MG786189, MG775032, MK284926 and MH476458. The full list is given in Table A2. All data generated or analyzed during this study are included in this published article.

ORCID

Vida Časaitė  <https://orcid.org/0000-0001-5334-1165>

Rolandas Meškys  <https://orcid.org/0000-0003-3125-9827>

REFERENCES

- Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A., & Handelsman, J. (2009). Functional metagenomics reveals diverse B-lactamases in a remote Alaskan soil. *ISME Journal*, 3(2), 243–251. <https://doi.org/10.1038/ismej.2008.86>
- Bianchi, P., Varela, R. F., Bianchi, D. A., Kempainen, M., Iribarren, A. M., & Lewkowicz, E. (2017). Selection of microbial biocatalysts for the reduction of cyclic and heterocyclic ketones. *Process Biochemistry*, 58(April), 137–144. <https://doi.org/10.1016/j.procbio.2017.04.039>
- Celik, A., Speight, R. E., & Turner, N. J. (2005). Identification of broad specificity P450CAM variants by primary screening against indole as substrate. *Chemical Communications (Cambridge, England)*, 29, 3652–3654. <https://doi.org/10.1039/b506156c>
- Chen, R., Li, C., Pei, X., Wang, Q., Yin, X., & Xie, T. (2014). Isolation an aldehyde dehydrogenase gene from metagenomics based on semi-nest touch-down PCR. *Indian Journal of Microbiology*, 54(1), 74–79. <https://doi.org/10.1007/s12088-013-0405-0>
- 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. [https://doi.org/10.1016/S0006-291X\(03\)01087-8](https://doi.org/10.1016/S0006-291X(03)01087-8)
- Duong-Ly, K. C., & Gabelli, S. B. (2015). Affinity purification of a recombinant protein expressed as a fusion with the maltose-binding protein (MBP) tag. *Methods in Enzymology*, 559, 17–26. <https://doi.org/10.1016/j.jcviro.2015.09.001>
- 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. *Journal of Bacteriology*, 177(23), 6983–6988. <https://doi.org/10.1128/jb.177.23.6983-6988.1995>
- 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. <https://doi.org/10.1126/science.6353574>
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4), 783–791. <https://doi.org/10.2307/2408678>
- Furuya, T., Takahashi, S., Ishii, Y., Kino, K., & Kirimura, K. (2004). Cloning of a gene encoding flavin reductase coupling with dibenzothiophene monooxygenase through coexpression screening using indigo production as selective indication. *Biochemical and Biophysical Research Communications*, 313(3), 570–575. <https://doi.org/10.1016/j.bbrc.2003.11.157>
- Hicks, K. A., Yuen, M. E., Zhen, W. F., Gerwig, T. J., Story, R. W., Kopp, M. C., & Snider, M. J. (2016). Structural and biochemical characterization of 6-hydroxynicotinic acid 3-monooxygenase, a novel decarboxylative hydroxylase involved in aerobic nicotinate degradation. *Biochemistry*, 55(24), 3432–3446. <https://doi.org/10.1021/acs.biochem.6b00105>
- Huijbers, M. M. E., Montersino, S., Westphal, A. H., Tischler, D., & van Berkel, W. J. H. (2014). Flavin dependent monooxygenases. *Archives of Biochemistry and Biophysics*, 544, 2–17. <https://doi.org/10.1016/j.abb.2013.12.005>
- Kennedy, J., O'Leary, N. D., Kiran, G. S., Morrissey, J. P., O'Gara, F., Selvin, J., & Dobson, A. D. W. (2011). Functional metagenomic strategies for the discovery of novel enzymes and biosurfactants with biotechnological applications from marine ecosystems. *Journal of Applied Microbiology*, 111(4), 787–799. <https://doi.org/10.1111/j.1365-2672.2011.05106.x>
- Kotchoni, S. O., Kuhns, C., Ditzer, A., Kirch, H. H., & Bartels, D. (2006). Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to abiotic stress and protects plants against lipid peroxidation and oxidative stress. *Plant, Cell and Environment*, 29(6), 1033–1048. <https://doi.org/10.1111/j.1365-3040.2005.01458.x>
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Lyu, Y., LaPointe, G., Zhong, L., Lu, J., Zhang, C., & Lu, Z. (2017). Heterologous expression of aldehyde dehydrogenase in *Lactococcus lactis* for acetaldehyde detoxification at low pH. *Applied Biochemistry and Biotechnology*, 184(2), 570–581. <https://doi.org/10.1007/s12010-017-2573-6>
- Maruthamuthu, M., Jiménez, D. J., Stevens, P., & van Elsas, J. D. (2016). A multi-substrate approach for functional metagenomics-based screening for (hemi)cellulases in two wheat straw-degrading microbial consortia unveils novel thermoalkaliphilic enzymes. *BMC Genomics*, 17(1), 1–16. <https://doi.org/10.1186/s12864-016-2404-0>

- Mavrodi, D. V., Bonsall, R. F., Delaney, S. M., Soule, M. J., Phillips, G., & Thomashow, L. S. (2001). Functional analysis of genes for biosynthesis of pyocyanin and phenazine-1-carboxamide from *Pseudomonas aeruginosa* PAO1. *Journal of Bacteriology*, 183(21), 6454–6465. <https://doi.org/10.1128/JB.183.21.6454>
- McClay, K., Boss, C., Keresztes, I., & Steffan, R. J. (2005). Mutations of toluene-4-monooxygenase that alter regioselectivity of indole oxidation and lead to production of novel indigoid pigments. *Applied and Environmental Microbiology*, 71(9), 5476–5483. <https://doi.org/10.1128/AEM.71.9.5476-5483.2005>
- Nagayama, H., Sugawara, T., Endo, R., Ono, A., Kato, H., Ohtsubo, Y., ... Tsuda, M. (2015). Isolation of oxygenase genes for indigo-forming activity from an artificially polluted soil metagenome by functional screening using *Pseudomonas putida* strains as hosts. *Applied Microbiology and Biotechnology*, 99(10), 4453–4470. <https://doi.org/10.1007/s00253-014-6322-2>
- Nakashima, N., & Tamura, T. (2004a). A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35°C. *Biotechnology and Bioengineering*, 86(2), 136–148. <https://doi.org/10.1002/bit.20024>
- Nakashima, N., & Tamura, T. (2004b). Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to expression. *Applied and Environmental Microbiology*, 70(9), 5557–5568. <https://doi.org/10.1128/AEM.70.9.5557>
- Nei, M., & Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. New York, NY: Oxford University Press.
- O'Connor, K. E., Dobson, A. D. W., & Hartmans, S. (1997). Indigo formation by microorganisms expressing styrene monooxygenase activity. *Applied and Environmental Microbiology*, 63(11), 4287–4291.
- Oyobiki, R., Kato, T., Katayama, M., Sugitani, A., Watanabe, T., Einaga, Y., ... Doi, N. (2014). Toward high-throughput screening of NAD(P)-dependent oxidoreductases using boron-doped diamond microelectrodes and microfluidic devices. *Analytical Chemistry*, 86(19), 9570–9575. <https://doi.org/10.1021/ac501907x>
- Pavel, A. B., & Vasile, C. I. (2012). PyElph - a software tool for gel images analysis and phylogenetics. *BMC Bioinformatics*, 13(1), <https://doi.org/10.1186/1471-2105-13-9>
- Qian, S., & Dabbs, E. (1993). Nocardioform arsenic resistance plasmid characterization and improved *Rhodococcus* cloning vectors. *Plasmid*, 29(1), 74–79. <https://doi.org/10.1006/plas.1993.1010>
- Reisinger, C., van Assema, F., Schürmann, M., Hussain, Z., Remler, P., & Schwab, H. (2006). A versatile colony assay based on NADH fluorescence. *Journal of Molecular Catalysis B: Enzymatic*, 39(1–4), 149–155. <https://doi.org/10.1016/j.molcatb.2006.01.014>
- Reyes-Duarte, D., Ferrer, M., & García-Arellano, H. (2012). Lipases and phospholipases. In G. Sandoval (Ed.), *Methods in molecular biology*, 861, 101–113. <https://doi.org/10.1007/978-1-61779-600-5>
- Rüther, U. (1980). Construction and properties of a new cloning vehicle, allowing direct screening for recombinant plasmids. *MGG Molecular & General Genetics*, 178(2), 475–477. <https://doi.org/10.1007/BF00270503>
- Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), 406–425. <https://doi.org/citeulike-articleid:93683>
- Sambrook, J., Fritsch, E., & Maniatis, T. (1989). *Molecular cloning: A laboratory manual*. Long Island, NY: Cold Spring Harbor.
- Seok, J. Y., Yang, J., Choi, S. J., Lim, H. G., Choi, U. J., Kim, K.-J., ... Jung, G. Y. (2018). Directed evolution of the 3-hydroxypropionic acid production pathway by engineering aldehyde dehydrogenase using a synthetic selection device. *Metabolic Engineering*, 47(March), 113–120. <https://doi.org/10.1016/j.ymben.2018.03.009>
- Shang, M., Chan, V. J., Wong, D. W. S., & Liao, H. (2018). A novel method for rapid and sensitive metagenomic activity screening. *MethodsX*, 5, 669–675. <https://doi.org/10.1016/J.MEX.2018.06.011>
- Shi, S., Ma, F., Sun, T., Li, A., Zhou, J., & Qu, Y. (2013). Biotransformation of indole to indigo by the whole cells of phenol hydroxylase engineered strain in biphasic systems. *Applied Biochemistry and Biotechnology*, 169(4), 1088–1097. <https://doi.org/10.1007/s12010-012-0069-y>
- Singh, S., Brocker, C., Koppaka, V., Ying, C., Jackson, B., Thompson, D. C., & Vasiliou, V. (2014). Aldehyde dehydrogenases in cellular responses to oxidative/ electrophilic stress. *Free Radical Biology and Medicine*, 1(303), 89–101. <https://doi.org/10.1016/j.freeradbiomed.2012.11.010>
- Stanislaukiene, R., Gasparavičiute, R., Vaitekūnas, J., Meskiene, R., Rutkiene, R., Časaite, V., & Meskys, R. (2012). Construction of *Escherichia coli*-*Arthrobacter Rhodococcus* shuttle vectors based on a cryptic plasmid from *Arthrobacter rhombi* and investigation of their application for functional screening. *FEMS Microbiology Letters*, 327(1), 78–86. <https://doi.org/10.1111/j.1574-6968.2011.02462.x>
- Stankevičiūtė, J., Vaitekūnas, J., Petkevičius, V., Gasparavičiūtė, R., Tauraitė, D., & Meškys, R. (2016). Oxyfunctionalization of pyridine derivatives using whole cells of *Burkholderia* sp. MAK1. *Scientific Reports*, 6(1), 39129. <https://doi.org/10.1038/srep39129>
- Suenaga, H., Ohnuki, T., & Miyazaki, K. (2007). Functional screening of a metagenomic library for genes involved in microbial degradation of aromatic compounds. *Environmental Microbiology*, 9(9), 2289–2297. <https://doi.org/10.1111/j.1462-2920.2007.01342.x>
- Tsuji, H., Ogawa, T., Bando, N., & Sasaoka, K. (1986). Purification and properties of 4-aminobenzoate hydroxylase, a new monooxygenase from *Agaricus bisporus*. *Journal of Biological Chemistry*, 261(28), 13203–13209.
- Uemura, T., Kita, A., Watanabe, Y., Adachi, M., Kuroki, R., & Morimoto, Y. (2016). The catalytic mechanism of decarboxylative hydroxylation of salicylate hydroxylase revealed by crystal structure analysis at 2.5 Å resolution. *Biochemical and Biophysical Research Communications*, 469(2), 158–163. <https://doi.org/10.1016/j.bbrc.2015.11.087>
- Van Berkel, W. J. H., Eppink, M. H. M., Middelhoven, W. J., Vervort, J., & Rietjens, M. C. M. (1994). Catabolism of 4-hydroxybenzoate in *Candida parapsilosis* proceeds through initial oxidative decarboxylation by a FAD-dependent 4-hydroxybenzoate 1-hydroxylase. *FEMS Microbiology Letters*, 121, 207–215. [https://doi.org/10.1016/0378-1097\(94\)90128-7](https://doi.org/10.1016/0378-1097(94)90128-7)
- Varaljay, V. A., Satagopan, S., North, J. A., Witte, B., Dourado, M. N., Anantharaman, K., ... Tabita, F. R. (2016). Functional metagenomic selection of ribulose 1,5-bisphosphate carboxylase/oxygenase from uncultivated bacteria. *Environmental Microbiology*, 18(4), 1187–1199. <https://doi.org/10.1111/1462-2920.13138>
- Wexler, M., Bond, P. L., Richardson, D. J., & Johnston, A. W. B. (2005). A wide host-range metagenomic library from a waste water treatment plant yields a novel alcohol/aldehyde dehydrogenase. *Environmental Microbiology*, 7(12), 1917–1926. <https://doi.org/10.1111/j.1462-2920.2005.00854.x>
- Willett, A., Joint, I., Gilbert, J. A., Trimble, W., & Mühling, M. (2012). Isolation and initial characterization of a novel type of Baeyer-Villiger monooxygenase activity from a marine microorganism. *Microbial Biotechnology*, 5(4), 549–559. <https://doi.org/10.1111/j.1751-7915.2012.00337.x>
- Ye, X., Peng, Y., Niu, Z., Luo, X., & Zhang, L. (2018). Novel approach for the rapid screening of banned aromatic amines in dyed textiles using a chromogenic method. *Analytical and Bioanalytical Chemistry*, 410(11), 2701–2710. <https://doi.org/10.1007/s00216-018-0941-x>

How to cite this article: Časaite V, Sadauskas M, Vaitekūnas J, et al. Engineering of a chromogenic enzyme screening system based on an auxiliary indole-3-carboxylic acid monooxygenase. *MicrobiologyOpen*. 2019;e795. <https://doi.org/10.1002/mbo3.795>

APPENDIX 1

TABLE A1 Source of chemicals

Aldrich, Buchs, Switzerland
5-Methylfurfural, 1-naphthaldehyde, 3-cyclohexene-1-carboxaldehyde, 3-hydroxybenzaldehyde, <i>trans</i> -cinnamaldehyde, vanillin, phenylacetaldehyde, furfural, 3,4-dihydroxybenzaldehyde, benzaldehyde, salicylaldehyde, pyrrole-2-carboxaldehyde, 4-(dimethylamino)benzaldehyde, cyclohexanecarboxaldehyde, 1,2-dimethyl-1H-I3C, 5-methylindole-3-carboxaldehyde, 4-nitroindole-3-carboxaldehyde, 1H-benzo[<i>g</i>]indole-3-carboxaldehyde, 2-phenylindole-3-carboxaldehyde, 6-isopropylindole-3-carboxaldehyde, nicotinic acid, 2-picoline acid, 4-picoline acid, indole-6-carboxylic acid, 5-nitroindole-3-carboxylic acid, salicylic acid, 2-phenylpropionaldehyde, 1,6,7,8-tetrahydrocyclopenta[<i>g</i>]indole-3-carboxaldehyde, 5-bromindole-3-carboxaldehyde, 5-benzoyloxyindole-3-carboxaldehyde, 1-methylindole-3-carboxaldehyde, 4-nitroacetanilide, 4-nitrobenzanilide, 1H-indole-5-carboxamide, hippuric acid, <i>N,N,N',N'</i> -tetramethyl- <i>O</i> -(1H-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU), dimethylformamide NADH, FAD
Combi Blocks, SanDiego, USA
4-Benzoyloxyindole-3-carboxaldehyde, 6-benzoyloxyindole-3-carboxaldehyde, benzo[<i>b</i>]thiophene-3-carboxaldehyde, 5-hydroxy pyrazine-2-carboxylic acid, indoline-2-carboxylic acid, indole-4-carboxylic acid, indole-5-carboxylic acid, indole-7-carboxylic acid
Fluka, Steinheim, Switzerland
Indole-2-carboxylic acid, nicotinamide, L-leucin- <i>p</i> -nitroanilide.
Sigma, St. Louis, USA
7-Methylindole-3-carboxylic acid, indole 3-carboxylic acid (I3CA), indole-3-carboxaldehyde (I3C)
Reanal, Budapest, Hungary
Glycyl-L-leucine, L-valyl-L-valine
Merck, Darmstadt, Germany
3-Amino-1-propanol, z-lys(z)-OH, boc-lys(z)-OH, ethyl acetate, triethylamine
Thermo Fisher Scientific Vilnius, Lithuania
5-Bromo-4-chloro-indolyl- β -D-galactopyranoside (X-gal), isopropyl- β -D-thiogalactopyranoside (IPTG)

TABLE A2 Clones from soil metagenomic libraries

Clone name	GenBank accession number
JU61	MG770119
pER2AH2	MG770120
pALDR177	MG770121
pER2AH	MG770122
VMIX	MG770123
pALDSV3	MG770124
pALDMO9	MG770126
pALDBSal	MG770127
pALDBS21	MG770128
pALDMO11	MG770129
pEGA1	MG770130
pEMMO	MG770131
pALD458	MG770132
URAGR	MG770133
pRG1	MK284926
pRG2	MG770134
pALDJU6	MG770135
DON4	MG770136
pALDGA1	MG770137
pALD442	MG770138
pMILC	MG786188
pNVS	MG786189
pKVIABam8	MG770125
KVIA (16S RNA gene)	MG775032
MO13	MH476458

TABLE A3 Solubility of Icm at different induction conditions in *Escherichia coli*. Condition A—induction for 4 hr with 0.5 mM IPTG at 30°C, condition B—induction for 4 hr with 0.05 mM IPTG at 30°C, and condition C—overnight induction with 0.05 mM IPTG at 16°C. For Strep-Tag MBP-His₆-KVIA, anhydrotetracycline (200 µg/L) was used instead of IPTG. ND: not detected

	Condition A			Condition B			Condition C		
	Total amount of Icm (mg/L)	Amount of soluble Icm (mg/L)	% soluble	Total amount of Icm (mg/L)	Amount of soluble Icm (mg/L)	% soluble	Total amount of Icm (mg/L)	Amount of soluble Icm (mg/L)	% soluble
Icm	10.25	ND	ND	5.9	ND	ND	11.7	ND	ND
His ₆ -Icm	14.9	ND	ND	11.1	ND	ND	19.1	ND	ND
His ₆ -MBP-His ₆ -Icm	19.6	0.4	1.9	18.3	6.8	37.3	44.4	9	20.2
Strep-Tag MBP-His ₆ -Icm	39.2	3.3	8.4	29.1	7.3	24.9	35	21.4	61

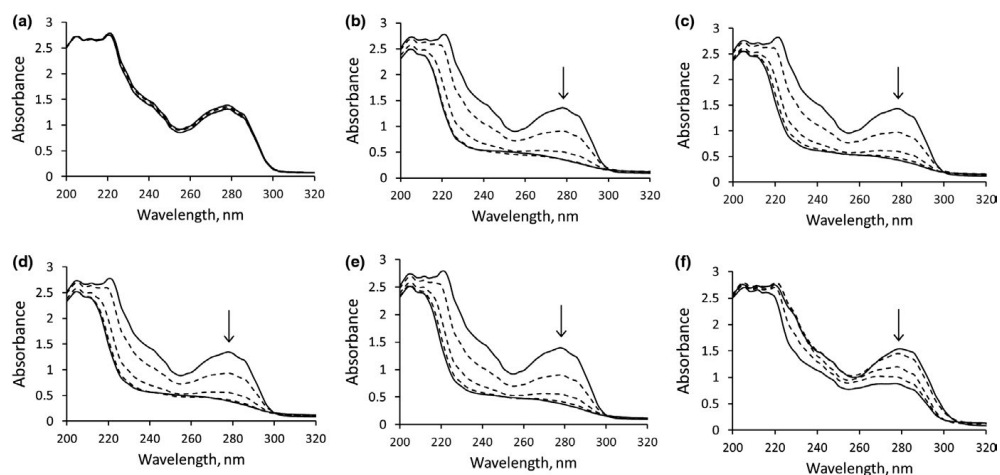


FIGURE A1 Whole-cell consumption of I3CA by recombinant cells with Icm. (a) *Escherichia coli* BL21 (DE3) cells (negative control), (b) *E. coli* BL21 (DE3) cells with Icm, (c) cells with His₆-Icm, (d) cells with His₆-MBP-His₆-Icm, (e) cells with Strep-Tag MBP-His₆-Icm, (f) *Rhodococcus erythropolis* SQ1 cells with Icm. Consumption was monitored hourly during a period of 5 hr

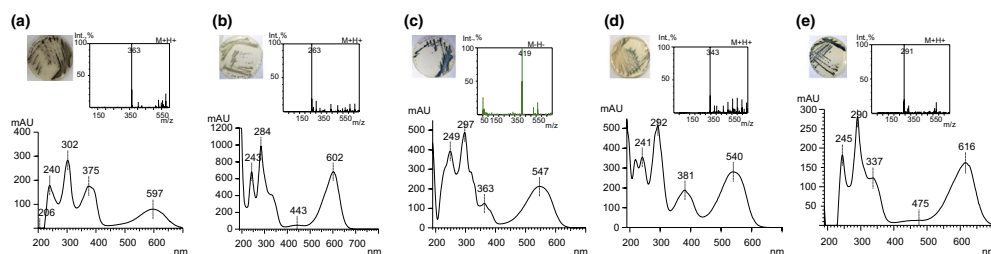


FIGURE A2 Formation of blue pigment in the presence of Icm and different indole carboxaldehydes. *Escherichia coli* DH5α (pACYC-KVIA/pALDR177) cells on the plates, UV and Mass spectra of formed compounds: (a) 1*H*-benzo[*g*]indole-3-carboxaldehyde, (b) I3C, (c) 5-bromindole-3-carboxaldehyde, (d) 1,6,7,8-tetrahydrocyclopenta[*g*]indole-3-carboxaldehyde, (e) 5-methylindole-3-carboxaldehyde

Publication II

Sadauskas, M., Statkevičiūtė, R., Vaitekūnas, J., Petkevičius, V., Časaitė, V., Gasparavičiūtė, R., Meškys, R. Enzymatic synthesis of novel water-soluble indigoid compounds. *Dyes and Pigments*, 2020, 173, 107882
DOI: 10.1016/j.dyepig.2019.107882

This material is licensed for reuse by Elsevier Ltd. All rights reserved.



Enzymatic synthesis of novel water-soluble indigoid compounds

Mikas Sadauskas*, Roberta Statkevičiūtė, Justas Vaitekūnas, Vytautas Petkevičius, Vida Časaite, Renata Gasparavičiūtė, Rolandas Meškys

Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Life Sciences Center, Vilnius University, Sauletekio al. 7, Vilnius, LT-10257, Lithuania

ARTICLE INFO

Keywords:

Indigo
Indirubin
Bacterial flavin-dependent monooxygenases

ABSTRACT

This work was aimed at expanding the diversity of indigoid compounds with novel characteristics by employing biological catalysts. A total of 16 novel indigoid compounds were synthesized using indole derivatives as substrates containing either aminomethyl-, hydroxymethyl-, carboxaldehyde or carboxyl groups at positions 4, 5, 6 or 7 of indole ring. Two different monooxygenase systems – a flavin dependent monooxygenase Hind8 and the mutant G109Q of multicomponent soluble diiron monooxygenase PML were employed to achieve the conversion of all those substrates. Characterization of purified indigo dicarboxylic acids revealed that the produced indigoids were soluble in water (solubility increased more than 1000-fold compared to indigo), also in methanol and DMSO. A bioconversion of indole-7-carboxylic acid resulted in the mixture of indigo-7,7'-dicarboxylic acid and indirubin-7,7'-dicarboxylic acid. Both of those products were separated from each another and purified thus providing the basis for the purification of water-soluble indigoids of similar structure. These findings show that (1) biological catalysts can be an easy approach for synthesizing novel indigoids with important characteristics and (2) introduction of carboxyl group can increase the water-solubility of indigo and indirubin significantly, thus providing a new tool in search for bioactive indigoids.

1. Introduction

Indigo is one of the oldest examples of green biotechnology, when leaf extracts of *Indigofera* spp. were used for cotton dyeing more than 6000 years ago [1]. In the 20th century, the rate of synthetic indigo production increased and reached approx. 17000 tons a year in 2012 [2]. These amounts are used for cotton and jeans dyeing mainly. Composed of two aromatic indolinone ring systems, indigo is highly insoluble in water; however, two carbonyl groups are also present, which can be reduced to form leuco-indigo, a water-soluble form of indigo. These features are exploited during a vat dyeing.

Another ancient indigoid compound is indirubin, a structural isomer of indigo. Indirubin is not being used in textile industry due to low stability, albeit having a clear purple color. Instead, indirubin has been known to be a constituent of the Chinese herbal medicine qing-dai, also known as Indigo naturalis. This medicine has long been used as an anti-inflammatory drug [3]. Indirubin can act as an inhibitor of cyclin-dependent kinase 1 [4] and glycogen synthase kinase [5], inhibit the proliferation of cancer cells and also the proliferation of tumor-derived endothelial cells [6,7]. In addition, isoindigo, another structural isomer

of indigo, has become an excellent tool for the production of organic electronics, particularly organic photovoltaics and organic field effect transistors [8].

Substituent chemical groups can be introduced into the indigo backbone and the resulting compounds are regarded as indigoids. Those compounds receive increasing attention due to exceptional characteristics. Owing to their reversible two-electron reduction and oxidation, indigoids can act as both electron donors and acceptors, thus providing the basis for production of indigoids-based materials for green organic electronics [9]. Due to low toxicity and chemical stability in aerated conditions, halogenated indigoids, such as 5-bromoindigo, 6,6'-dichloroindigo and Tyrian purple (6,6'-dibromoindigo), regarded as one of the most precious ancient compounds, are potential compounds for semiconductor synthesis [10]. Indigo *N,N'*-bis(arylimine)s (Nindigo), modified variants of indigo containing two β -diketiminatetype motifs, were shown to form coordination complexes with different redox active metals [11–13] giving these indigoids novel absorbance and redox properties.

Different methods exist for the synthesis of indigoids. Chemical synthesis has been the preferred method until recently, when enzymatic

* Corresponding author.

E-mail addresses: mikas.sadauskas@bchi.vu.lt (M. Sadauskas), statkeviciuteroberta@gmail.com (R. Statkevičiūtė), justas.vaitekunas@bchi.vu.lt (J. Vaitekūnas), vytautas.petkevicius@bchi.vu.lt (V. Petkevičius), vida.casaite@bchi.vu.lt (V. Časaite), renata.gasparaviciute@bchi.vu.lt (R. Gasparavičiūtė), rolandas.meskys@bchi.vu.lt (R. Meškys).

<https://doi.org/10.1016/j.dyepig.2019.107882>

Received 31 July 2019; Received in revised form 12 September 2019; Accepted 12 September 2019

Available online 13 September 2019

0143-7208/ © 2019 Elsevier Ltd. All rights reserved.

synthesis emerged as an environmental-friendly alternative method offering additional benefits. Monooxygenase or multicomponent dioxygenase enzymes have been identified as the ones capable of oxidation of indole derivatives to corresponding indigoids. Some oxygenases are highly specific to indole as a natural substrate, while others can accept substituted indoles due to enzyme promiscuity. Since the first description of indigo bioproduction in *E. coli* using recombinant naphthalene dioxygenase genes [14], a plethora of different proteins were used to achieve the synthesis of a wide range of indigoid compounds. For example, different cytochrome P450 monooxygenases and their mutants were used for production of indigoid derivatives: dichloroindigo, dibromoindigo, dihydroxyindigo, dimethylindigo, dinitroindigo, dicyanoindigo, diaminoindigo [15], 5,5'-dibenzoyloxyindirubin [16] and indirubin [17]. Also, naphthalene dioxygenase and toluene monooxygenase have been demonstrated to convert substituted indoles to corresponding indigoid pigments [18,19]. In addition, the oxidation of indole to indigoid pigments itself has been used as a screening platform leading to identification of biodegradation genes of 2-hydroxypyridine and selection of mutant enzymes with novel activities [20,21]. Recently, flavin-dependent monooxygenases have been identified as responsible for the initial step of indole biodegradation [22,23] via indole epoxidation [24] leading to formation of indoxyl and indigo. As such, indole appears to be a natural substrate of certain group E flavin-dependent monooxygenases [24]. This feature has been largely exploited for high-yield synthesis of indigo reaching the yield of up to 1 mg L^{-1} [25] as well as for the synthesis of modified indigoids [24].

Enzymatic synthesis of indigoids offers several advantages over chemical synthesis. The latter employs environmental-damaging conditions, such as aniline, cyanide, formaldehyde, strong bases and heating to 300°C [26]. Meanwhile, enzymatic synthesis only requires conditions suitable for cultivation of conventional microbial cellular factories, typically *E. coli*. Moreover, genetically modified microorganisms were engineered for *de novo* indigo biosynthesis from glucose or L-tryptophan [15,27]. On the other hand, chemical synthesis allows the generation of pure asymmetrical indigoids, while enzymatic synthesis of asymmetrical indigoids often results in a mixture of products. Such indigoids often possess fairly different biological and physical properties compared to symmetrical form [28,29].

Indigo dye is well-known for low solubility in aqueous solutions and high solubility in organic solvents. Several indigoids were created to facilitate the hydrophilicity. The most notable is indigo carmine (indigo-5,5'-disulfonic acid, indigotine) which is soluble in water (10 g L^{-1}) and has wide industrial applications. Due to solubility in water, indigo carmine has been used as pH, ozone and redox indicator [30]. When sodium ions in indigo carmine have been replaced with large organic cations, the solubility of such indigo derivative in water reaches 1.6 M while maintaining the ability of both reversible oxidation and reduction allowing the creation of redox flow cells at neutral pH [31]. Water-soluble indirubin derivatives (mostly indirubin-3'-oximes) have also been synthesized and possessed novel biological activities [32,33]. This implies that insolubility, while providing the reason for indole-based vat dyeing or development of organic electronics, limits the biological activity of indigoids. Hence, water-soluble indigo-based compounds offer possibilities for different and novel applications.

The majority of enzymatically produced indigoid compounds reported to date were halogenated indoles, methoxy-, cyano-, nitro- and similar derivatives [15,16,24]. This research aimed at increasing the variety of indigoids with altered functions by applying enzymatic synthesis of novel pigments. To achieve this goal, a metagenomic screening was performed using the formation of blue colonies on aminomethyl-, hydroxymethyl-, carboxaldehyde or carboxylic derivatives of indole as a screening platform.

2. Results

2.1. Selection and characterization of indole derivatives-oxidizing clones

Two approaches were used to identify oxygenases active toward hydroxymethyl-, aminomethyl-, carboxaldehyde- and carboxy-group containing indoles. First, metagenomic libraries containing approximately 0.3–0.5 Gb of total cloned DNA per library with an average insert size of 3–12 kb were screened for indigo-forming colonies in the presence of indole. Then the positive hits (80) were further tested for ability to oxidize indoles harboring the aforementioned side chains at positions 4, 5, 6 and 7. Among tested metagenomic clones a hit named Hind8 was able to oxidize a wide variety of indole derivatives with hydroxymethyl-, aminomethyl- and carboxaldehyde. The DNA sequence analysis of this clone (sequence accession number MN124688) revealed a single open reading frame (1488 bp) which had the highest (87%) nucleotide sequence identity to hypothetical FAD-containing EthA-type monooxygenase gene from *Aminobacter* sp MSH1 (CP026265). A phylogenetic analysis of amino acid sequence of the oxygenase encoded by the hit Hind8 showed that the protein formed a branch with flavin-dependent monooxygenases of group B (Fig. S1).

For the second approach, a soluble diiron monooxygenase PML and its mutants [21] were tested for conversion of the above-mentioned indoles to corresponding indigoids. It was found, that in contrast to the wild-type enzyme, the G109Q mutant showed an expanded range of substrates and was active toward carboxyindoles. This mutant was chosen for further analysis.

To compare the substrate specificity of these two screened proteins, the oxygenase gene from the hit Hind8 was amplified by PCR and cloned into a pLATE31 vector for gene expression. The *E. coli* BL21(DE3) cells transformed with pLATE31-Hind8 plasmid produced a soluble 55-kDa recombinant Hind8 protein according to the SDS-PAGE analysis (Fig. S2).

Next, a new expression host was developed by deleting the tryptophanase-encoding *tnaA* gene in *E. coli* strain HMS174(DE3) in order to avoid background indigo production from endogenous tryptophan. The cells transformed with pLATE31-Hind8 plasmid were tested on agar plates supplemented with IPTG and different indole derivatives, and the pigment production was compared with *E. coli* HMS174(DE3) Δ *tnaA* harboring the PML G109Q oxygenase (Fig. 1).

Based on this analysis, it was concluded that both enzymes could accept a wide range of indoles but some differences in activity were also observed for both oxygenases.

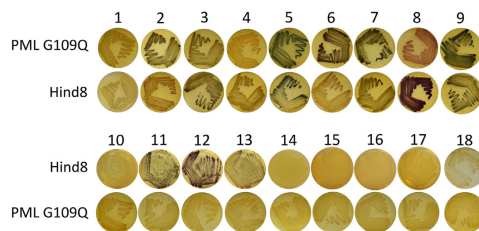


Fig. 1. Pigment formation by *E. coli* HMS174(DE3) Δ *tnaA* producing Hind8 and PML G109Q oxygenases on LB agar plates supplemented with different indole derivatives: 1 – 4-hydroxymethylindole, 2 – 5-hydroxymethylindole, 3 – 6-hydroxymethylindole, 4 – 7-hydroxymethylindole, 5 – indole-4-carboxaldehyde, 6 – indole-5-carboxaldehyde, 7 – indole-6-carboxaldehyde, 8 – indole-7-carboxaldehyde, 9 – indole, 10 – 4-aminomethylindole, 11 – 5-aminomethylindole, 12 – 6-aminomethylindole, 13 – 7-aminomethylindole, 14 – indole-4-carboxylic acid, 15 – indole-5-carboxylic acid, 16 – indole-6-carboxylic acid, 17 – indole-7-carboxylic acid, 18 – no substrate.

2.2. Whole cells bioconversion of indole derivatives

To determine the substrate specificity of Hind8 and PML G109Q toward indole derivatives in more detail, *E. coli* HMS174(DE3) Δ maA cells with recombinant pLATE31-Hind8 or pET28-PML-G109Q plasmids were used for whole-cell bioconversion. In total, 16 modified indoles were tested as possible substrates. Those included indoles containing either an aminomethyl-, hydroxymethyl-, carboxaldehyde or carboxyl group at positions 4, 5, 6 or 7 of indole ring. The activity was determined following the substrate consumption, extracting the colored bioconversion products and analysis of the formed products by RP-HPLC/MS. It is important to note that these techniques did not allow the separation between indigo, indirubin and isoindigo isoforms, therefore all indigoids described below, except for conversion products of indole carboxylic acids, are regarded as indigoid compounds.

2.2.1. Bioconversion of aminomethylindoles

Out of four different aminomethylindoles that were used, the PML G109Q oxidized only 4-aminomethylindole to corresponding indigoid compound. Meanwhile, Hind8 was active toward other three aminomethylindoles (5-aminomethylindole, 6-aminomethylindole and 7-aminomethylindole) forming the corresponding indigoids, but did not use 4-aminomethylindole as substrate. The observed molecular mass of the major peaks of di(aminomethyl)-indigoids (320 Da, Figs. S3–S5) corresponded to the theoretical mass (320.3 Da). Minor peaks visible in the RP-HPLC chromatograms of di(aminomethyl)-indigoids could be attributed to other side products, but the exact identification was prevented by weak ionization of these products. The above-mentioned differences in substrate specificity between the two monooxygenases were also visible on agar plates as Hind8 harboring cells produced blue colonies on 5-aminomethylindole, 6-aminomethylindole and 7-aminomethylindole, while colonies with PML G109Q turned blue only on 4-aminomethylindole (Fig. 1). However, a whole-cell bioconversion of 4-aminomethylindole by PML G109Q proceeded very slowly and inefficiently and did not yield a sufficient amount of product for MS and TLC analyses.

2.2.2. Bioconversion of hydroxymethylindoles

No differences in substrate specificity were observed between Hind8 and PML G109Q when 4-, 5-, 6- and 7-hydroxymethylindoles were used as substrates. All four compounds were successfully converted to the corresponding indigoid pigments both on agar plates and during whole-cell bioconversion experiments. Molecular mass of the formed pigments (322 Da, Figs. S6–S9) also corresponded to calculated mass (322.3 Da). After extraction, pure di(hydroxymethyl)indigoid products were obtained, except for 5-hydroxymethylindole, where an unidentified pigment composed less than 10% of the total extracted pigment amount (Fig. S3).

2.2.3. Bioconversion of indole carboxaldehydes

During bioconversion of indole carboxaldehydes, it was observed that *E. coli* cells could reduce aldehyde groups forming the corresponding hydroxymethylindoles at a significant rate. Therefore, mixtures of indigoid compounds were obtained during bioconversion of indole carboxaldehydes, consisting of di(hydroxymethyl)indigoid (molecular mass 322 Da), indigoid dicarboxaldehydes (mass 318 Da), and an asymmetric hydroxymethylindigoid-carboxaldehyde (molecular mass 320 Da) (Figs. S10–S13). In spite of the fact that a mixture of indigoids was obtained, the target compounds (indole dicarboxaldehydes) were successfully identified as comprising approx. 1/3 of the total indigoids in the reaction mixture. This suggested that both Hind8 and PML G109Q were capable of converting all tested indoles to corresponding indigoid dicarboxaldehydes.

2.2.4. Bioconversion of indole carboxylic acids

No development of colored colonies was observed on agar plates

supplemented with indole carboxylic acids (indole-4-carboxylic acid, indole-5-carboxylic acid, indole-6-carboxylic acid and indole-7-carboxylic acid) neither with Hind8 nor PML G109Q containing cells. Surprisingly, *E. coli* cells with recombinant proteins (both Hind8 and PML G109Q) were able to transform indole-5-carboxylic acid, indole-6-carboxylic acid and indole-7-carboxylic acid to the colored compounds in liquid media. However, the activity toward the carboxylic acids was different: the Hind8 oxygenase converted indole-6- and indole-7-carboxylic acid at a faster rate than the PML G109Q enzyme, but the latter appeared to be more specific to indole-5-carboxylic acid (Figs. S15–S17, Table S1). Still, both catalysts were inactive toward indole-4-carboxylic acid (Fig. S14 D). It is worth noting that neither the wild-type PML nor any other PML mutant were able to oxidize any of indole carboxylic acids.

Usually, after bioconversion, the indigoid compounds remained inside the cells and could be retrieved by centrifugation, followed by extraction with organic solvents. However, this was not the case for indigo dicarboxylic acids as the colored compounds clearly remained in the supernatant after centrifugation and the cell pellet remained only slightly yellow. This suggested that indigo dicarboxylic acids could be soluble in water. To test this, indigo dicarboxylic acids were purified and their solubility in different solvents was analyzed.

2.3. Characterization of indigo dicarboxylic acids

To characterize the obtained indigo dicarboxylic acids, RP-FPLC was used for purification of the formed reaction products. The purity of indigo dicarboxylic acids was confirmed by RP-HPLC/MS and TLC (Fig. S18–S21 and S22) and was estimated to be > 95%. In order to achieve the highest yield, PML G109Q was used for oxidation of indole-5-carboxylic acid and Hind8 – for oxidation of indole-6- and indole-7-carboxylic acid.

Oxidation of indole-7-carboxylic acid yielded a mixture of two indigoids in an approx. ratio 1:3 – the first one with retention time of 7.1 min and the second one with retention time of 7.6 min (Fig. 2). Those indigoids were successfully separated from each other and purified. The indigoid with retention time of 7.1 min possessed absorbance maximum at 614 nm, molecular weight of 350 Da and a clear blue color suggesting indigo-7,7'-dicarboxylic acid. Meanwhile, another indigoid had absorbance maximum at 548 nm, which was similar to the absorbance maximum of indirubin (546 nm) [34], an identical molecular mass to the first indigoid (350 Da) and a clear purple color. The yield of 7.1 min indigoid was insufficient for NMR analysis, however, the structure of the indigoid with retention time of 7.6 min was successfully determined by ^1H NMR and was confirmed to be indirubin-7,7'-dicarboxylic acid (Fig. S24).

Both indigo-5,5'-dicarboxylic acid and indigo-6,6'-dicarboxylic acid were produced as pure products and without the notable formation of an indirubin isomer. Indigo-5,5'-dicarboxylic acid possessed identical absorbance peak to indigo-7,7'-dicarboxylic acid (614 nm). The absorbance maximum of indigo-6,6'-dicarboxylic acid in the long wave region was different (639 nm) from other indigo dicarboxylic acids. In addition, this compound gave a distinct green-blue color, probably due to a strong absorption in the range of 400–450 nm (Fig. 3). The structure of indigo-5,5'-dicarboxylic acid was also confirmed by ^1H NMR (Fig. S23).

To gain insights into the solubility patterns of indigo and indirubin dicarboxylic acids, solvents such as methanol, DMSO and Milli-Q H_2O were used in solubility experiments. All four indigo dicarboxylic acids possessed more than a 1000-fold better solubility in water (8–16 mM) compared to indigo (Table 1), and it was comparable to the solubility of indigo carmine. Indigo dicarboxylic acids were slightly less soluble in methanol, with a notable exception of indirubin-7,7'-dicarboxylic acid. However, molar absorptivity coefficients of the analyzed compounds in water were rather low (Table 1).

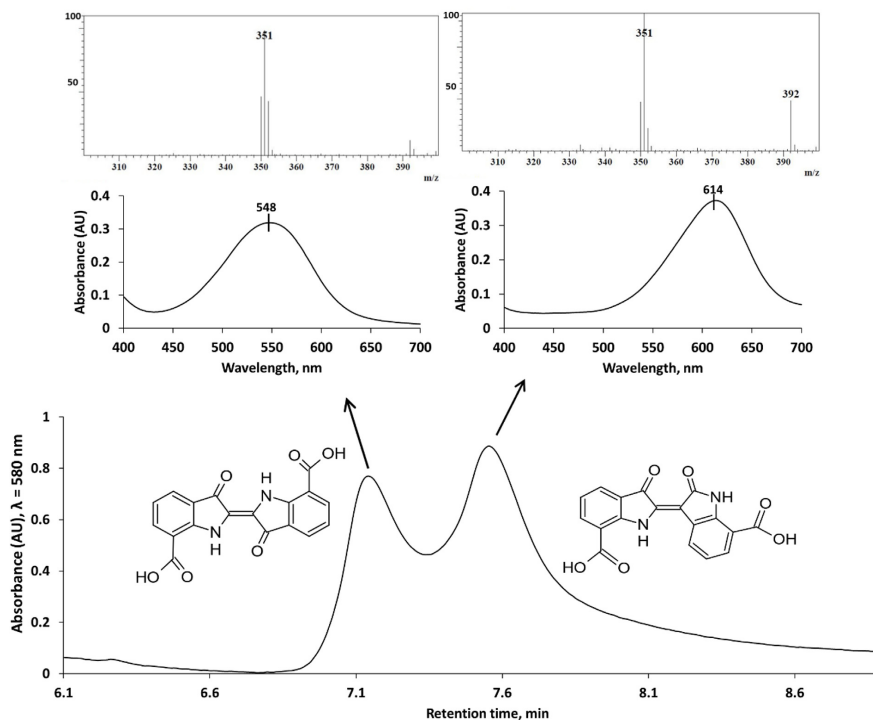


Fig. 2. RP-HPLC chromatogram (bottom), absorbance spectra (middle) and mass spectra (top) of indigo compounds obtained during whole cells bioconversion of indole-7-carboxylic acid.

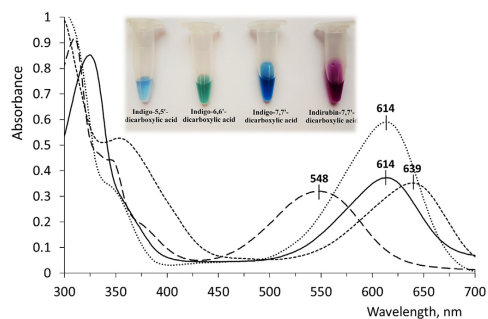


Fig. 3. Absorbance spectra of indigo dicarboxylic acids, dissolved in water. Solid line – indigo-5,5'-dicarboxylic acid, dashed line – indigo-6,6'-dicarboxylic acid, long-dashed line – indigo-7,7'-dicarboxylic acid, dotted line – indirubin-7,7'-dicarboxylic acid. Inset shows different colors of these indigo compounds. Inset shows aqueous solutions of indigo dicarboxylic acids. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Solubility of indigo dicarboxylic acids in different solvents. ND – not determined. * -aggregates are observed at higher concentration [37].

	Solubility, mM			λ_{\max} , nm	ϵ at λ_{\max} , $M^{-1}cm^{-1}$
	H ₂ O	Methanol	DMSO		
Indigo-5,5'-dicarboxylic acid	16	3	50	614 (H ₂ O)	2260
Indigo-6,6'-dicarboxylic acid	16	6	> 50	639 (H ₂ O)	3200
Indigo-7,7'-dicarboxylic acid	8	1.5	25	614 (H ₂ O)	5130
Indirubin-7,7'-dicarboxylic acid	13	> 13	> 50	548 (H ₂ O)	1430
Indigo [31]	0.0037	ND	1*	610 (DMSO)	22140
Indirubin [35]	0.0048	1.1	38	546 (DMSO)	9340
Indigo carmine [36]	21	ND	ND	608 (H ₂ O)	15000

2.4. Reduction of indigo dicarboxylic acids by pyrroloquinoline quinone (PQQ)-dependent glucose dehydrogenase

Soluble, stable and cheap redox mediators are of high demand for development of various biocatalytic devices and processes such as biosensors, biofuel cells and bioconversion methods [38,39]. PQQ-dependent oxidoreductases can oxidize a wide range of molecules, including carbohydrates and alcohols [40,41], but practical applications of such enzymes require an efficient enzyme reoxidation process

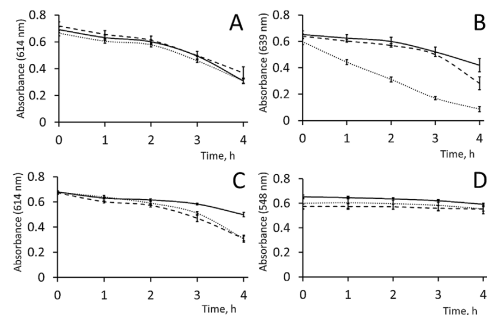


Fig. 4. Reduction of indigo dicarboxylic acids in the presence of glucose dehydrogenase. A – indigo-5,5'-dicarboxylic acid, B – indigo-6,6'-dicarboxylic acid, C – indigo-7,7'-dicarboxylic acid, D – indirubin-7,7'-dicarboxylic acid. Dotted lines – GDH and glucose, straight lines – no glucose, dashed lines – no GDH. Error bars represent standard deviations in three independent experiments.

[42–45], thus providing the basis for the search of appropriate electron acceptors.

Since the indigo dicarboxylic acids and indirubin dicarboxylic acid were soluble in water and appeared to be stable for a reasonable period of time, all four compounds prepared in this study were tested as possible electron acceptors for a recombinant PQQ-dependent glucose dehydrogenase (GDH) from *Acinetobacter calcoaceticus* strain LMD [42]. Among tested indigo dicarboxylic acids, only indigo-6,6'-dicarboxylic acid was reduced by GDH which was visible by both disappearance of color and the decrease in absorbance at 639 nm (Fig. 4, B). Other indigo dicarboxylic acids demonstrated slow decrease in absorbance even in the absence of glucose or GDH, suggesting spontaneous reduction of indigo dicarboxylic acids in this system. However, neither the reduction of indirubin-7,7'-dicarboxylic occurred nor did it serve as an electron acceptor for GDH.

3. Discussion

In this work, a range of novel indigo compounds were produced using two different bacterial oxygenase systems. The notable features of these new indigoids were high water solubility of indigo dicarboxylic acids and indirubin-7,7'-dicarboxylic acid. Since increasing the water solubility of indirubin and its derivatives significantly increased the extent of biological activity (see Introduction), the introduction of carboxyl groups could be a simple method for increasing the solubility of other biologically-active indigoids. Also, the functional groups of newly synthesized indigoids allow further modifications and generation of novel combinations of indigo compounds.

Recently, the synthesis of indigo dicarboxaldehydes was described in a patent application [46]. The potential usage of these compounds included dyeing synthetic textiles with desired colors, depending on the selected indigo derivative. Bioconversion of indole carboxaldehydes, described here, yielded a mixture of reduced indigo pigments. In spite of that, both oxygenase systems described here (Hind8 and PML G109Q) were confirmed to convert indole carboxaldehydes to corresponding indigoids. It was most likely the bioconversion host (*E. coli*), which was responsible for the reduction of aldehyde group. Such endogenous reducing activity of *E. coli* is well documented. The production of benzaldehyde ultimately resulted in the accumulation of benzyl alcohol instead of benzaldehyde and only the deletion of six putative aldo-keto reductases decreased the conversion of aldehydes into alcohols significantly [47]. Thus, *E. coli* strains with deleted aldo-keto reductase and possibly other genes could serve as an efficient platform for

production of indigo dicarboxaldehydes. Alternatively, other bioconversion hosts might be used to avoid the undesired alcohol formation.

No color development was observed on LB agar plates supplemented with indole carboxylic acid and cells with either Hind8 or PML G109Q. However, both oxygenase systems were capable of converting indole carboxylic acids to corresponding indigoids when a whole-cell approach was used. It should be mentioned that when LB medium was substituted with a minimal medium, a formation of blurry color was visible throughout the plate (Fig. S27). It might be suspected that the glucose or other medium components reduced the formed pigments, but no significant reduction of indigo dicarboxylic acids at least by glucose was observed with purified indigoids (Fig. 4). So far, the reasons for invisible indigo dicarboxylic acids on LB agar plates remain obscure.

Purification procedure for soluble indigoids described here can be compared to common extraction of water-insoluble indigoids using organic solvents. Although requiring more equipment and being more laborious, the procedure described here eliminates the need for harmful organic solvents. Also, the cells remain intact after the purification and can be reused for bioconversion.

According to phylogenetic analysis and the absence of flavin reductase gene adjacent to monoxygenase gene, the Hind8 enzyme most likely belongs to group B flavin monoxygenases. These enzymes contain a tightly bound FAD cofactor and depend on NADPH as coenzyme [48]. In turn, group B monoxygenases do not require flavin reductase component as the oxidation of NADPH, reduction of FAD and monoxygenation of substrate occur in the same enzyme. Purification of Hind8 was not successful, rendering the detection of a bound cofactor impossible.

It is worth noting that indoles substituted at position 4 were difficult to oxidize for both Hind8 and PML G109Q. 4,4'-(aminomethyl)-indigo was detected during bioconversion in traces only thus preventing further characterization of this indigo. In general, a bioconversion of 4-substituted indoles to corresponding indigoids is fairly rare in literature. For example, Namgung and co-workers tested three 4-substituted indole derivatives as substrates for cytochrome P450 class enzyme, and only one (4-nitroindole) was oxidized to 4,4'-dinitroindigo [15]. This implies that substitution at position 4 either hinders the positioning of the substrate in the active site pocket, or the reaction is prevented by some other means.

4. Conclusions

In this work, a total of 16 different indigo compounds were synthesized using two monoxygenase systems: bacterial flavin-dependent monoxygenase Hind8 and soluble diiron monoxygenase PML G109Q. 4,4'-di(hydroxymethyl)indigo, 5,5'-di(hydroxymethyl)indigo, 6,6'-di(hydroxymethyl)indigo, 7,7'-di(hydroxymethyl)indigo, 5,5'-di(aminomethyl)-indigo, 6,6'-di(aminomethyl)-indigo, 7,7'-di(aminomethyl)-indigo, indigo-4,4'-dicarboxaldehyde, indigo-5,5'-dicarboxaldehyde, indigo-6,6'-dicarboxaldehyde and indigo-7,7'-dicarboxaldehyde were extracted and characterized by RP-HPLC/MS. Purified indigo-5,5'-dicarboxylic acid, indigo-6,6'-dicarboxylic acid, indigo-7,7'-dicarboxylic acid and indirubin-7,7'-dicarboxylic acid were all soluble in water, methanol and DMSO. Among all novel indigoids, only the oxidation of indole-7-carboxylic acid resulted in the mixture of indigo and indirubin forms, both of which were separated from each another and characterized. Indigo-6,6'-dicarboxylic acid, but not other water-soluble indigoids served as an electron acceptor during glucose oxidation in the presence of glucose dehydrogenase.

5. Materials and methods

5.1. Reagents and bacterial strains

Reagents used in this study were purchased from the following

suppliers: indole (Sigma-Aldrich), 4-hydroxymethylindole, 5-hydroxymethylindole, 6-hydroxymethylindole, 7-hydroxymethylindole, 4-aminomethylindole, 5-aminomethylindole, 6-aminomethylindole, 7-aminomethylindole, indole-4-carboxaldehyde, indole-5-carboxaldehyde, indole-6-carboxaldehyde, indole-7-carboxaldehyde, indole-4-carboxylic acid, indole-5-carboxylic acid, indole-6-carboxylic acid, indole-7-carboxylic acid (all from Combi-Blocks, USA). All other reagents used in this study were of analytical or higher grade. *Escherichia coli* DH5 α was used as cloning host for metagenomic screening and plasmid amplification, *E. coli* HMS174(DE3) Δ *tnaA* (see below) was used as protein synthesis platform and as whole-cells for bioconversion of indole derivatives. All *E. coli* strains were cultivated in LB medium. When necessary, cells were suspended in buffer A (10 mM PBS, pH 7.7).

5.2. Construction and screening of metagenomic libraries

E. coli DH5 α cells were transformed with the metagenomic libraries (source – different soil samples) prepared as described previously [49,50] and plated on LB plates supplemented with indole or indole derivatives. After an overnight incubation at 30 °C, plates were inspected for formation of indigo-producing (blue) colonies. Plasmids from target colonies were isolated, sequenced and analyzed for potential oxygenase genes.

Phylogenetic analysis was performed with MEGA7 [51] using the Neighbor-Joining method, a total number of bootstrap trials was set to 1000. Amino acid sequences of selected oxygenases were retrieved from NCBI Genbank [52].

5.3. Construction of *E. coli* HMS174(DE3) Δ *tnaA* mutant

E. coli strain BW25113 JW3686 from Keio collection [53] served as the genetic source for the *tnaA:FRT-kan^R-FRT* deletion cassette. The deletion cassette was amplified using *dtnaAF* and *dtnaAR* DNA primers (Table 2). Chromosomal in-frame gene deletions in *E. coli* and subsequent *kan^R* marker removal were accomplished via a Quick & Easy *E. coli* Gene Deletion Kit (Gene Bridges GmbH) using *E. coli* HMS174(DE3) as a parental strain.

5.4. Cloning and expression of *Hind8* oxygenase encoding gene

The *Hind8* oxygenase encoding gene was amplified with primers *Hind8-F* and *Hind8-R* (Table 2) using Phusion Green Hot Start II High-Fidelity PCR Master Mix (Thermo Scientific) and pUC-*Hind8* plasmid as DNA matrix. The fragment obtained was gel-extracted with GeneJET Gel Extraction Kit (Thermo Scientific) and cloned to pLATE31 plasmid with aLICator LIC Cloning and Expression Kit 1 (Thermo Scientific) according to manufacturer's protocol and verified by sequencing. The *E. coli* HMS174(DE3) *tnaA* strain was transformed with the obtained pLATE31-*Hind8* plasmid. For gene expression, cells were cultivated in liquid LB medium with agitation (30 °C, 180 RPM) until reaching OD₆₀₀ = 0.8, then supplemented with IPTG to a final concentration of 0.1 mM, transferred immediately to 16 °C and cultivated overnight. Cells were collected by centrifugation, washed with buffer A and resuspended in the same buffer. Cells were lysed by sonication and the soluble fraction was obtained by centrifugation of the lysate at 4 °C and 20000 g for 15 min. Protein synthesis was visualized with 12% SDS-

PAGE.

5.5. Whole cells bioconversion of indole derivatives

E. coli cells with recombinant *Hind8* protein were suspended in buffer A supplemented with succinate (5 mM) to reach the 2x concentration of initial culture and the indole substrate was added to a concentration of 2 mM. Incubation of whole cells was performed at 16 °C overnight with agitation (180 RPM). Formation of indigoid compounds was monitored visually as the color of cell culture turned blue. Obtained indigoids (except for indigo dicarboxylic acids) were extracted with DMF. Typically, cell pellet was mixed well with 1 ml of DMF, incubated for 30 min at room temperature and centrifuged at 16000 g for 3 min. The supernatant was subjected to RP-HPLC and TLC. Indigo and indirubin dicarboxylic acids were extracted by pelleting the whole-cell bioconversion mixture and the supernatant was subjected to RP-FPLC for purification.

Expression of and whole cells bioconversion with the soluble diiron monooxygenase PML G109Q mutant was performed as described in Ref. [21], except that *E. coli* HMS174(DE3) Δ *tnaA* was used as expression host instead of *E. coli* BL21 (DE3).

Bioconversion rates were calculated as described [20].

5.6. Analytical techniques

Reversed-phase (RP) preparative Fast Performance Liquid Chromatography (FPLC) was performed with C18 Reveleris® Flash column (Grace) at constant flow of 2 ml min⁻¹. Column was equilibrated with 3 column volumes (CV) of buffer A and indigoid compounds were eluted with dH₂O. Column was regenerated with 3 CV of 20% ethanol. Fractions containing colored compounds were analyzed by RP-HPLC, pooled and subjected to the second round of preparative RP-FPLC with identical conditions to increase the purity of target compounds. Then chromatographically pure fractions were pooled, vacuum-dried and dissolved in selected solvents.

RP-HPLC/MS was performed as described in Ref. [23]. For TLC separation (silicagel matrix, Merck) of indigoids (except indigo dicarboxylic acids), the mobile phase was composed of chloroform:methanol (9:1). For indigo dicarboxylic acids, TLC plates were developed with 1,4-dioxane/2-propanol/water/NH₄OH, 4/2/2/1 as the mobile phase.

¹H NMR spectra were recorded in hexadeuterodimethyl sulfoxide (DMSO-*D*₆) on a Bruker Ascend 400 at 400 MHz. Spectra were calibrated with respect to the solvent signal (DMSO-*D*₆, ¹H δ = 2.50).

Indirubin-7,7'-dicarboxylic acid. Purple solid. ¹H NMR (400 MHz, DMSO-*d*₆): δ = 6.95 (t, *J* = 7.5 Hz, 2H, CH), 7.62 (d, *J* = 7.4 Hz, 1H, CH), 7.61–7.68 (dd, *J* = 7.7, 1.3 Hz, 1H, CH), 7.95–7.97 (dd, *J* = 7.5, 1.3 Hz, 1H, CH), 8.69–8.71 (m, 1H, CH), 10.89 (s, 1H, NH), 12.10 (s, 1H, NH).

Indigo-5,5'-dicarboxylic acid. Blue solid. ¹H NMR (400 MHz, D₂O): δ = 6.32 (d, *J* = 8.4 Hz, 2H, CH), 7.37 (dd, *J* = 8.5, 1.7 Hz, 2H, CH), 7.60 (d, *J* = 1.7 Hz, 2H, CH).

5.7. Determination of solubility and optical characteristics of indigoids

Purified indigo dicarboxylic acids in dried form were dissolved in certain amount of either dH₂O, methanol or DMSO and incubated overnight at room temperature. Following centrifugation at 20000 × g for 3 min, the supernatant was vacuum-dried and the amount of dissolved substance was calculated by weighting the dried indigoid powder. Solubility of indigo dicarboxylic acids in DMSO was determined as described [31]. Absorbance of indigoid solutions with known concentrations was measured with Lambda 25 UV/VIS spectrometer (PerkinElmer) and molar absorptivity coefficients were calculated using Beer's Law. Milli-Q water was used as aqueous solvent in solubility experiments.

Table 2
Oligonucleotides used in this study.

Oligo name	Oligo sequence
<i>dtnaAF</i>	5'- AATTGTGCGATCACCGCCCTT -3'
<i>dtnaAR</i>	5'- AAGGCACCCAGAAAAACCA -3'
<i>Hind8-F</i>	5'- AGAAGGAGATATAACTATGTGTGCGAGAAGAAGAAAGTCTGC -3'
<i>Hind8-R</i>	5'- GTGGTGTGATGGTGTATGGCCCGACTATACCGCGCG -3'

5.8. Glucose dehydrogenase electron transfer activity

Reaction mixture consisted of 20 mM glucose, 5 μ l (1 mg ml⁻¹) of purified recombinant glucose dehydrogenase from *Acinetobacter calcoaceticus* LMD [42], 0.2–0.5 mM indigo or indirubin dicarboxylic acid all added to buffer A. Mixture was incubated at room temperature and absorbance at 614 (indigo-5,5'- and indigo-7,7'-dicarboxylic acid), 639 (indigo-6,6'-dicarboxylic acid) or 548 nm (indirubin-7,7'-dicarboxylic acid) was measured every hour.

5.9. Sequence accession numbers

The sequence of the clone Hind8 was deposited to GenBank under accession number MN124688, accession number for the wild-type PML sequence is MK037457. In this work, a G109Q mutant of PML was used.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding sources

This work was supported by the European Union's Horizon 2020 research and innovation program [BlueGrowth: Unlocking the potential of Seas and Oceans] under grant agreement no. 634486 (project INMARE). M.S. and R.S. were also funded by the European Social Fund under the No 09.3.3-LMT-K-712 "Development of Competences of Scientists, other Researchers and Students through Practical Research Activities" measure, Grant No. 09.3.3-LMT-K-712-15-0023.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

Rita Buksnaitiene (Department of Organic Chemistry) is greatly acknowledged for ¹H NMR analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.dyepig.2019.107882>.

References

- [1] Splitstoser JC, Dillehay TD, Wouters J, Claro A. Early pre-Hispanic use of indigo blue in Peru. *Sci. Adv.* 2016;2:e1501623.
- [2] Mahapatra NN. Textile dyes. CRC Press; 2016.
- [3] Lin Y-K, et al. Anti-inflammatory effects of the extract of indigo naturalis in human neutrophils. *J Ethnopharmacol* 2009;125:51–8.
- [4] Marko D, et al. Inhibition of cyclin-dependent kinase 1 (CDK1) by indirubin derivatives in human tumour cells. *Br J Canc* 2001;84:283–9.
- [5] Leclerc S, et al. Indirubins inhibit glycogen synthase kinase-3 beta and CDK5/p25, two protein kinases involved in abnormal tau phosphorylation in Alzheimer's disease. A property common to most cyclin-dependent kinase inhibitors? *J Biol Chem* 2001;276:251–60.
- [6] Li Z, et al. Indirubin inhibits cell proliferation, migration, invasion and angiogenesis in tumor-derived endothelial cells. *Oncotargets Ther* 2018;11:2937–44.
- [7] Blažević T, et al. Indirubin and indirubin derivatives for counteracting proliferative diseases. *Evid Based Complement Altern Med* 2015. <https://doi.org/10.1155/2015/654098>.
- [8] Stalder R, Mei J, Graham KR, Estrada LA, Reynolds J R Isoindigo. A versatile electron-deficient unit for high-performance organic electronics. *Chem Mater* 2014;26:664–78.
- [9] Irimia-Vladu M, et al. Indigo – from jeans to semiconductors: indigo – a natural pigment for high performance ambipolar organic field effect transistors and circuits. *Adv. Mater.* 2012;24: 321–321.
- [10] Glowacki ED, et al. Indigo and tyrian purple – from ancient natural dyes to modern

- organic semiconductors. *Isr J Chem* 2012;52:540–51.
- [11] Nawn G, et al. Redox-active bridging ligands based on indigo diimine ("Nindigo") derivatives. *Inorg Chem* 2011;50:9826–37.
- [12] Fortier S, et al. Probing the redox non-innocence of dinuclear, three-coordinate Co (ii) nindigo complexes: not simply β -diketimate variants. *Chem Commun* 2012;48:11082.
- [13] Oakley SR, et al. "Nindigo": synthesis, coordination chemistry, and properties of indigo diimines as a new class of functional bridging ligands. *Chem Commun* 2010;46:6753–5.
- [14] Ensley BD, et al. Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science* 1983;222:167–9.
- [15] Namgung S, et al. Ecofriendly one-pot biosynthesis of indigo derivative dyes using CYP102G4 and PrnA halogenase. *Dyes Pigments* 2019;162:80–8.
- [16] Wu Z-L, Podust LM, Guengerich FP. Expansion of substrate specificity of Cytochrome P450 2A6 by random and site-directed mutagenesis. *J Biol Chem* 2005;280:41090–100.
- [17] Hu S, et al. Altering the regioselectivity of cytochrome P450 BM-3 by saturation mutagenesis for the biosynthesis of indirubin. *J Mol Catal B Enzym* 2010;67:29–35.
- [18] McClay K, Boss C, Keresztes I, Steffan JR. Mutations of Toluene-4-monoxygenase that alter regioselectivity of indole oxidation and lead to production of novel indigoid pigments. *Appl Environ Microbiol* 2005;71:5476–83.
- [19] Zhang X, et al. Cloning and expression of naphthalene dioxygenase genes from *Comamonas* sp. MQ for indigoids production. *Process Biochem* 2013;48:581–7.
- [20] Petkevičius V, et al. A Biocatalytic synthesis of heteroaromatic N-oxides by whole cells of *Escherichia coli* expressing the multicomponent, soluble di-iron mono-oxygenase (SDIMO) PmlABCDDEF. *Adv Synth Catal* 2019;361:2456–65.
- [21] Petkevičius V, Vaitekūnas J, Vaitkus D, Čėnas N, Meškys R. Tailoring a soluble diiron mono-oxygenase for synthesis of aromatic N-oxides. *Catalysts* 2019;9:356.
- [22] Qu Y, et al. Unveiling the biotransformation mechanism of indole in a *Cupriavidus* sp. strain. *Mol Microbiol* 2017;106:905–18.
- [23] Sadauskas M, Vaitekūnas J, Gasparavičiūtė R, Meškys R. Genetic and biochemical characterization of indole biodegradation in *Acinetobacter* sp. Strain O153. *Appl. Environ. Microbiol.* AEM. 2017;17:01453. <https://doi.org/10.1128/AEM.01453-17>.
- [24] Heine T, Großmann C, Hofmann S, Tischler D. Indigoid dyes by group E mono-oxygenases: mechanism and biocatalysis. *Biol Chem* 2019;400:939–50.
- [25] Dai C, et al. Application of an efficient indole oxygenase system from *Cupriavidus* sp. SHE for indigo production. *Bioproc Biosyst Eng* 2019. <https://doi.org/10.1007/s00449-019-02189-4>.
- [26] Steingruber E. Indigo and indigo colorants. 2004. <https://doi.org/10.1002/14356007.a14.149.pub2>.
- [27] Murdock D, Ensley BD, Serdar C, Thalen M. Construction of metabolic operons catalyzing the de novo biosynthesis of indigo in *Escherichia coli*. *Bio Technology* 1993;11:381.
- [28] Klein LL, et al. A novel indigoid anti-tuberculosis agent. *Bioorg Med Chem Lett* 2014;24:268–70.
- [29] Lavinda O, et al. Singular thermochromic effects in dyeings with indigo, 6-bromoindigo, and 6,6'-dibromoindigo. *Dyes Pigments* 2013;96:581–9.
- [30] Tratnyek PG, et al. Visualizing redox chemistry: probing environmental oxidation-reduction reactions with indicator dyes. *Chem Educ* 2001;6:172–9.
- [31] Carretero-González J, Castillo-Martínez E, Armand M. Highly water-soluble three-redox state organic dyes as bifunctional analytes. *Energy Environ Sci* 2016;9:3521–30.
- [32] Ginzinger W, et al. Water-soluble cationic derivatives of indirubin, the active anticancer component from Indigo naturalis. *Chem Biodivers* 2012;9:2175–85.
- [33] Cheng X, et al. Identification of a water-soluble indirubin derivative as potent inhibitor of insulin-like growth factor 1 receptor through structural modification of the parent natural molecule. *J Med Chem* 2017;60:4949–62.
- [34] Seixas de Melo J, Moura AP, Melo MJ. Photophysical and spectroscopic studies of indigo derivatives in their keto and leuco forms. *J Phys Chem A* 2004;108:6975–81.
- [35] Indirubin. ALX-270-361 - Enzo life sciences Available at: <http://www.enzolifesciences.com/ALX-270-361/indirubin/> Accessed: 24th July 2019.
- [36] Indigotindisulfonic acid Available at: <https://www.drugbank.ca/drugs/DB11577> Accessed: 24th July 2019.
- [37] Bernardino ND, Brown-Xu S, Gustafson TL, de Faria DLA. Time-resolved spectroscopy of indigo and of a Maya blue analog. *J Phys Chem C* 2016;120:21905–14.
- [38] Nasar A, Perveen R. Applications of enzymatic biofuel cells in bioelectronic devices – a review. *Int J Hydrogen Energy* 2019;44:15287–312.
- [39] Bollella P, Gorton L. Enzyme based amperometric biosensors. *Curr. Opin. Electrochem.* 2018;10:157–73.
- [40] Laurinavicius V, et al. Bioelectrochemical application of some PQQ-dependent enzymes. *Bioelectrochemistry Amst. Neth.* 2002;55:29–32.
- [41] Marcinkevičienė L, et al. Biocatalytic properties of quinohemoprotein alcohol dehydrogenase IIG from *Pseudomonas putida* HK5. *Chemija* 2012;23:223–32.
- [42] Tetianec L, Bratkovskaja I, Kulyš J, Casalta V, Meškys R. Probing reactivity of PQQ-dependent carboxylate dehydrogenases using artificial electron acceptor. *Appl Biochem Biotechnol* 2011;163:404–14.
- [43] Tetianec L, et al. Characterization of methylated azopyridine as a potential electron transfer mediator for electroenzymatic systems. *Process Biochem* 2017;54:41–8.
- [44] Ratautas D, Ramonas E, Marcinkevičienė L, Meškys R, Kulyš J. Wiring gold nanoparticles and redox enzymes: a self-sufficient nanocatalyst for the direct oxidation of carbohydrates with molecular oxygen. *ChemCatChem* 2018;10:971–4.
- [45] Ramonas E, Ratautas D, Dagnys M, Meškys R, Kulyš J. Highly sensitive amperometric biosensor based on alcohol dehydrogenase for determination of glycerol in human urine. *Talanta* 2019;200:333–9.
- [46] Kaplan G, et al. Use of indigo derivatives for dyeing synthetic textiles, novel indigo derivatives and process for dyeing synthetic textiles. 2019.

- [47] Kunjapur AM, Tarasova Y, Prather KLJ. Synthesis and accumulation of aromatic aldehydes in an engineered strain of *Escherichia coli*. *J Am Chem Soc* 2014;136:11644–54.
- [48] van Berkel WJH, Kamerbeek NM, Fraaije MW. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J Biotechnol* 2006;124:670–89.
- [49] Časaitė, V. et al. Engineering of a chromogenic enzyme screening system based on an auxiliary indole-3-carboxylic acid monooxygenase. *Microbiology* 0, e795.
- [50] Urbelienė N, et al. Application of the uridine auxotrophic host and synthetic nucleosides for a rapid selection of hydrolases from metagenomic libraries. *Microb. Biotechnol.* 2019;12:148–60.
- [51] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33:1870–4.
- [52] Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res* 2016;44:D67–72.
- [53] Baba T, et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2006;2.

Publication III

Sadauskas, M., Vaitekūnas, J., Gasparavičiūtė, R., Meškys, R. Indole Biodegradation in *Acinetobacter* sp. Strain O153: Genetic and Biochemical Characterization. *Applied and Environmental Microbiology*, 2017, 83(19):e01453-17
DOI: 10.1128/AEM.01453-17



Indole Biodegradation in *Acinetobacter* sp. Strain O153: Genetic and Biochemical Characterization

Mikas Sadauskas, Justas Vaitekūnas, Renata Gasparavičiūtė, Rolandas Meškys

Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Life Sciences Center, Vilnius University, Vilnius, Lithuania

ABSTRACT Indole is a molecule of considerable biochemical significance, acting as both an interspecies signal molecule and a building block of biological elements. Bacterial indole degradation has been demonstrated for a number of cases; however, very little is known about genes and proteins involved in this process. This study reports the cloning and initial functional characterization of genes (*iif* and *ant* cluster) responsible for indole biodegradation in *Acinetobacter* sp. strain O153. The catabolic cascade was reconstituted *in vitro* with recombinant proteins, and each protein was assigned an enzymatic function. Degradation starts with oxidation, mediated by the *lifC* and *lifD* flavin-dependent two-component oxygenase system. Formation of indigo is prevented by *lifB*, and the final product, anthranilic acid, is formed by *lifA*, an enzyme which is both structurally and functionally comparable to cofactor-independent oxygenases. Moreover, the *iif* cluster was identified in the genomes of a wide range of bacteria, suggesting the potential of widespread *lif*-mediated indole degradation. This work provides novel insights into the genetic background of microbial indole biodegradation.

IMPORTANCE The key finding of this research is identification of the genes responsible for microbial biodegradation of indole, a toxic *N*-heterocyclic compound. A large amount of indole is present in urban wastewater and sewage sludge, creating a demand for an efficient and eco-friendly means to eliminate this pollutant. A common strategy of oxidizing indole to indigo has the major drawback of producing insoluble material. Genes and proteins of *Acinetobacter* sp. strain O153 (DSM 103907) reported here pave the way for effective and indigo-free indole removal. In addition, this work suggests possible novel means of indole-mediated bacterial interactions and provides the basis for future research on indole metabolism.

KEYWORDS indole, biodegradation, bacterial metabolism, *Acinetobacter*, bacterial signaling, cofactor-independent oxygenases

Indole is an *N*-heterocyclic aromatic compound derived mainly by TnaA tryptophanase from L-tryptophan in *Escherichia coli* (1) and is widely found in natural environments. Indole acts as cell-to-cell signaling molecule that regulates the expression of several virulence genes (2–4), promotes biofilm formation (5–7), and mediates complex predator-prey interactions (8, 9). At high concentrations, indole and its derivatives exhibit toxic activity to both prokaryotic cells and animals and are even considered mutagens (10). Toxic indole concentrations reportedly vary for different microorganisms in the range of 0.5 to 5 mM (11). The main mechanisms of indole toxicity are reported to be an alteration of membrane potential with subsequent inhibition of cell division (12), depletion of ATP levels (13), and an inhibition of acyl-homoserine lactone (AHL)-based quorum sensing by regulator misfolding (14).

In order to utilize aromatic compounds as an energy source, microorganisms have

Received 3 July 2017 Accepted 26 July 2017

Accepted manuscript posted online 4 August 2017

Citation Sadauskas M, Vaitekūnas J, Gasparavičiūtė R, Meškys R. 2017. Indole biodegradation in *Acinetobacter* sp. strain O153: genetic and biochemical characterization. *Appl Environ Microbiol* 83:e01453-17. <https://doi.org/10.1128/AEM.01453-17>.

Editor Volker Müller, Goethe University, Frankfurt am Main

Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Address correspondence to Mikas Sadauskas, mikas.sadauskas@bchi.vu.lt.

to cope with the problem of high-resonance energy that stabilizes the aromatic ring system (15). A common strategy is the use of oxygenases and O_2 , which itself requires the activation of dioxygen. In addition to being thermodynamically unfavorable, reactions between dioxygen (triple state) and most of the organic compounds (singlet state) are not possible due to a spin barrier (16). Diverse elements, including transition metals (iron, manganese, and copper) or organic cofactors (flavins and pterin), are used extensively by oxygenases to form a superoxide, a reactive singlet state form of dioxygen (17). A remarkable group of cofactor-independent oxygenases have been described which require neither an organic cofactor nor a metal to catalyze the incorporation of (di)oxygen into a single molecule of an organic substrate (18, 19). Establishment of the catalytic mechanisms for this group of enzymes provides interesting mechanistic insights into substrate-assisted oxygen activation (20, 21).

To defend against the toxicity of indole, bacteria that encounter indole have established enzymatic detoxification systems, notably the oxidation of indole to insoluble nontoxic indigoid pigments (22, 23), and biodegradation mechanisms (24, 25). A number of indole-degrading bacterial microorganisms (26–28) as well as bacterial consortia (29) were reported previously, but no genetic background in these reports has been specified. Several possible intermediates in bacterial indole degradation are also known, but proteins with specific enzyme activities that drive the degradation cascade have not been identified in this context. In this paper, the identification of an indole degradation gene cluster (*iif*) in indole-degrading *Acinetobacter* sp. strain O153 is reported. The catabolic pathway was reconstituted *in vitro* with the recombinant proteins encoded by the *iif* genes, and the function of each enzyme was identified by analyzing the reaction products.

RESULTS

Screening for indole-degrading bacteria. Numerous microbiome samples obtained from invertebrates as well as soil samples were used for the screening of indole-degrading bacteria as these environments potentially contain indole. Due to the known toxicity of indole, this compound was not used as a sole carbon source for selection of indole-mineralizing bacteria in this study. Instead, a derivative of indole was hypothesized to form a dead end pigment in the indole-degrading microorganisms. Such a strategy enabled the selection of desirable bacteria on nutrient-rich medium supplemented with the chromogenic substrate 5-bromoindoline.

A bacterial colony isolated from the intestine of the crustacean *Orconectes limosus*, which is a widespread invasive species in Europe, was able to convert 5-bromoindoline to insoluble indigoid pigment. The sequence of the 16S rRNA gene of this strain showed 99% similarity to the 16S rRNA gene sequence of *Acinetobacter guillouiae* strain NBRC 110550, and therefore strain O153 was designated *Acinetobacter* sp. O153. Notably, this strain did not produce indigoid pigment when grown on Luria-Bertani (LB) agar plates with indole (data not shown).

To test the ability of *Acinetobacter* sp. O153 to assimilate indole as a carbon (C) and/or nitrogen (N) source, cells of strain O153 were cultured in minimal medium with succinate or indole as a C source and NH_4Cl or indole as an N source. No growth was observed with indole as the only C source or the only source for both N and C (Fig. 1A). However, moderate growth (up to an optical density at 600 nm [OD_{600}] of 0.3 in 12 h) was achieved with indole (1 mM) as a nitrogen source and succinate as the sole carbon source. While this effect was observed with indole concentrations ranging from 0.5 mM to 1.5 mM, no growth was recorded with concentrations of 0.1 mM and 2 mM (Fig. 1B), which were most likely insufficient and toxic, respectively.

Whole-cell experiments were performed to determine whether the expression of enzymes involved in indole metabolism are inducible or constitutive. After overnight growth in a minimal medium with succinate and with or without indole, cells of strain O153 were resuspended in buffer containing 1 mM indole. Indole was transformed immediately by indole-induced O153 cells (Fig. 2). In contrast, uninduced cells con-

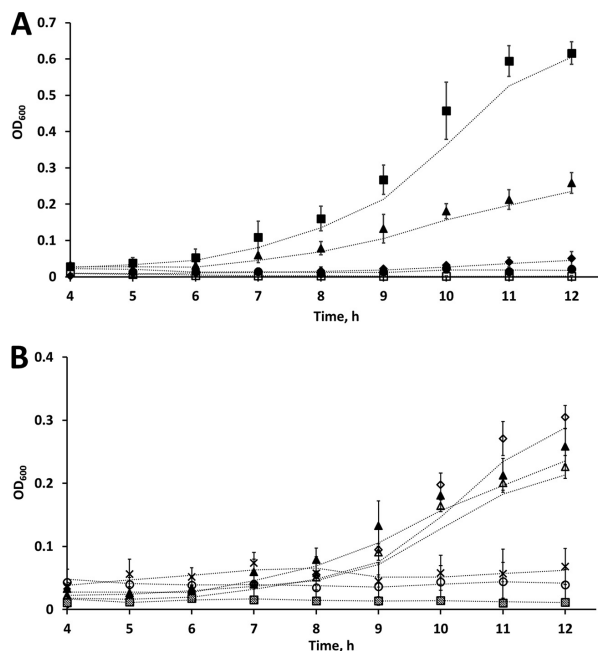


FIG 1 Growth kinetics of *Acinetobacter* sp. strain O153. (A) Growth in minimal medium with different carbon and nitrogen sources. (B) Growth in minimal medium supplemented with different concentrations of indole as a nitrogen source and succinate as a carbon source. Filled squares, NH₄Cl (5 g/liter) and succinate (5 mM) (positive control); filled triangles, indole (1 mM) and succinate; filled rhombus, NH₄Cl and indole (1 mM); filled circles, indole (1 mM); open squares, no N source and succinate (negative control); open circles, 0.1 mM indole; crosses, 0.5 mM indole; open triangles, 0.75 mM indole; open rhombus, 1.5 mM indole; striped squares, 2 mM indole. Dashed lines represent trend lines using a moving average data approximation (period = 2).

sumed indole significantly more slowly and with a lag period. These results indicated that indole mineralization (biodegradation) was an inducible process.

Identification and characterization of genes involved in indole degradation.

Screening of an *Acinetobacter* sp. O153 genomic library using *E. coli* DH5 α as a host strain revealed one clone (O153H14) which was able to convert 5-bromoindoline to a purple pigment, a characteristic feature of the wild-type strain O153. Clone O153H14 was bearing a pUC19 vector with a 12-kb genomic DNA fragment. The DNA sequencing of this fragment revealed the presence of nine open reading frames (ORFs) (Fig. 3). Bioinformatic analysis of the ORFs in this fragment using NCBI's blastn algorithm and the Conserved Domain Database (CCD) (30) allowed the identification of two different sets of genes (Table 1). The *iif* cluster comprised the first set, and the second set encoded a multisubunit anthranilate dioxygenase. The two sets were separated by a putative *araC* transcription regulator. Analysis of the *iif* cluster suggested both high sequence identity and structural similarity to *iif* genes, which were shown to comprise a functional operon induced by indole in *Acinetobacter baumannii* ATCC 19606 (23). Hence, genes and proteins presented here were given the same nomenclature and were called *iif* homologs. Five genes of the *iif* operon encoded putative enzymes, as described by conserved domain analysis: *iifA* encodes a dienelactone hydrolase family protein, *iifB* encodes a short-chain dehydrogenase, *iifC* encodes an oxygenase, *iifD*

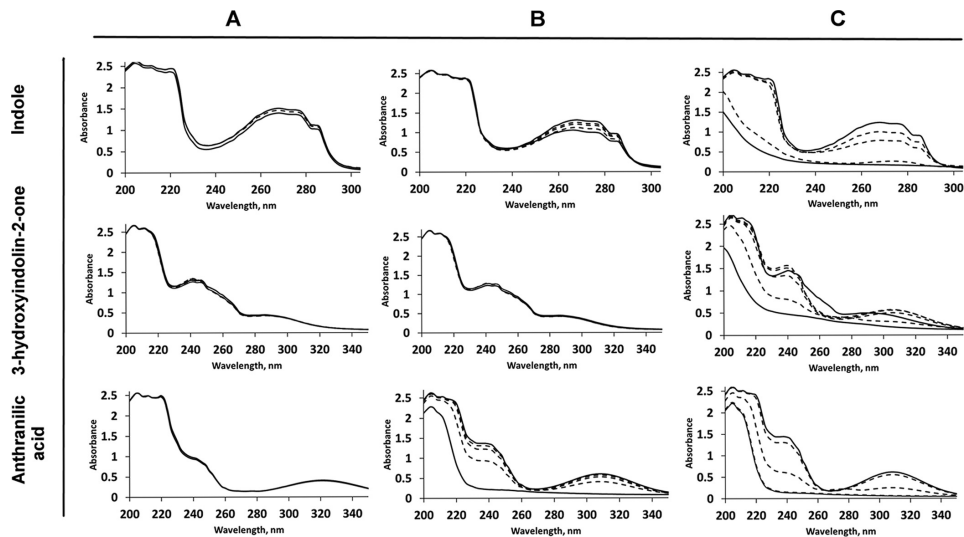


FIG 2 Whole-cell bioconversion assays using different substrates and cells of *Acinetobacter* sp. O153. (A) Substrates without cells (negative control). (B) Bioconversion with uninduced (overnight culture grown without indole) cells. (C) Bioconversion with induced (overnight culture grown in the presence of 1 mM indole) cells. Solid lines indicate spectra of initial and final products; dashed lines indicate spectra of intermediate products during a 6-h bioconversion cycle.

encodes a flavin reductase, and *iifE* encodes a putative phenol degradation protein. The *ant* cluster was composed of three genes encoding different subunits of the anthranilate dioxygenase and appeared to be of typical composition: *antA* and *antB* encoded large and small subunits of the anthranilate dioxygenase oxygenase component, respectively, while *antC* coded for the electron transfer component protein.

Distribution of *iif* homologs among bacterial genomes. The prevalence of *iif* operon genes as well as genes of anthranilate dioxygenase among available microbial genomes was analyzed. Mainly organisms from the genera *Acinetobacter*, *Pseudomonas*, and *Burkholderia* were identified, including some notoriously pathogenic strains of *A. baumannii*, *Pseudomonas fluorescens*, and *Pseudomonas syringae* (Fig. 3). Spatial separation of the two operons was found to be a common indication outside the genus *Acinetobacter*, stretching to as long as 200 kb (*Burkholderia cepacia* ATCC 25416 and *Alcaligenes faecalis* ZD02) or even separate replicons (*Burkholderia contaminans* MS14, *Cupriavidus metallidurans* CH34, and *Cupriavidus basilensis* 4G11). Also, aberrant composition of the *iif* operon was predicted in several cases with four genomes (*P. fluorescens* NCIMB 11764, *C. metallidurans* CH34, *C. basilensis* 4G11, and *Ralstonia solanacearum* PS107) lacking the *iifE* homolog.

Enzyme activity of recombinant *lif* proteins. To identify the enzymatic functions of *lif* proteins, *iifA*, *iifB*, *iifC*, *iifD*, and *iifE* genes were expressed in *E. coli* BL21(DE3) with an N-terminal His tag, purified, and tested for activity toward indole, which was monitored by high-performance liquid chromatography–mass spectrometry (HPLC-MS). Molecular masses of the purified proteins observed in SDS-PAGE gels corresponded well to the theoretical masses (46 kDa for *lifA*, 28.1 kDa for *lifB*, 46.3 kDa for *lifC*, and 19.9 kDa for *lifD*) (see Fig. S1 in the supplemental material).

lifC was found to be capable of using indole as a substrate only in the presence of *lifD* flavin reductase, flavin adenine dinucleotide (FAD) cofactor, and NAD(P)H (Fig. 4, peak 3). The product of this reaction was a blue insoluble indigoid pigment (Fig. S2).

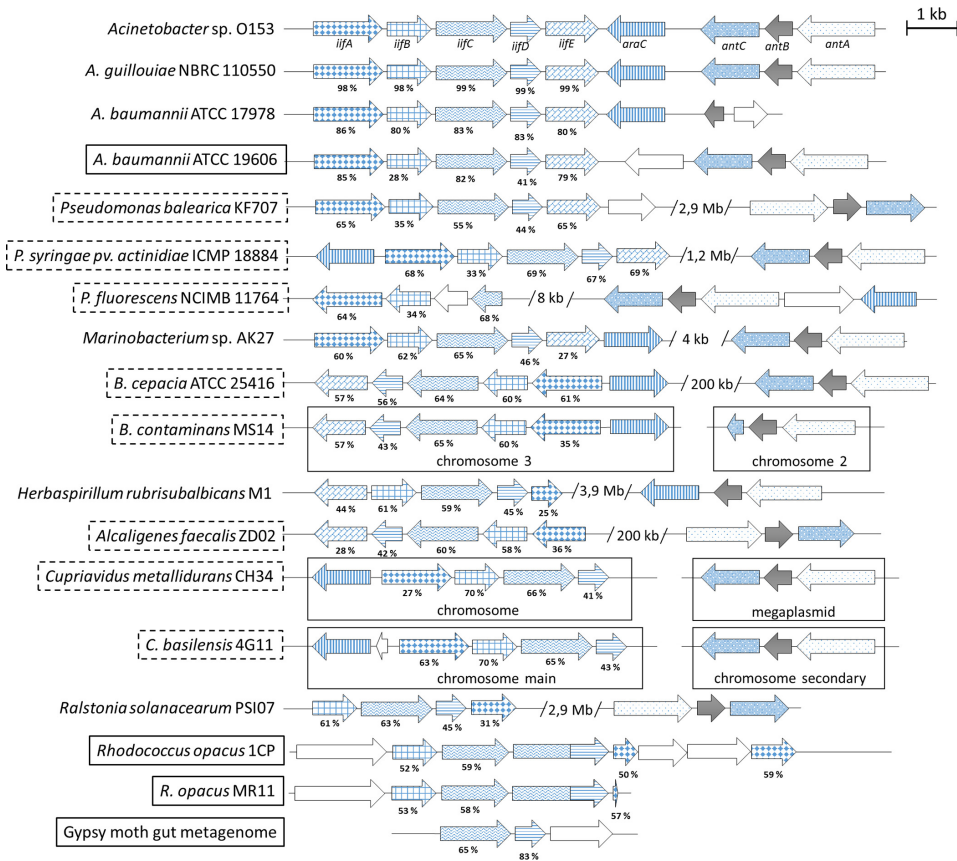


FIG 3 Organization and distribution of *iif* and *ant* genes in different microbial genomes. Genes are represented by arrows (drawn to scale as indicated). Homologous genes are highlighted in the same pattern according to the scheme in the top line for *Acinetobacter* sp. O153. White arrows, indole degradation unrelated/unknown. Strains and genomic fragments boxed in solid lines indicate reported activity of certain *lif*-homologous proteins; dashed boxes highlight strains for which indole biodegradation activity was reported at the genus level. Identities (percent) of amino acid sequences between *lif* proteins of strain O153 and homologs are indicated under the corresponding ORFs.

Next, when *lifB* was added to the reaction mixture (*lifC*, *lifD*, FAD, NADPH, and indole), indigo formation was abolished, and a product with a retention time (4.75 min) identical to that of 3-hydroxyindolin-2-one was observed (Fig. 4, peak 1). The UV-visible light (UV-Vis) absorption spectra of this compound and 3-hydroxyindolin-2-one were very similar, with peaks at 211 nm, 253, and 291 nm (Fig. S3E and F). Also, both of these compounds demonstrated identical molecular masses: 150 (3-hydroxyindolin-2-one + H⁺) and 132 (3-hydroxyindolin-2-one - H₂O + H⁺) (Fig. S3E and F). Moreover, 3-hydroxyindolin-2-one was consumed by wild-type indole-induced *Acinetobacter* sp. O153 cells in a similar way as indole (Fig. 2). 3-Hydroxyindolin-2-one was rapidly transformed into anthranilic acid, which was then slowly consumed by indole-induced cells. Meanwhile, no consumption was observed with uninduced cells. These results

TABLE 1 Predicted conserved domains and putative functions of proteins encoded by genes identified in a 12-kb genomic fragment of *Acinetobacter* sp. O153

Protein	Conserved domain (identifier)	E value	Closest homolog with known function			Identity (%)	Reference
			Protein ^a	GenBank accession no. or identifier	Function		
IIfA	α/β hydrolase (COG0412) Snoal-like polyketide cyclase (pfam07366)	2.3e-53 0.04	C1CD from <i>P. knackmussii</i> B13	AB871539.1	Dienelactone hydrolase	34	66
IIfB	Short-chain dehydrogenase (cd05233)	1.2e-56	FabG from <i>V. cholerae</i> N16961	VC2021	Reduction of β -ketoacyl-ACP to 3-hydroxyacyl-ACP	33	67
IIfC	Flavin reductase (COG1853)	6.8e-40	IIfC from <i>A. baumannii</i> ATCC 19606	F911_02004	Monoxygenase	82	23
IIfE	MetA-pathway of phenol degradation (cd01768)	1.3e-57	IIfD from <i>A. baumannii</i> ATCC 19606	F911_02003	Flavin reductase	81	23
IIfR	AraC-binding-like domain (pfam14525)	4e-39	Ppu2725 from <i>P. putida</i> F1	4RL8	Outer membrane channel of unknown function	21	68
AntA	Rieske non-heme iron oxygenase (cd03469)	9e-46	IIfR from <i>A. baumannii</i> ATCC 19606	F911_02001	Transcriptional regulator	74	23
AntB	C-terminal catalytic domain of the oxygenase alpha subunit (cd08879)	9e-83	AntA from <i>Acinetobacter</i> sp. ADP1	AAC34813.1	Oxygenase component of anthranilate dioxygenase	93	31
AntC	Ring hydroxylating beta subunit (cd00667)	2e-51	AntB from <i>Acinetobacter</i> sp. ADP1	AAC34814.1	Oxygenase component of anthranilate dioxygenase	94	31
	Benzoate dioxygenase reductase (cd06209)	3e-107	AntC from <i>Acinetobacter</i> sp. ADP1	AAC34815.1	Reductase component of anthranilate dioxygenase	87	31
	2Fe-2S cluster (cd00207)	3e-21					

^a*P. knackmussii*, *Pseudomonas knackmussii*; *V. cholerae*, *Vibrio cholerae*.

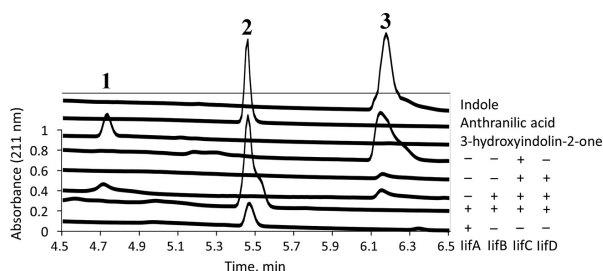


FIG 4 Stacked HPLC chromatograms (211 nm) of enzymatic reaction products and standards. Top three curves, standards; middle four curves, reaction products of indole and lif proteins as indicated on the right; bottom curve, reaction product of 3-hydroxyindolin-2-one and lifA. Numbers indicate peaks of three major compounds: peak 1, 3-hydroxyindolin-2-one; peak 2, anthranilic acid; peak 3, indole.

indicated that 3-hydroxyindolin-2-one was the reaction product of an lifB-catalyzed reaction and did not act as an inducer for indole-degrading proteins.

When four lif proteins (lifC, lifD, lifB, and lifA) were used in a reaction with indole, a compound with m/z 138 $[M+H^+]$ was identified, corresponding to $[137+H^+]$, with the molecular mass of anthranilic acid (Fig. S3A and B). In addition, this reaction product and anthranilic acid shared identical retention times (Fig. 4, peak 2) and UV-Vis absorption spectra (Fig. S3A and B). Importantly, lifA protein was also capable of using 3-hydroxyindolin-2-one as the substrate to form anthranilic acid (Fig. 4, bottom curve). Notably, conversion of 3-hydroxyindolin-2-one to anthranilic acid by lifA did not require any cofactor or metal ion and did not occur with other substrates of similar structure (isatin or *N*-formyl-anthranilic acid). The presence of lifE protein in any of the above-mentioned reaction mixtures did not change the outcomes of the reactions. Thus, analysis of reaction products formed by lifA both supported the involvement of 3-hydroxyindolin-2-one as an intermediate in indole degradation and showed the formation of anthranilic acid, a possible end product of lifCDBA-catalyzed indole biodegradation.

The kinetic parameters of lifA, lifC, and lifD proteins were determined (Table 2). lifD flavin reductase showed a preference for flavin mononucleotide (FMN) and NADH. In spite of that, FAD, but not FMN or riboflavin, was identified as the only flavin cofactor suitable for indole oxidation performed by lifC (data not shown).

Identification of lifA as a possible cofactor-independent oxygenase. Structural and functional approaches were used to analyze the enzymatic mechanism of lifA. The conversion of 3-hydroxyindolin-2-one to anthranilic acid by purified recombinant lifA

TABLE 2 Kinetic parameters of lifD, lifC and lifA proteins^a

Protein and substrate or cofactor	K_m (μ M)	K_{cat} (min^{-1})	K_{cat}/K_m ($\text{min}^{-1} \mu\text{M}^{-1}$)
lifD			
NADH	29 ± 0.1	1.21 ± 0.15	0.041
NADPH	7.9 ± 0.4	0.13	0.017
FMN	2.85 ± 0.19	2.7 ± 0.45	0.94
FAD	7.07 ± 0.12	7.02 ± 0.64	0.99
Riboflavin	8.35 ± 0.83	3.1 ± 0.43	0.38
lifC			
Indole	250 ± 13	0.11 ± 0.01	0.00047
lifA			
3-Hydroxyindolin-2-one	281.7 ± 13.4	2476 ± 310	17.52

^aValues are means \pm standard deviations.

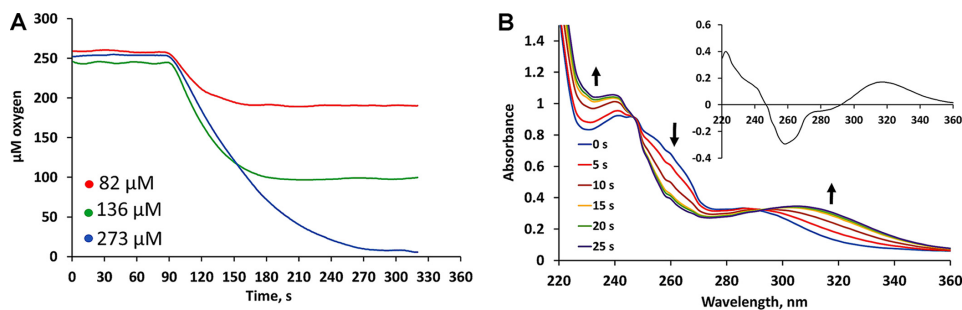


FIG 5 Functional characterization of lifA. (A) Oxygen consumption by lifA. Red, green, and blue indicate different substrate concentrations as indicated. Reactions were initiated at 90 s. (B) UV spectra of the transition between 3-hydroxyindolin-2-one and anthranilic acid, catalyzed by lifA. The inset shows the spectral difference between the substrate and the product.

did not require the addition of any cofactor or metal ions. No cofactor-indicating peaks were observed in the UV-Vis spectrum of the purified lifA protein (Fig. S4). Addition of 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.5 mM HgCl₂, or 3 mM EDTA did not inhibit the reaction, suggesting a nonhydrolytic metal ion-independent enzymatic conversion. Moreover, oxygen was consumed in the above-mentioned reaction in a substrate-dependent manner (Fig. 5A); i.e., addition of 3-hydroxyindolin-2-one equivalent to the oxygen concentration resulted in complete consumption of oxygen, while a 1:2 molar ratio of substrate/oxygen consumed half of the oxygen present in an air-restricted reaction cell. Notably, the reaction did not occur in the absence of dioxygen (Fig. S5). Oxidation of 3-hydroxyindolin-2-one to anthranilic acid by lifA was monitored spectrophotometrically, and a decrease in absorbance at 260 nm as well as increases in absorbance at 225 nm and 315 nm was observed (Fig. 5B). The presence of two isobestic points indicated that no intermediates were formed during this reaction.

Protein sequence analysis of lifA performed using the CDD indicated the presence of two domains: an N-terminal diene lactone hydrolase (DLH)-like domain and C-terminal SnoaL-like domain (Fig. S6B). Attempts to obtain separate recombinant domains were unsuccessful due to insolubility, suggesting a major structural role of both domains during protein folding. A three-dimensional model of lifA was generated with the I-TASSER platform. A standard α/β hydrolase fold was located in the N-terminal domain and was composed of a β -sheet of seven strands, with the first strand antiparallel to the others. A typical catalytic triad, common to the majority of hydrolases, was located in this domain and was composed of Cys124, Asp173, and His204 (Fig. S6A). Residues of this putative catalytic triad were located close to each other in the three-dimensional model. Although the overall sequence identity between the N-terminal domain of lifA and well-described cofactor-independent dioxygenases Hod and Qdo was low (17.9% with Hod and 19.1% with Qdo), the structural resemblance was clearly evident (Fig. S7).

DISCUSSION

The present article deals with identification of genes in *Acinetobacter* sp. strain O153 that are required for degradation of the *N*-heterocyclic compound indole. These genes are clustered into an *lif* operon, with five genes coding for potential enzymes. The initial description of proteins encoded by genes of the *lif* operon clearly showed the ability of lifC to oxidize indole leading to formation of indigo pigment (23). Here, the ability of lifC to oxidize indole to indigo was confirmed, but no appearance of indigo was observed when three proteins, namely, lifC, lifD, and lifB, were incubated with indole *in vitro* although the concentration of indole readily decreased. Incorporation of all four

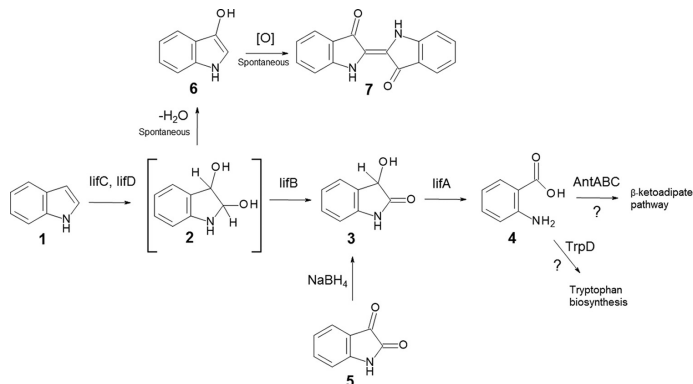


FIG 6 Proposed indole biodegradation pathway in *Acinetobacter* sp. O153: step 1, indole; step 2, indole-2,3-dihydrodiol; step 3, 3-hydroxyindolin-2-one; step 4, anthranilic acid; step 5, isatin; step 6, 1H-indol-3-ol (indoxyl); step 7, indigo. Brackets indicate an intermediate not experimentally detected.

proteins (lifC, lifD, lifB, and lifA) into the reaction mixture resulted in the formation of anthranilic acid. Importantly, genes encoding a multisubunit anthranilate dioxygenase are located close to the *lif* operon in most of the organisms potentially capable of degrading indole or located in separate chromosomes, as in the case of *Cupriavidus metallidurans* CH34 and *C. basiliensis* 4G11 (Fig. 3). Anthranilate dioxygenase converts anthranilic acid to catechol (31), which is then catabolized through a well-established β -ketoadipate pathway (32). Such a degradation pathway where indole is oxidized to metabolites, directly reaching the tricarboxylic acid (TCA) cycle, might enable microorganisms to both counteract the toxic effect of indole and use it as a carbon and energy source. Although this may be true for some other microorganisms with reported indole utilization as a sole carbon source and an *lif-ant* combination identified, strain O153 did not grow with indole as the only carbon source. This suggests difficulties in coping with the toxic effects of indole (see the introduction and references cited therein) or insufficient metabolic activity to shuttle the indole degradation products into the TCA cycle, for example, as a result of low expression levels or relatively poor catalytic efficiency of catechol dioxygenase (33, 34). Still, anthranilic acid was consumed by the cells of the strain O153 with no product-indicating spectra suggesting utilization of anthranilate in other cellular processes, possibly as a substrate for anthranilate phosphoribosyltransferase during the synthesis of tryptophan (35). Strain O153 used indole (at concentrations ranging from 0.5 to 1.5 mM) as the sole nitrogen source, indicating a physiological benefit when the bacteria occupy indole-containing environments.

Based on the results presented above, a scheme for indole biodegradation in *Acinetobacter* sp. strain O153 is proposed (Fig. 6). Degradation starts with indole oxidation at the C-2 and C-3 positions, forming indole-2,3-dihydrodiol. This compound is known to be rather unstable (22) and therefore was not detected. A spontaneous dehydration of indole-2,3-dihydrodiol results in the formation of indoxyl, which is prone to auto-oxidation and forms an insoluble indigo pigment. However, the loss of a water molecule could be prevented by lifB, forming a stable intermediate. The hypothetical short-chain dehydrogenase lifB performs oxidation at the C-2 position to obtain 3-hydroxyindolin-2-one. Anthranilic acid, formed by lifA, is then the end product of lifCDBA-catalyzed indole biodegradation.

Proteins with notable sequence similarity to lif proteins were described earlier, leading to coherent concepts and hypotheses. First, the MoxY (65% sequence identity

to *lifC*) and *MoxZ* (38% sequence identity to *lifD*) proteins with indole oxidation activity were suggested to be involved in the production of a quorum-sensing-inducing compound (36). Later, a genomic fragment of *Rhodococcus opacus* 1CP containing several genes with high structural similarity to the *lif* operon was published (37), emphasizing its role in styrene biodegradation. Monooxygenase activity of this fragment was detected by indole oxidation to indigo and was attributed specifically to *StyA2B*, a styrene-epoxidizing fusion protein containing both oxidation and flavin-reduction activities (47% sequence identity between the N-terminal segment of *StyA2B* and *lifC*; 52% sequence identity between *lifD* and the C-terminal segment of *StyA2B*). ORFs 6 and 9, comprising a possible structural homolog of *lifA* of this fragment and forming a hypothetical diene lactone hydrolase, were interrupted by transposase, most likely rendering this protein inactive. A complete set of *lif*-homologous genes was then found in *R. opacus* MR1, including an adjacent multisubunit anthranilate dioxygenase, leading to the hypothesis that the locus described therein might be involved in the degradation of heteroaromatic compounds (38). Finally, the indole detoxification ability of the *lifC*-*lifD* system as well as the flavin oxidoreductase activity of *lifD* was demonstrated in representatives of the genus *Acinetobacter* (23, 39). From the current perspective, it appears that a link between high metabolic versatility of acinetobacters and indole as the substrate for *lif* proteins is clearly evident, suggesting biodegradation as the fate of indole when exposed to *lif*-containing bacteria.

In the proposed model of indole biodegradation, *lifC* should act as a dioxygenase, incorporating oxygen atoms at both the C-2 and C-3 positions. However, *lifC* appears to have typical structural elements of flavin-containing monooxygenases (23). Although difficult to detect experimentally, indole dioxygenation was reported for a number of bacterial dioxygenases and monooxygenases (25, 40). Indole-2,3-dihydrodiol, a product of possible dioxygenase attack, was observed in the biotransformation of indole by *Cupriavidus* sp. strain KK10 (41). Notably, several species of *Cupriavidus* were identified to possess the *iif-ant* cluster (Fig. 3). Indole-2,3-dihydrodiol was not observed directly as a reaction intermediate during the indole degradation described here, which might be attributed to the instability of this compound (22). Importantly, the formation of indigo from indole-2,3-dihydrodiol was detected in an indole oxidation reaction catalyzed by naphthalene dioxygenase (22). A possible mechanism of indole-2,3-dihydrodiol formation might be hydrolysis of indole 2,3-epoxide, which was identified as an intermediate during indole epoxidation performed by a styrene monooxygenase (42) and was suggested to form 2,3-*trans* diol spontaneously or enzymatically (43). The opening of the epoxide to a diol was presumed to be a plausible pathway for indole-2,3-dihydrodiol formation (44). Moreover, the oxidation of several *N*-heterocyclic compounds catalyzed by monooxygenases when two oxygen atoms (one from oxygen, another from water) were inserted was described recently (45–47). A similar mechanism for bacterial degradation of indole acetic acid was described (48) involving the intermediate dioxindole-3-acetic acid (DOAA), a structural analogue of 3-hydroxyindolin-2-one in indole degradation. Formation of DOAA was not attributed to a single enzymatic step, and either a second hydroxylation or hydroxy/keto oxidoreduction mechanisms were suggested. Given these points, deciphering the mechanism of *lifC* requires further structural studies.

The role of *lifE* in the degradation of indole remains poorly understood and might even be ruled out based on experimental evidence. First, no changes were observed when recombinant *lifE* protein was added to a reaction mixture consisting of indole and other *lif* proteins; i.e., such a reaction still yielded anthranilic acid with complete substrate consumption. Next, the expression levels of *lifE* did not change upon addition of indole, according to quantitative reverse transcription-PCR (qRT-PCR) results (23). Finally, the gene coding for *lifE* appears to be absent in genomes of several microorganisms that possess the *iif* operon. Interestingly, the presence of a putative signal sequence was detected in this protein, with a cleavage site between residues 24 and 25 (AQA-YD), using SignalP (49). Also, this protein showed low sequence identity to the Pput2725 membrane channel from *Pseudomonas putida* F1. Taken together, such

inconsistent insights provide no precise functional description for this protein, but they suggest possible functions, such as facilitation of indole uptake or other metabolic activities, that are worth exploring.

As already noted, a distinct group of extraordinary dioxygenases catalyzes the incorporation of oxygen atoms into organic substrates with no requirements for cofactors. Structural and functional properties of *lifA* appear to be very similar to those of the best described cofactor-independent dioxygenases such as 1H-3-hydroxy-4-oxoquinoline 2,4-dioxygenase (Qdo) from *Pseudomonas putida* 33/1 and 1H-3-hydroxy-4-oxoquinoline 2,4-dioxygenase (Hod) from *Arthrobacter ilicis* R61a (50). All of these enzymes share an α/β structural fold and a speculative catalytic triad, which was demonstrated to act rather as a dyad as mutants of catalytic Ser retained substantial activity (18, 50). Catalytic properties of *lifA* also resemble those of cofactor-free dioxygenase from an aerobic Gram-positive coccus participating in indole biodegradation described by Fujioka and Wada (51). This cofactor-free dioxygenase was able to convert 2-oxo-3-hydroxyindoline to stoichiometric amounts of anthranilic acid and CO₂ with the consumption of equivalent amounts of dioxygen (51). As it stands, the establishment of *lifA* as an oxygenase catalyzing the ring-opening reaction requires additional experiments to elucidate the fate of the C-2 carbon atom of 3-hydroxyindolin-2-one during the oxidation process as well as site-directed mutagenesis to prove the role of putative catalytic amino acid residues, and these experiments are under way.

Notably, a number of bacterial genera with *iif* and *ant* operons identified in this work have already been described as indole-degrading microorganisms. *B. unamae* strain CK43B was found to degrade indole under aerobic conditions when supplemented with gallic acid or pyrogallol (27), showing a similar cometabolic effect of indole degradation as in the case of strain O153. Importantly, indoxyl was detected as the reaction intermediate as well as compounds of the β -ketoacid pathway, anthranilic acid, and catechol, signifying the possible relationship between *lif* proteins and indole biodegradation in the genus *Burkholderia*. *Cupriavidus* sp. strain SHE was also reported to be able to perform indole biodegradation (28); however, no intermediates were identified, except for isatin. A compound with an *m/z* corresponding to that of isatin was also observed during *lif*-catalyzed indole biodegradation. Nonetheless, based on results presented here indicating (i) that, in contrast to indole, isatin was not consumed by resting cells during an uptake assay and (ii) that isatin was not a substrate for any of the *lif* proteins described here, we suggest that isatin might not be an intermediate during *lif*-catalyzed indole degradation but, rather, appeared as a dead end product of spontaneous oxidation or an experimental artifact. While indole biodegradation characteristics were also evaluated for representatives of the genera *Pseudomonas* (*P. aeruginosa* strain Gs) and *Alcaligenes* (*Alcaligenes* sp. strain In3) (26, 52), no data are available regarding the biodegradation of indole in the genus *Acinetobacter* to date. Such a high prevalence of the *iif* operon might be explained by a dual beneficial effect of indole degradation. Indole, produced mainly by *E. coli* in the intestine, can decrease the virulence of *P. aeruginosa* (53). This effect was later attributed to the indole-caused inhibition of AHL-based quorum-sensing signaling by altering the folding of the AqsR regulator (14, 54). Such an effect was reported for a number of Gram-negative bacteria, including *Chromobacterium violaceum*, *Pseudomonas chlororaphis*, *Serratia marcescens* (55), and *Acinetobacter oleivorans* DR1 (14), a strain which was identified here to possess the *iif-ant* operon combination (Fig. 3). In a complex and competition-prone environment such as the gastrointestinal tract, indole production might become an advantage if the producer does not utilize AHL-based signaling, which is the case for *E. coli* (56). The degradation of indole and its utilization as an additional carbon and energy source using the *lif-AntABC* enzymatic system described here might be a possible solution for AHL producers, including members of the genus *Acinetobacter* (57), to counteract the toxic indole effect and retain fitness. Taken together, the concordance between *de facto* indole-degrading microbial genera and the prevalence of the *iif* and *ant* operon combination suggests that proteins encoded by the *iif* operon might be responsible for

TABLE 3 Primers used in this study^a

Primer name	Sequence (5'–3')	Purpose	Reference or source
CMBFNdel	GATCATATGTCAGGCCAAGATATTGAA	Amplification and cloning of <i>iifA</i> gene	This work
CMBRXhol	CTCTCGAGTTAAAACGATTGGCCTTCA		
SCDFNdel	GAGTGGCATATGGATATTGAATTGAATCAG	Amplification and cloning of <i>iifB</i> gene	This work
SCDRXhol	ATCTCGAGTCAACGCTTCATCGGCTA		
APRLF	GATCATATGCTGCTGATCGTATAG	Amplification and cloning of <i>iifC</i> gene	This work
APRLR	CTACTCGAGTTAAGCCACTTTTTGAC		
<i>iifDFN</i> del	TGGCTAGCATGAATATCAACACATC	Amplification and cloning of <i>iifD</i> gene	This work
<i>iifDRX</i> hol	CTGCTCGAGTTAAATACTCAGTTTC		
MetAFNcol	CTCATGGCCAATAATGTGATCAAAAG	Amplification and cloning of <i>iifE</i> gene	This work
MetARHindIII	GAAGCTTTCAGAAGTGATGAATATAAC		
27F	AGAGTTTGATCMGGCTCAG	Amplification and sequencing of 16S rRNA gene	59
1492R	TACGGYATCTTGTACGACTT		

^aUnderlined nucleotide sequences are introduced sequences for recognition of restriction endonucleases.

indole biodegradation in the majority of microorganisms described to date as indole degraders.

MATERIALS AND METHODS

Bacterial strains, chemicals, and standard techniques. *Escherichia coli* DH5 α and BL21(DE3) strains were used as cloning and protein expression hosts, respectively. Transformation of bacteria by electroporation was performed as described previously (58). *E. coli* strains carrying plasmids were cultivated in Luria-Bertani (LB) medium supplemented with antibiotics (100 μ g/ml ampicillin or 50 μ g/ml kanamycin). Plasmid DNA was isolated using a ZR Plasmid Miniprep Classic kit (Zymo Research). Indole, isatin, and anthranilic acid were purchased from Sigma-Aldrich; 5-bromoindoline was obtained from Combi-Blocks, Inc. All other chemicals used in this study were of analytical grade. All media and reagent solutions were prepared with Milli-Q water (Merck Millipore).

Isolation and identification of an indole-oxidizing bacterial strain. Indole-oxidizing microorganisms were isolated from the intestine of *Orconectes limosus*. Crustacean samples were collected at Lake Ruzis, Alytus district, Lithuania. Intestines were suspended in 0.9% NaCl solution and plated on LB agar plates supplemented with 1 mM 5-bromoindoline. Following an overnight incubation at 30°C, purple pigment-forming colonies were isolated and maintained as pure cultures. The 16S rRNA gene of the isolate strain was amplified with universal bacterial primers 27F and 1492R (Table 3) according to the method of Frank et al. (59) and sequenced. Phylogenetic analysis was performed with the blastn algorithm (<https://blast.ncbi.nlm.nih.gov>) (58) against the database of 16S rRNA sequences.

Genomic DNA of strain O153 was extracted as described previously (60), digested with HindIII (ThermoFisher Scientific) to obtain 5- to 20-kb DNA fragments, which were cloned into pUC19 (ThermoFisher Scientific). The resulting library was transformed into *E. coli* DH5 α ; cells were plated on LB agar with 5-bromoindoline, incubated overnight at 37°C, and screened for purple pigment-producing colonies.

Whole-cell bioconversion assays and growth conditions. For growth kinetics experiments, an overnight culture of strain O153 grown in LB medium was centrifuged, washed with potassium phosphate buffer, and resuspended in 0.9% NaCl. Cells were diluted in fresh M9 medium with different carbon (5 mM succinate or 1 mM indole) and nitrogen (5 g/liter NH₄Cl or different concentrations of indole) sources. Cell suspensions were incubated at 30°C with shaking (180 rpm). Growth was monitored by measuring the absorbance at 600 nm. At least three independent experiments were performed.

For whole-cell experiments, an overnight culture of *Acinetobacter* sp. O153 was grown in M9 medium (3.5 g/liter Na₂HPO₄, 1.5 g/liter KH₂PO₄, 2.5 g/liter NaCl, 0.2 g/liter MgSO₄, 0.01 g/liter CaCl₂) supplemented with 5 mM succinate and 1 mM indole (induced cells) or with succinate and 5 g/liter NH₄Cl (uninduced cells). Cells were pelleted by centrifugation, washed with 20 mM potassium phosphate buffer, pH 7.8, and resuspended in M9 minimal medium to a final OD₆₀₀ of 0.8 in a volume of 1 ml containing a 1 mM concentration of the appropriate substrate. Cells were incubated at 30°C with shaking. A sample without cells was used as an evaporation (negative) control. Absorbance spectra of the cell-free supernatants were recorded every hour to monitor substrate consumption.

Cloning, expression, and purification of proteins involved in indole degradation. All reagents used for the cloning experiments were purchased from ThermoFisher Scientific (Lithuania). Target genes were amplified by PCR in a Mastercycler ep Gradient S (Eppendorf) instrument using the primer pairs listed in Table 3, Maxima Hot Start Green PCR master mix, and a colony of *Acinetobacter* sp. O153 cells. PCR conditions were as follows: initial denaturation at 98°C for 4 min, followed by 30 cycles of denaturation at 98°C for 30s, annealing at various temperatures for 30s, elongation at 72°C for 90s, and a final elongation at 72°C for 10 min. DNA fragments obtained were cloned into pTZ57R/T vector according to the manufacturer's protocol and sequenced. Resulting plasmids were digested with restriction endonucleases that recognize the primer-introduced sequences (Tables 3 and 4), and fragments were ligated into pET28-(+) (Novagen) previously digested with the same restriction endonucleases. Expression plasmids with N-terminally encoded 6 \times His tags were transformed into *E. coli* BL21(DE3) for protein expression. Protein expression and purification were carried out as described

TABLE 4 Plasmids used in this study

Plasmid	Description and/or purpose	Source
pTZ57R/T	TA cloning of PCR products	ThermoFisher Scientific (Lithuania)
pUC19	Construction of genomic library	ThermoFisher Scientific (Lithuania)
pET28-c(+)	Expression of recombinant proteins	Novagen (Germany)
pET28- <i>iifA</i>	<i>iifA</i> gene, amplified by PCR with CMBFNdel and CMBRXhol primers, digested with NdeI/XhoI, and cloned into pET28-c(+)	This work
pET28- <i>iifB</i>	<i>iifB</i> gene, amplified by PCR with SCDFNdel and SCDRXhol primers, digested with NdeI/XhoI, and cloned into pET28-c(+)	This work
pET28- <i>iifC</i>	<i>iifC</i> gene, amplified by PCR with CMBFNdel and CMBRXhol primers, digested with NdeI/XhoI, and cloned into pET28-c(+)	This work
pET28- <i>iifD</i>	<i>iifD</i> gene, amplified by PCR with <i>iifDFN</i> del and <i>iifDRX</i> hol primers, digested with NheI/XhoI, and cloned into pET28-c(+)	This work
pET21- <i>iifE</i>	<i>iifE</i> gene, amplified by PCR with MetAFNcoI and MetARHindIII primers, digested with NcoI/HindIII, and cloned into pET28-c(+)	This work

earlier (61). Protein concentration was determined with a Coomassie protein assay kit (Pierce) using bovine serum albumin as the standard.

Enzyme assays and detection of reaction products. Flavin reductase activity of the purified *iifD* protein was determined from the decrease of the absorbance at 340 nm due to oxidation of NADH or NADPH ($\epsilon_{340} = 6,220 \text{ M}^{-1} \text{ cm}^{-1}$), using a spectrophotometer and was performed at room temperature. A total reaction volume of 1 ml contained 50 mM Tris-HCl, pH 7.5, variable amounts of NADH or NADPH, and flavins (FAD, FMN, or riboflavin). Kinetic parameters for NADH and NADPH were determined using a constant concentration of FAD (20 μM) and various concentrations of NAD(P)H (5 to 200 μM). In addition, various concentrations of flavin (1 to 40 μM) and a constant concentration of NADH (200 μM) were used for analysis of the enzyme kinetics. Reactions were initiated by adding 0.3 μg of the enzyme. One unit (U) of enzyme activity was defined as the amount of enzyme catalyzing the oxidation of 1 μmol of NAD(P)H per minute.

Indole oxidation activity of *iifC* was determined as a function of indigo formation in the reaction mixture. An initial indole oxidation rather than spontaneous dimerization was assumed to be a rate-limiting step. A typical reaction mixture contained 50 mM Tris-HCl, pH 7.5, 100 μM flavin, 250 μM NADH or NADPH, and various concentrations of indole (10 to 500 μM) and was initiated by adding 3 μg of *iifD* as well as 3 μg of *iifC*. Reaction mixtures were incubated at room temperature for 15 min, and indigo particles were pelleted by centrifugation at $16,000 \times g$ for 5 min. The precipitant was dissolved in 100 μl of dimethylformamide (DMF); dissolving was facilitated by incubating the mixture at 37°C for 15 min. Indigo concentration was determined as described previously (62), using the molar absorption coefficient $\epsilon_{620} = 14,000 \text{ M}^{-1} \text{ cm}^{-1}$. One unit of enzyme activity was defined as the amount catalyzing the formation of 1 μmol of indigo per minute.

Oxygenase activity of *iifA* was measured from the increase of absorbance at 315 nm (formation of anthranilic acid from 3-hydroxyindolin-2-one). A typical assay mixture contained various concentrations (10 to 300 μM) of 3-hydroxyindolin-2-one (synthesized as described below) in 50 mM Tris-HCl, pH 7.5, air-saturated (oxygen concentration, $\sim 250 \mu\text{M}$) buffer. Reactions were initiated by adding 3 μg of *iifA*, and the concentration of anthranilic acid formed was determined using the molar absorption coefficient $\epsilon_{315} = 1,600 \text{ M}^{-1} \text{ cm}^{-1}$ (calculated using the Beer-Lambert law and data on the absorbance of 0.05 to 5 mM anthranilic acid in 50 mM Tris-HCl, pH 7.5). One unit of enzyme activity was defined as the amount catalyzing the formation of 1 μmol of anthranilic acid per minute under the conditions described above. All kinetic parameters were determined by fitting the data to Lineweaver-Burk (double-reciprocal) plots and performing a linear regression. All kinetic experiments were performed in duplicate, and average means were derived.

Enzymatic conversion of indole by *iif* proteins was analyzed by monitoring substrate consumption and formation of corresponding reaction products in *in vitro* reactions. Under standard conditions, a total volume of 250 μl of reaction mixture contained 50 mM Tris-HCl, pH 7.5, 50 μM flavin cofactor, 100 μM NAD(P)H, 1 mM indole, and 3 μg of each purified *iif* protein. The combinations of proteins tested were as depicted in Fig. 4. Reaction mixtures were incubated at room temperature for 1 h with shaking (700 rpm; Eppendorf Thermomixer Comfort). Substrate consumption and formation of products were analyzed spectrophotometrically in PowerWave XS plate reader (BioTek Instruments, Inc.) or mixed with an equal volume of acetonitrile, centrifuged at $16,000 \times g$ for 5 min, and analyzed by HPLC-MS.

Oxygen consumption and anaerobic assay. The measurement of oxygen consumption was performed with a homemade computer-assisted membrane oxygen electrode and 10.0-ml glass cell. The concentration of oxygen was assumed to be 0.25 mM in air-saturated 50 mM Tris-HCl, pH 7.5, buffer solution at 25°C. Different concentrations of 3-hydroxyindolin-2-one (82, 136, and 273 μM) were used, and reactions were initiated by adding 3 μg of *iifA*.

Anaerobic conditions were generated in an air-restricted cell by argon bubbling of Tris-HCl buffer solution, pH 7.5, at 25°C for 10 min. Prior to bubbling, traces of oxygen were eliminated from argon gas by passage through a solution of alkaline pyrogallol (benzene-1,2,3-triol).

HPLC-MS. HPLC-MS analyses were performed using a high-performance liquid chromatography system (CBM-20A controller, two LC-2020AD pumps, SIL-30AC auto sampler, and CTO-20AC column oven; Shimadzu, Japan) equipped with a photodiode array (PDA) detector (Shimadzu, Japan) and mass spectrometer (LCMS-2020; Shimadzu, Japan) equipped with an electrospray ionization (ESI) source. The chromatographic separation was conducted using a YMC Pack Pro column (3 by 150 mm; YMC, Japan) at 40°C and a mobile phase that consisted of 0.1% formic acid water solution (solvent A) and acetonitrile (solvent B) delivered in the 0 to 60% gradient elution mode. Mass scans were measured from m/z 100 to m/z 700 at a 350°C interface temperature, 250°C desolvation line (DL) temperature, $\pm 4,500$ V interface voltage, and neutral DL/Qarray, using N_2 as nebulizing and drying gas. Mass spectrometry data were acquired in both positive and negative ionization modes. The data were analyzed using LabSolutions liquid chromatography-mass spectrometry (LCMS) software.

Chemical synthesis of 3-hydroxyindolin-2-one. Reduction of isatin to 3-hydroxyindolin-2-one was performed according to Bergonzini and Melchiorre (63). Isatin (1.5 mmol) was added in small portions to a stirred suspension of sodium borohydride (0.75 mmol; 2 eq) in 20 ml of a 1:1 water-ethanol mixture at room temperature. The mixture was vigorously stirred until the suspension became colorless (about 10 min). The mixture was extracted with chloroform (3 times, 10 ml each). The combined organic extracts were dried ($MgSO_4$), and the solvent was evaporated under reduced pressure. The residual material was dissolved in 5 ml of deionized water and was purified by reverse-phase chromatography (12 g; C_{18} cartridge) to separate the 3-hydroxyindolin-2-one from the pigments formed during the extraction and evaporation procedures. Prior the purification, the column was equilibrated with water. A mobile phase that consisted of water and methanol was delivered in the gradient elution mode. The collected fractions were analyzed by HPLC-MS. The fractions containing a pure product were combined, and the solvent was removed under reduced pressure.

The structures of the bioconversion products were determined using 1H nuclear magnetic resonance (NMR) and ^{13}C NMR. 1H NMR spectra were recorded in hexadeuteriodimethyl sulfoxide ($DMSO-D_6$) or $CDCl_3$ on a Bruker Ascend 400 at 400 MHz, and ^{13}C NMR spectra were recorded on a Bruker Ascend 400 at 100 MHz. All products were dissolved in deuterated dimethyl sulfoxide. Spectra were calibrated with respect to the solvent signal ($CDCl_3$, $^1H \delta = 7.26$, $^{13}C \delta = 77.2$; $DMSO-D_6$, $^1H \delta = 2.50$, $^{13}C \delta = 39.5$).

Sequence alignments and structure modeling. Genome sequences containing genes with high sequence similarity to *iif* genes were identified with the blastn suite (<https://blast.ncbi.nlm.nih.gov>) by using each of the five *iif* genes individually. Hits with E values less than $5e-10$ were pooled and used as a database in MultiGeneBlast (64). A homology search was performed using an *iif* locus of *A. guillouiae* genomic DNA (ranging from nucleotides 1661941 to 1671563) as a query with default parameters except that the maximum distance between genes in the locus was increased to 100 kb. Positive variants were manually proofread with respect to the structural organization of the *iif* and *ant* genes as well as possible orientation in separate DNA fragments.

The model of tertiary structures of the N-terminal (amino acids 1 to 233) and C-terminal (amino acids 252 to 374) domains of the *lifA* protein were obtained using the I-TASSER platform (<http://zhanglab.cmb.med.umich.edu/I-TASSER>) (65). Structures of dienelactone hydrolases (PDB accession numbers 1DIN and 1Z18) were used as threading templates for modeling of the N-terminal domain, and structures of putative polyketide cyclases (PDB accession numbers 4LQG and 3F9S) were used as templates for the C-terminal domain. Models with the highest C-scores (1.42 and 0.26 for N- and C-terminal domains, respectively) were selected for further structural analysis and comparison with structures of cofactor-independent oxygenases.

Accession number(s). Sequence data described in this paper have been submitted to GenBank database under the following accession numbers: 16S rRNA gene of strain *Acinetobacter* sp. O153, [KX955254](https://doi.org/10.1093/nar/kx955254); *iifA*, [KX955255](https://doi.org/10.1093/nar/kx955255); *iifB*, [KX955256](https://doi.org/10.1093/nar/kx955256); *iifC*, [KX955257](https://doi.org/10.1093/nar/kx955257); *iifD*, [KX955258](https://doi.org/10.1093/nar/kx955258); *iifE*, [KY700688](https://doi.org/10.1093/nar/ky700688). The strain *Acinetobacter* sp. O153 has been deposited in the DSMZ under accession number DSM 103907.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/AEM.01453-17>.

SUPPLEMENTAL FILE 1, PDF file, 0.8 MB.

ACKNOWLEDGMENTS

We are grateful to Regina Vidziunaite and I. Bratkovskaja for technical assistance with oxygen consumption measurements and A. Lauryenas for generation of anaerobic conditions.

We have no conflicts of interest to declare.

This work was supported by the Research Council of Lithuania (project no. MIP-042/2012).

REFERENCES

- Deeley MC, Yanofsky C. 1981. Nucleotide sequence of the structural gene for tryptophanase of *Escherichia coli* K-12. *J Bacteriol* 147:787–796.
- Hirakawa H, Inazumi Y, Masaki T, Hirata T, Yamaguchi A. 2005. Indole induces the expression of multidrug exporter genes in *Escherichia coli*.

- Mol Microbiol 55:1113–1126. <https://doi.org/10.1111/j.1365-2958.2004.04449.x>.
- Mueller RS, Beyhan S, Saini SG, Yildiz FH, Bartlett DH. 2009. Indole acts as an extracellular cue regulating gene expression in *Vibrio cholerae*. *J Bacteriol* 191:3504–3516. <https://doi.org/10.1128/JB.01240-08>.
 - Yang Q, Pande GSJ, Lin B, Rubin RA, Vora GJ, Defoirdt T. 2017. Indole signaling and microalgal auxins decrease the virulence of *Vibrio campbellii*, a major pathogen of aquatic organisms. *Environ Microbiol* 19:1987–2004. <https://doi.org/10.1111/1462-2920.13714>.
 - Martino PD, Fursy R, Bret L, Sundararaju B, Phillips RS. 2003. Indole can act as an extracellular signal to regulate biofilm formation of *Escherichia coli* and other indole-producing bacteria. *Can J Microbiol* 49:443–449. <https://doi.org/10.1139/w03-056>.
 - Sasaki-Imamura T, Yano A, Yoshida Y. 2010. Production of indole from L-tryptophan and effects of these compounds on biofilm formation by *Fusobacterium nucleatum* ATCC 25586. *Appl Environ Microbiol* 76:4260–4268. <https://doi.org/10.1128/AEM.00166-10>.
 - Kim D, Sitepu IR, Hashidoko Y. 2013. Induction of biofilm formation in the Betaproteobacterium *Burkholderia unamae* CK43B exposed to exogenous indole and gallic acid. *Appl Environ Microbiol* 79:4845–4852. <https://doi.org/10.1128/AEM.01209-13>.
 - Erb M, Veyrat M, Robert CA, Xu H, Frey M, Ton J, Turlings TC. 2015. Indole is an essential herbivore-induced volatile priming signal in maize. *Nat Commun* 6:6273. <https://doi.org/10.1038/ncomms7273>.
 - Lee J-H, Kim Y-G, Kim M, Kim E, Choi H, Kim Y, Lee J. 2017. Indole-associated predator-prey interactions between the nematode *Caenorhabditis elegans* and bacteria. *Environ Microbiol* 19:1776–1790. <https://doi.org/10.1111/1462-2920.13649>.
 - Ochiai M, Wakabayashi K, Sugimura T, Nagao M. 1986. Mutagenicities of indole and 30 derivatives after nitrite treatment. *Mutat Res* 172:189–197. [https://doi.org/10.1016/0165-1218\(86\)90056-X](https://doi.org/10.1016/0165-1218(86)90056-X).
 - Lee JH, Lee J. 2010. Indole as an intercellular signal in microbial communities. *FEMS Microbiol Rev* 34:426–444. <https://doi.org/10.1111/j.1574-6976.2009.00204.x>.
 - Chimerel C, Field CM, Pinerio-Fernandez S, Keyser UF, Summers DK. 2012. Indole prevents *Escherichia coli* cell division by modulating membrane potential. *Biochim Biophys Acta* 1818:1590–1594. <https://doi.org/10.1016/j.bbame.2012.02.022>.
 - Kim J, Hong H, Heo A, Park W. 2013. Indole toxicity involves the inhibition of adenosine triphosphate production and protein folding in *Pseudomonas putida*. *FEMS Microbiol Lett* 343:89–99. <https://doi.org/10.1111/1574-6968.12135>.
 - Kim J, Park W. 2013. Indole inhibits bacterial quorum sensing signal transmission by interfering with quorum sensing regulator folding. *Microbiology* 159:2616–2625. <https://doi.org/10.1099/mic.0.070615-0>.
 - Fuchs G, Boll M, Heider J. 2011. Microbial degradation of aromatic compounds—from one strategy to four. *Nat Rev Microbiol* 9:803–816. <https://doi.org/10.1038/nrmicro2652>.
 - Sawyer DT. 1991. Oxygen chemistry. Oxford University Press, Inc., Oxford, United Kingdom.
 - Kovaleva EG, Lipscomb JD. 2008. Versatility of biological non-heme Fe(II) centers in oxygen activation reactions. *Nat Chem Biol* 4:186–193. <https://doi.org/10.1038/nchembio.71>.
 - Steiner RA, Janssen HJ, Roversi P, Oakley AJ, Fetzner S. 2010. Structural basis for cofactor-independent dioxygenation of *N*-heteroaromatic compounds at the α / β -hydrolase fold. *Proc Natl Acad Sci U S A* 107:657–662. <https://doi.org/10.1073/pnas.0909033107>.
 - Fetzner S, Steiner RA. 2010. Cofactor-independent oxidases and oxygenases. *Appl Microbiol Biotechnol* 86:791–804. <https://doi.org/10.1007/s00253-010-2455-0>.
 - Thierbach S, Bui N, Zapp J, Chhabra SR, Kappl R, Fetzner S. 2014. Substrate-assisted O₂ activation in a cofactor-independent dioxygenase. *Chem Biol* 21:217–225. <https://doi.org/10.1016/j.chembiol.2013.11.013>.
 - Hernandez-Ortega A, Quesne MG, Bui S, Heyes DJ, Steiner RA, Scrutton NA, de Visser SP. 2015. Catalytic mechanism of cofactor-free dioxygenases and how they circumvent spin-forbidden oxygenation of their substrates. *J Am Chem Soc* 137:7474–7487. <https://doi.org/10.1021/jacs.5b03836>.
 - Enslley BD, Ratzkin BJ, Osslund TD, Simon MJ, Wackett LP, Gibson DT. 1983. Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science* 222:167–169. <https://doi.org/10.1126/science.6353574>.
 - Lin GH, Chen HP, Shu HY. 2015. Detoxification of indole by an indole-induced flavoprotein oxygenase from *Acinetobacter baumannii*. *PLoS One* 10:e0138798. <https://doi.org/10.1371/journal.pone.0138798>.
 - Fetzner S. 1998. Bacterial degradation of pyridine, indole, quinoline, and their derivatives under different redox conditions. *Appl Microbiol Biotechnol* 49:237–250. <https://doi.org/10.1007/s002530051164>.
 - Arora PK, Sharma A, Bae H. 2015. Microbial degradation of indole and its derivatives. *J Chem* 2015:129159. <https://doi.org/10.1155/2015/129159>.
 - Yin B, Gu JD, Wan N. 2005. Degradation of indole by enrichment culture and *Pseudomonas aeruginosa* Gs isolated from mangrove sediment. *Int Biodeter Biodegr* 56:243–248. <https://doi.org/10.1016/j.ibiod.2005.10.001>.
 - Kim D, Rahman A, Sitepu IR, Hashidoko Y. 2014. Accelerated degradation of exogenous indole by *Burkholderia unamae* strain CK43B exposed to pyrogallol-type polyphenols. *Biosci Biotech Biochem* 77:1722–1727. <https://doi.org/10.1271/bbb.130282>.
 - Qu Y, Shen E, Ma Q, Zhang Z, Liu Z, Shen W, Wang J, Li D, Li H, Zhou J. 2015. Biodegradation of indole by a newly isolated *Cupriavidus* sp. SHE. *J Environ Sci (China)* 34:162–132.
 - Gu JD, Berry DF. 1991. Degradation of substituted indoles by an indole-degrading methanogenic consortium. *Appl Environ Microbiol* 57:2622–2627.
 - Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222–D226. <https://doi.org/10.1093/nar/gku1221>.
 - Bundy BM, Cambell AL, Neidle EL. 1998. Similarities between the *antABC*-encoded anthranilate dioxygenase and the *benABC*-encoded benzoate dioxygenase of *Acinetobacter* sp. strain ADP1. *J Bacteriol* 180:4466–4474.
 - Harwood CS, Parales RE. 1996. The beta-ketoadipate pathway and the biology of self-identity. *Annu Rev Microbiol* 50:553–590. <https://doi.org/10.1146/annurev.micro.50.1.553>.
 - Neidle EL, Hartnett C, Bonitz S, Ormston LN. 1988. DNA sequence of the *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase I structural gene *catA*: evidence for evolutionary divergence of intradiol dioxygenases by acquisition of DNA sequence repetitions. *J Bacteriol* 170:4874–4880. <https://doi.org/10.1128/jb.170.10.4874-4880.1988>.
 - Suzuki K, Ichimura A, Ogawa N, Hasebe A, Miyashita K. 2002. Differential expression of two catechol 1,2-dioxygenases in *Burkholderia* sp. strain TH2. *J Bacteriol* 184:5714–5722. <https://doi.org/10.1128/JB.184.20.5714-5722.2002>.
 - Cohn W, Crawford IP. 1976. Regulation of enzyme synthesis in the tryptophan pathway of *Acinetobacter calcoaceticus*. *J Bacteriol* 127:367–379.
 - Guan C, Ju J, Borlee BR, Williamson LL, Shen B, Raffa KF, Handelsman J. 2007. Signal mimics derived from a metagenomic analysis of the gypsy moth gut microbiota. *Appl Environ Microbiol* 73:3669–3676. <https://doi.org/10.1128/AEM.02617-06>.
 - Tischler D, Eulberg D, Lakner S, Kaschabek SR, van Berkel WJ, Schlömann M. 2009. Identification of a novel self-sufficient styrene monoxygenase from *Rhodococcus opacus* 1CP. *J Bacteriol* 191:4996–5009. <https://doi.org/10.1128/JB.00307-09>.
 - Tischler D, Gröning JAD, Kaschabek SR, Schlömann M. 2012. One-component styrene monoxygenase: an evolutionary view on a rare class of flavoproteins. *Appl Biochem Biotechnol* 167:931–944. <https://doi.org/10.1007/s12010-012-9659-y>.
 - Gröning JAD, Kaschabek SR, Schlömann M, Tischler D. 2014. A mechanistic study on SMOB-ADP1: an NADH:flavin oxidoreductase of the two-component styrene monoxygenase of *Acinetobacter baylyi* ADP1. *Arch Microbiol* 196:829–845. <https://doi.org/10.1007/s00203-014-1022-y>.
 - Han X, Wang W, Xiao X. 2008. Microbial biosynthesis and biotransformation of indigo and indigo-like pigments. *Sheng Wu Gong Cheng Xue Bao* 24:921–926. (In Chinese.) [https://doi.org/10.1016/S1872-2075\(08\)60043-6](https://doi.org/10.1016/S1872-2075(08)60043-6).
 - Fukuoka K, Tanaka K, Ozeki Y, Kanaly RA. 2015. Biotransformation of indole by *Cupriavidus* sp. KK10 proceeds through *N*-heterocyclic and carbocyclic-aromatic ring cleavage and production of indigoids. *Int Biodeter Biodegr* 97:13–24.
 - O'Connor KE, Dobson AD, Hartmans S. 1997. Indigo formation by microorganisms expressing styrene monoxygenase activity. *Appl Environ Microbiol* 63:4287–4291.
 - Boyd DR, Larkin MJ, Reid KA, Sharma ND, Wilson K. 1997. Metabolism of

- naphthalene, 1-naphthol, indene, and indole by *Rhodococcus* sp. strain NCIMB 12038. *Appl Environ Microbiol* 63:151–155.
44. Linhares M, Rebelo SLH, Simoes MMQ, Silva AMS, Neves MGPMS, Cavaleiro JAS, Freire C. 2014. Biomimetic oxidation of indole by Mn(III)porphyrins. *Appl Catal A Gen* 470:427–433. <https://doi.org/10.1016/j.apcata.2013.11.023>.
 45. Chaiyen P. 2010. Flavoenzymes catalyzing oxidative aromatic ring-cleavage reactions. *Arch Biochem Biophys* 493:62–70. <https://doi.org/10.1016/j.abb.2009.08.021>.
 46. Stanislaukiene R, Gasparaviciute R, Vaitekunas J, Meskiene R, Rutkiene R, Casaitė V, Meskys R. 2012. Construction of *Escherichia coli*-*Arthrobacter-Rhodococcus* shuttle vectors based on a cryptic plasmid from *Arthrobacter rhombi* and investigation of their application for functional screening. *FEMS Microbiol Lett* 327:78–86. <https://doi.org/10.1111/j.1574-6968.2011.02462.x>.
 47. Kutanovas S, Stankeviciute J, Urbelis G, Tauraitė D, Rutkiene R, Meskys R. 2013. Identification and characterization of tetramethylpyrazine catabolic pathway in *Rhodococcus jostii* TMP1. *Appl Environ Microbiol* 79:3649–3657. <https://doi.org/10.1128/AEM.00011-13>.
 48. Donoso R, Leiva-Novoa P, Zúñiga A, Timmermann T, Recabarren-Gajardo G, González B. 21 October 2016. Biochemical and genetic bases of indole-3-acetic acid (auxin phytohormone) degradation by the plant-growth-promoting rhizobacterium *Paraburkholderia phytofirmans* PsJN. *Appl Environ Microbiol* <https://doi.org/10.1128/AEM.01991-16>.
 49. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785–786. <https://doi.org/10.1038/nmeth.1701>.
 50. Fischer F, Kunne S, Fetzner S. 1999. Bacterial 2,4-dioxygenases: new members of the α/β hydrolase-fold superfamily of enzymes functionally related to serine hydrolases. *J Bacteriol* 181:5725–5733.
 51. Fujioka M, Wada H. 1968. The bacterial oxidation of indole. *Biochim Biophys Acta* 158:70–78. [https://doi.org/10.1016/0304-4165\(68\)90073-1](https://doi.org/10.1016/0304-4165(68)90073-1).
 52. Claus G, Kutzner HJ. 1983. Degradation of indole by *Alcaligenes* spec. *Syst Appl Microbiol* 4:169–180. [https://doi.org/10.1016/S0723-2020\(83\)80046-0](https://doi.org/10.1016/S0723-2020(83)80046-0).
 53. Lee J, Attila C, Cirillo SLG, Cirillo JD, Wood TK. 2009. Indole and 7-hydroxyindole diminish *Pseudomonas aeruginosa* virulence. *Microb Biotechnol* 2:75–90. <https://doi.org/10.1111/j.1751-7915.2008.00061.x>.
 54. Chu W, Zere TR, Weber MW, Wood TK, Whiteley M, Hidalgo-Romano B, Valenzuela E, McLean RJC. 2012. Indole production promotes *Escherichia coli* mixed-culture growth with *Pseudomonas aeruginosa* by inhibiting quorum signaling. *Appl Environ Microbiol* 78:411–419. <https://doi.org/10.1128/AEM.06396-11>.
 55. Hidalgo-Romano B, Gollihar J, Brown SA, Whiteley M, Valenzuela E, Jr, Kaplan HB, Wood TK, McLean RJC. 2014. Indole inhibition of *N*-acylated homoserine lactone-mediated quorum signaling is widespread in Gram-negative bacteria. *Microbiology* 160:2464–2473. <https://doi.org/10.1099/mic.0.081729-0>.
 56. Ahmer BM. 2004. Cell-to-cell signaling in *Escherichia coli* and *Salmonella enterica*. *Mol Microbiol* 52:933–945. <https://doi.org/10.1111/j.1365-2958.2004.04054.x>.
 57. Chan KG, Atkinson S, Mathee K, Sam C-K, Chhabra SR, Camara M, Koh C-L, Williams P. 2011. Characterization of *N*-acetylhomoserine lactone-degrading bacteria associated with the *Zingiber officinale* (ginger) rhizosphere: Coexistence of quorum quenching and quorum sensing in *Acinetobacter* and *Burkholderia*. *BMC Microbiol* 11:51. <https://doi.org/10.1186/1471-2180-11-51>.
 58. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
 59. Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol* 74:2461–2470. <https://doi.org/10.1128/AEM.02272-07>.
 60. Sambrook J. 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
 61. Vaitekunas J, Gasparaviciute R, Rutkiene R, Tauraitė D, Meskys R. 2015. A novel 2-hydroxypyridine catabolic pathway in *Rhodococcus rhodochrous* PY11. *Appl Environ Microbiol* 82:1264–1273. <https://doi.org/10.1128/AEM.02975-15>.
 62. De Melo JS, Moura AP, Melo MJ. 2004. Photophysical and spectroscopic studies of indigo derivatives in their keto and leuco forms. *J Phys Chem* 108:6975–6981. <https://doi.org/10.1021/jp049076y>.
 63. Bergonzini G, Melchiorre P. 2012. Dioxindole in asymmetric catalytic synthesis: routes to enantioenriched 3-substituted 3-hydroxyoxindoles and the preparation of marmecyn A. *Angew Chem Int Ed Engl* 51:971–974. <https://doi.org/10.1002/anie.201107443>.
 64. Medema MH, Takano E, Breitling R. 2013. Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* 30:1218–1223. <https://doi.org/10.1093/molbev/mst025>.
 65. Yang Y, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. The I-TASSER suite: protein structure and function prediction. *Nat Methods* 12:7–8. <https://doi.org/10.1038/nmeth.3213>.
 66. Pathak D, Ollis D. 1990. Refined structure of dienolactone hydrolase at 1.8 Å. *J Mol Biol* 214:497–525. [https://doi.org/10.1016/0022-2836\(90\)90196-5](https://doi.org/10.1016/0022-2836(90)90196-5).
 67. Hou J, Zheng H, Chruszcz M, Zimmerman MD, Shumilin IA, Osinski T, Demas M, Grimshaw S, Minor W. 2015. Dissecting the structural elements for the activation of β -ketoacyl-(acyl carrier protein) reductase from *Vibrio cholerae*. *J Bacteriol* 198:463–476. <https://doi.org/10.1128/JB.00360-15>.
 68. van den Berg B, Bhamidimarri SP, Winterhalter M. 2015. Crystal structure of a COG4313 outer membrane channel. *Sci Rep* 5:11927. <https://doi.org/10.1038/srep11927>.

Publication IV

Sadauskas, M., Statkevičiūtė, R., Vaitekūnas, J., Meškys, R.
Bioconversion of Biologically Active Indole Derivatives with Indole-3-
Acetic Acid-Degrading Enzymes from *Caballeronia glathei* DSM50014.
Biomolecules, 2020, 10(4):663
DOI: 10.3390/biom10040663

Article

Bioconversion of Biologically Active Indole Derivatives with Indole-3-Acetic Acid-Degrading Enzymes from *Caballeronia glathei* DSM50014

Mikas Sadauskas*, Roberta Statkevičiūtė, Justas Vaitekūnas and Rolandas Meškys

Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Life Sciences Center, Vilnius University, Saulėtekio al. 7, Vilnius, LT-10257, Lithuania; roberta.statkeviciute@gmc.stud.vu.lt (R.S.), justas.vaitekunas@bchi.vu.lt (J.V.), rolandas.meskys@bchi.vu.lt (R.M.)

* Correspondence: mikas.sadauskas@bchi.vu.lt; Tel.: +37062230055

Received: 26 March 2020; Accepted: 21 April 2020; Published: 24 April 2020

Abstract: A plant auxin hormone indole-3-acetic acid (IAA) can be assimilated by bacteria as an energy and carbon source, although no degradation has been reported for indole-3-propionic acid and indole-3-butyric acid. While significant efforts have been made to decipher the *Iac* (indole-3-acetic acid catabolism)-mediated IAA degradation pathway, a lot of questions remain regarding the mechanisms of individual reactions, involvement of specific *Iac* proteins, and the overall reaction scheme. This work was aimed at providing new experimental evidence regarding the biodegradation of IAA and its derivatives. Here, it was shown that *Caballeronia glathei* strain DSM50014 possesses a full *Iac* gene cluster and is able to use IAA as a sole source of carbon and energy. Next, *IacE* was shown to be responsible for the conversion of 2-oxoindole-3-acetic acid (Ox-IAA) intermediate into the central intermediate 3-hydroxy-2-oxindole-3-acetic acid (DOAA) without the requirement for *IacB*. During this reaction, the oxygen atom incorporated into Ox-IAA was derived from water. Finally, *IacA* and *IacE* were shown to convert a wide range of indole derivatives, including indole-3-propionic acid and indole-3-butyric acid, into corresponding DOAA homologs. This work provides novel insights into *Iac*-mediated IAA degradation and demonstrates the versatility and substrate scope of *IacA* and *IacE* enzymes.

Keywords: indole-3-acetic acid; indole-3-propionic acid; indole-3-butyric acid; biodegradation; *Caballeronia glathei*; bioconversion

1. Introduction

Indole and its derivatives comprise a group of biologically active *N*-heterocyclic compounds. Indole itself has recently been recognized as an interkingdom signaling molecule [1]. It is produced mainly by gut bacteria following the activity of tryptophanase [2], but can alter the physiology and metabolism of a very wide range of organisms, including the producer itself [3,4], other bacteria [5,6], eukaryotes [7], and even mammals [8,9]. The 3-substituted derivatives of indole bearing a carboxylic acid group are regarded as auxins, the plant growth-regulating hormones. These molecules can promote plant growth at basically all levels, including molecular, cellular, tissue, organ and whole plant levels [10]. The effect of auxins is concentration-dependent, requiring a strict regulation of auxin synthesis, degradation, conjugation, and import/export [11]. While indole-3-acetic acid (IAA or heteroauxin) is regarded as the most potent auxin with the strongest plant growth-regulating effects, other auxins that share similar structural scaffold were characterized as well, including indole-3-propionic acid (IPA) and indole-3-butyric acid (IBA) [12]. Surprisingly, the physiological effects of IPA, a product of bacterial tryptophan deamination, has recently been described not only in plants, but in the mammalian organisms as well (Figure 1). Increased concentration of IPA has been linked to lower risk of type 2 diabetes in humans [13,14]. Moreover,

IPA has been reported to provide beneficial effects for liver functions [15], perform as a biomarker for development of chronic kidney disease [16], and help to reduce the body weight in antibiotic-treated mice [17]. In addition, the activity of IPA against *M. tuberculosis* has been demonstrated [18] and later attributed to the suppression of tryptophan biosynthesis through the inhibition of anthranilate synthetase TrpE [19].

	Indole	IAA	IBA	IPA
Producers	Bacteria	Bacteria Plants	Plants Bacteria	Bacteria
Sources	Tryptophan	Tryptophan Anthranilate	Indole-3-acetic acid	Tryptophan
Physiological effects	Interkingdom signalling molecule	Promotion of plant growth	Precursor of IAA	Promotion of plant growth Antitubercular activity
Biodegraders	Bacteria	Bacteria	Unknown	Unknown
Biodegradation mechanism	Iif/Ind proteins Through anthranilate	Iac/Iad proteins Through catechol	Unknown	Unknown

Figure 1. Metabolism of biologically relevant derivatives of indole.

The aforementioned derivatives of indole-3-carboxylic acids can be present at large concentrations in bacteria-dominated niches, such as soil and gut, thus it would make sense for bacteria to use these compounds as carbon or energy sources. Indeed, bacterial degradation of indole and IAA has long been recognized [20,21]. Meanwhile no processes of biodegradation for IPA nor IBA has been reported so far (Figure 1). The bacterial degradation of IAA has been attributed to the activity of Iac proteins (indole acetic acid), encoded by *iac* gene cluster. Four bacterial strains with demonstrated aerobic IAA degradation capability and known genetic determinants for biodegradation—*Pseudomonas putida* 1290 [21], *Acinetobacter baumannii* ATCC19606 [22], *Paraburkholderia phytofirmans* PsJN [23], and *Enterobacter soli* LF7 [24]—were shown to follow the Iac-mediated biodegradation pathway. However, the information about biological cycle of IPA and IBA in different eco-niches is scarce. Numerous plants are able to convert the auxin precursor IBA into active IAA [25,26]. In particular, it has been suggested that, in *Arabidopsis*, a fatty acid beta-oxidation takes place in this process [27]. The fate of IPA is even less understood.

In the proposed mechanism of Iac-based biodegradation, IacA acts as an initial IAA oxygenase, producing 2-oxoindole-3-acetic acid (Ox-IAA) [22,28,29], which is then transformed to 3-hydroxy-2-oxindole-3-acetic acid (DOAA) by IacE, possibly involving IacB [23] as well. The end-product of Iac-mediated degradation is catechol [28], which is further oxidized by a catechol dioxygenase, the genes of which (*catABC*) are located in close proximity to the *iac* cluster [23,24,28]. Still, several questions remain unanswered in Iac-based biodegradation of IAA. While the end product of initial oxidation of IAA is presumed to be Ox-IAA, the most stable and easily identifiable end product of indole

oxidation reaction with IifC/IndA indole monooxygenases was indigo. The intermediate reaction product of IifC/IndA monooxygenases has long been elusive until a recent demonstration that indole-2,3-epoxide is the unstable intermediate [30]. Hence, the reaction product of IacA-catalyzed reaction and the reaction mechanism remain inconclusive. Also, IacE has been suggested to introduce another oxygen atom into Ox-IAA, therefore acting as an oxygenase. However, the amino acid sequence of IacE showed the highest sequence identity to short-chain dehydrogenases/reductases, which usually act as reductases [31] performing the NAD(P)(H)-dependent oxidoreduction of hydroxy/keto groups and usually does not incorporate oxygen. Finally, no intermediate has been proposed which could appear during the conversion of DOAA into catechol. Recently, an enzymatic decarboxylation of IAA to skatole by indoleacetate decarboxylase (Iad) has been identified [32] adding a new metabolic pathway to the IAA catabolism.

Caballeronia glathei, isolated as *Pseudomonas glathei* [33], later reclassified as *Burkholderia glathei* [34] and *Paraburkholderia glathei* [35], and finally as *C. glathei* [36], is a gram-negative bacterium and belongs to the class of β -Proteobacteria. *C. glathei* can inhabit different environments but is mainly found in soil. Although most Burkholderia-related microorganisms are relatively well-studied because of plant growth-promoting characteristics, biocontrol of plant diseases, or even being opportunistic pathogens for plants and humans [37], little is known about the ecological role of *C. glathei*. It has been demonstrated that *C. glathei* can establish a close relationship with soil-dominating fungi, which provides the bacterium with additional ecological fitness [38]. A recently published genome sequence of the type strain DSM50014 (GenBank RefSeq no. NZ_JFHC00000000.1) [39] contains a full set of iac genes, providing a framework for studying the functions of Iac proteins. Compared to other IAA-degrading strains that are publicly available, the iac locus in *C. glathei* DSM50014 is less interrupted by other genes (Figure 2) and comprises all iac genes that are known to date: iacA–I and iacR, iacS, iacT1, iacY. Therefore, the goal of this study was to analyze the initial steps of the catabolism of IAA in *C. glathei* and to offer additional data that would help to elucidate the Iac-mediated pathway of IAA biodegradation in detail. The second objective was to characterize a substrate scope (a range of converted/unconverted substrates) by the enzyme without specifying the substrate preference, which would require enzyme kinetics) of the IacA and IacE proteins

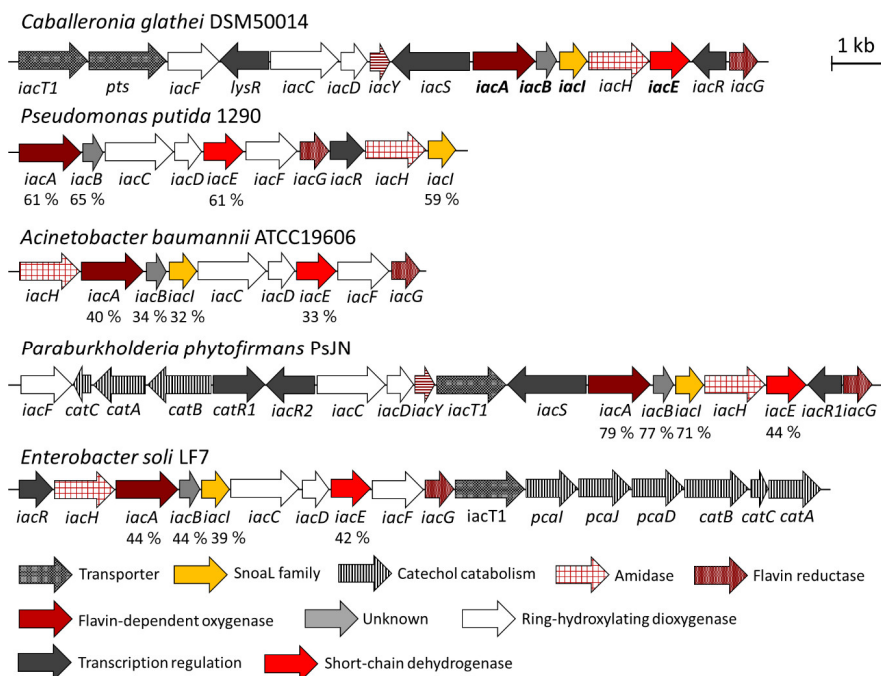


Figure 2. Distribution of *iac* and related genes in the genomes of known IAA-degrading microorganisms and *C. glathei* DSM50014. The genes of the enzymes that were studied in this work are highlighted in bold. Accession numbers of the proteins studied in this work are: *IacA*-WP_035925671, *IacB*-WP_035925668, *IacI*-WP_035925861, *IacE*-WP_035925663. The identity percentage between homologous genes that were studied in this report are also indicated.

2. Materials and Methods

2.1. Reagents, Bacterial Strains and Growth Conditions

All IAA derivatives used in this study (indole-3-acetic acid, 3-(2-hydroxyethyl)indole, ethyl-3-indole-acetate, 3-(3-hydroxypropyl)indole, 3-indoleacetonitrile, indole-3-acetamide, indole-3-butyric acid, DL-indole-3-lactic acid, indole-3-propionic acid, tryptamine, indole-3-acrylic acid and indole-3-carboxylic acid) and $H_2^{18}O$ were purchased from Sigma-Aldrich. All cloning and protein expression reagents were from ThermoFisher Scientific (Vilnius, Lithuania). All other reagents used in this study were of analytical or higher grade.

Caballeronia glathei strain DSM50014 was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ), Braunschweig, Germany. This strain was routinely cultivated in M1 medium (5 g L⁻¹ peptone, 3 g L⁻¹ meat extract, pH 7). For the IAA assimilation experiments, *C. glathei* DSM50014 was cultivated in M9 medium (3.5 g L⁻¹ Na₂HPO₄, 1.5 g L⁻¹ KH₂PO₄, 2.5 g L⁻¹ NaCl, 0.2 g L⁻¹ MgSO₄, 0.01 g L⁻¹ CaCl₂). *Escherichia coli* strains used in this study are listed in Supplementary Table S2. All *E. coli* strains were cultivated in LB medium. Ampicillin (50 µg mL⁻¹) and streptomycin (30 µg mL⁻¹) were added when necessary.

2.2. Cloning and Expression of *iac* Genes

Genomic DNA from *C. glathei* was extracted as described [40]. *iacA*, *iacE* and *iacB* genes were amplified from genomic DNA of *C. glathei* DSM50014 with oligonucleotides listed in Supplementary

Table S1. Cloning and expression of *iacA*, *iacB*, and *iacE* genes was performed following the general protocol described previously [41], except that the expression of IacE was induced with 0.1 mM IPTG and performed overnight at 16 °C. For protein co-expression, *iacA* was cloned to pET-28c(+) to obtain pET28-*iacA*. *iacB* and *iacE* were cloned to the MCS1 and MCS2, respectively, of pCDFDuet-1 to obtain pCDFDuet-*iacB* and pCDFDuet-*iacE*.

2.3. Whole-Cell Bioconversion

E. coli cells producing recombinant IacA, IacB and IacE proteins or their combinations were suspended in potassium phosphate buffer (10 mM, pH 7.7) supplemented with succinate (5 mM) to reach the 2× concentration of an initial culture and the IAA or derivative of IAA was added to a final concentration of 2 mM. Incubation of whole cells was performed with agitation (180 RPM, Innova44 Shaker, Eppendorf) at 30 °C overnight. Cells were removed by centrifugation at 16000× g for 5 min and the supernatant was subjected to HPLS/MS analysis.

2.4. Analytical Techniques

Substrate consumption and formation of products during the bioconversion experiments was analyzed spectrophotometrically by PowerWave XS plate reader (BioTek Instruments, Inc, Winooski, VT, USA) or the samples were mixed with an equal volume of acetonitrile and subjected to HPLC/MS analysis, which was performed as described [41]. Bioconversion efficiency of IacA was calculated as described [42], except that absorbance area at 280 nm was used rather than 254 nm. Three independent experiments were performed for each substrate and bioconversion efficiency of IacA is presented as mean ± SD.

IacE was purified by Ni-NTA affinity chromatography through C-terminal 6xHisTag as described in [41]. Enzymatic activity of purified IacE was monitored in reaction mixtures containing purified Ox-IAA, 1 mM of different cofactors (NAD⁺, NADP⁺, NADH or NADPH) and 1 µg of purified IacE. Reactions were incubated at 30 °C for different time intervals and analyzed with HPLC/MS as described above.

Purification of Ox-IAA and DOAA was performed by using reversed-phase preparative fast performance liquid chromatography essentially as described earlier [43]. Subsequently, ¹³C and ¹H NMR spectra of purified DOAA were recorded as described [42].

2-(3-hydroxy-2-oxoindolin-3-yl)acetic acid (DOAA). White solid. ¹H NMR (400 MHz, DMSO-d₆): δ = 1.89 (d, J = 14.9 Hz, 1H, CH), 2.28 (d, J = 14.9 Hz, 1H, CH), 6.76 (d, J = 7.6, 1H, CH), 6.89 (t, J = 7.5, 1H, CH), 7.14 (t, 1H, CH), 7.32 (d, J = 7.4, 1H, CH) 9.72 (s, 1H, CH) 10.12 (s, 1H, NH). ¹³C NMR (100 MHz, DMSO-d₆): δ = 42.5, 74.3, 109.7, 121.8, 124.3, 128.7, 134.8, 141.4, 174.2, 179.3.

2.5. Utilization of H₂¹⁸O

Following an overnight induction of expression or co-expression of genes *iacA*, *iacB* and *iacE* in *E. coli*, cells were washed with potassium phosphate buffer and suspended in H₂¹⁸O, containing 2 mM of IAA. For consumption of Ox-IAA, cells were suspended in H₂¹⁸O mixed with the solution of Ox-IAA (ratio 1:1). Bioconversion in H₂¹⁸O and analysis of reaction products was performed as described above.

3. Results

3.1. Identification of *Caballeronia glathei* DSM50014 as a biodegrader of IAA

Caballeronia glathei DSM50014 was obtained from the DSMZ collection bank and was identified to possess a full set of *iac* genes (sequence accession number NZ_JFHC01000015.1, nucleotide positions 99054–113486, Figure 2). First, this strain was tested for the capability to assimilate IAA. The growth was visible on M9 minimal medium agar plates supplemented with 1 mM IAA after the incubation period of five days (Supplementary Figure S1) confirming the ability of *C. glathei* DSM50014 to use IAA as a sole carbon and energy source.

In order to test whether the assimilation of IAA occurred through *Iac*-mediated pathway, whole cells of *C. glathei* were tested for the ability to consume DOAA, an intermediate in the *Iac*-mediated degradation pathway. IAA-induced *C. glathei* cells consumed both IAA and DOAA at the faster rate compared to the uninduced cells (Figure 3A,B), suggesting that both IAA degradation in *C. glathei* DSM50014 was an inducible process and DOAA was an intermediate compound during the assimilation. On the other hand, the DOAA counterpart in the indole degradation process, 3-hydroxyindolin-2-one, was not consumed by *C. glathei* DSM50014 as opposed to a natural indole-degrader *Acinetobacter* sp. strain O153 (Figure 3C). The absence of absorbance spectra that could indicate possible conversion products showed that *C. glathei* cells fully assimilated IAA and DOAA. Taken together, these results confirmed that *C. glathei* strain DSM50014 followed an *Iac*-mediated degradation of IAA.

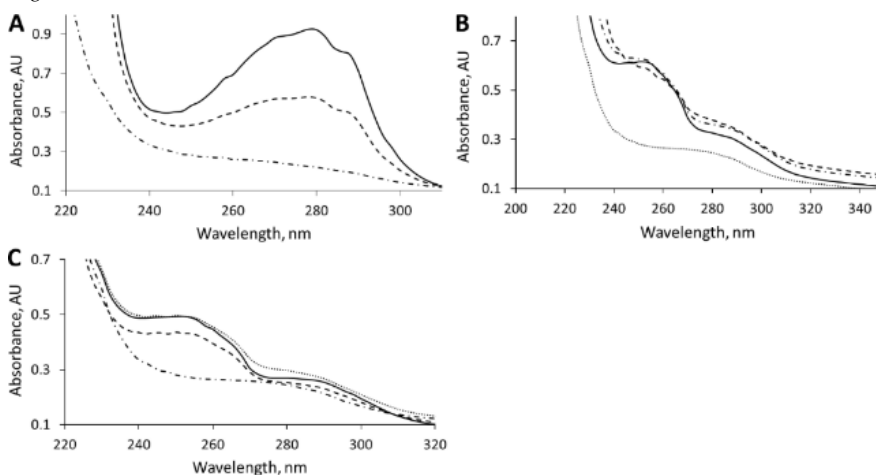


Figure 3. The whole-cell bioconversion of indole-3-acetic acid, 1 mM (A), 3-hydroxyindolin-2-one (B) and 3-hydroxy-2-oxindole-3-acetic acid (DOAA) (C). Bioconversion was performed at 30 °C overnight. Solid line—negative control (no cells), dashed line—*C. glathei* DSM50014, dash-dotted line—IAA (indole-3-acetic acid)-induced *C. glathei* DSM50014, dotted line—indole-induced *Acinetobacter* sp. O153.

3.2. *IacA*- and *IacE*-Catalyzed Reactions in IAA Degradation

To clarify the roles of individual *Iac* proteins, which were presumed to convert IAA into DOAA, *iacA*, *iacE*, and *iacB* genes were cloned to compatible plasmids and expressed in *E. coli* to obtain a bioconversion platform with different *Iac* protein combinations. *IacA* was annotated as a flavin-dependent acyl-CoA dehydrogenase (ACAD) family protein, *IacE* as a short-chain dehydrogenase/reductase, and no function prediction could be obtained from the sequence of *IacB* [23]. *IacG* was annotated as a flavin reductase and should provide a reduced flavin for the initial oxidation of IAA. However, as *E. coli* possesses numerous flavin reductases, which were shown to complement similar heterologous bioconversion reactions [41], *IacG* was omitted from the construction of IAA-converting *E. coli* strain. *IacA* has been demonstrated to oxidize IAA, but the exact reaction mechanism has not been elucidated, possibly due to instability of reaction intermediates as has been the case with biological indole oxidation [30]. Thus, the attention was focused on testing whether Ox-IAA was an intermediate compound in IAA biodegradation, and if so, which enzymes were required for the conversion of Ox-IAA into DOAA. Bioconversion of IAA by using *E. coli* cells expressing *IacAE* proteins resulted in accumulation of three compounds (Figure 4). The first compound with retention time of 5.2 min and molecular mass of 191 Da was also the major product during the conversion of IAA by *IacA*, assigning this peak as Ox-IAA. In both these reactions, a second product was also observed with a retention time of 5.5 min and molecular mass

of 382 (absorbance maxima at 241 nm and 293 nm), which could be a dimer of Ox-IAA, but the exact structure of this compound remained unidentified. The third compound with retention time of 4.6 min and molecular mass of 207 Da was only observed during the conversion of IAA by lacAE. It also possessed an absorbance spectrum resembling that of 3-hydroxyindolin-2-one, an intermediate compound (absorbance maxima at 254 nm and 291 nm) in indole biodegradation [41]. The latter product of IAA bioconversion was purified by using RP-FPLC. ^1H and ^{13}C NMR spectra (Supplementary Figure S19 and S20) corresponded to the ones published earlier [23], confirming that this compound was DOAA. *E. coli* with the combination of lacAB proteins did not produce DOAA, and reaction products profile was almost identical to that when using single lacA protein (Ox-IAA and the unidentified compound with 5.5 min retention time), except for a small peak at 6.9 min, which could not be identified. Also, no qualitative differences were observed between the bioconversion products of lacAE- and lacABE-carrying *E. coli* cells. The concentration of DOAA differed in these reactions, possibly because *E. coli* had to synthesize different sets of proteins. All these results confirmed that lacE catalyzed the second reaction after initial IAA oxygenation by lacA to obtain DOAA in *E. coli* system without the requirement of lacB.

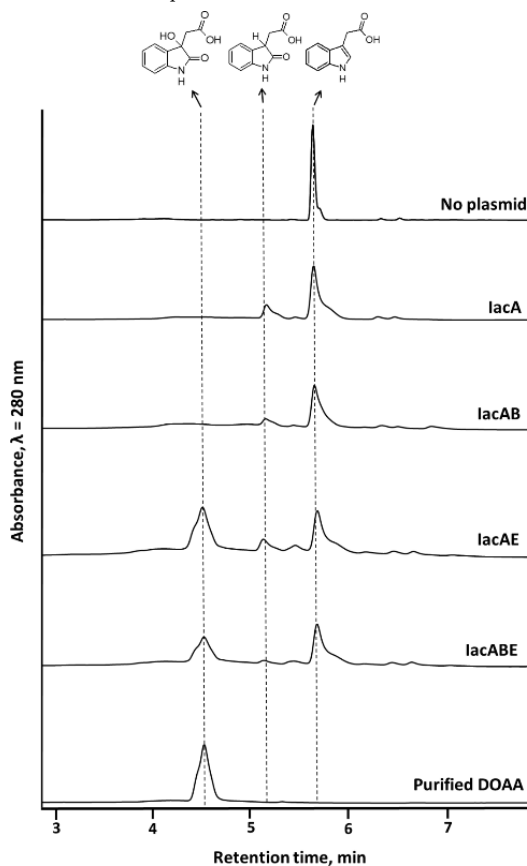


Figure 4. HPLC analysis of the bioconversion reaction mixture of IAA (1 mM) with *E. coli* whole-cells expressing different combination of lacA, lacB and lacE proteins. Formulas of the identified compounds are presented. Bioconversion reactions were carried out at 30 °C overnight.

The product of IacA-catalyzed IAA conversion, identified as Ox-IAA (retention time 5.3 min, molecular mass 191, absorbance maxima 241 nm and 293 nm) was also purified and tested as a substrate for both IacE- and IacB-expressing *E. coli*. Firstly, Ox-IAA was found to be unstable as it formed the product with unidentified structure as described above (retention time 5.5 min, molecular mass 382). In spite of that, Ox-IAA was converted to DOAA only in the presence of IacE (Figure 5) and the addition of IacB did not show any changes in the profile of the reaction of products. In addition, IAA-induced *C. glathei* DSM50014 cells were able to consume Ox-IAA completely (Figure 5), but did not consume the product with retention time 5.5 min and unknown structure, strongly suggesting that the latter compound can be a dead-end product. These results further supported the notion that IacB was not involved in the production of DOAA, and suggested that Ox-IAA was an intermediate compound in IAA biodegradation rather than a dead-end product.

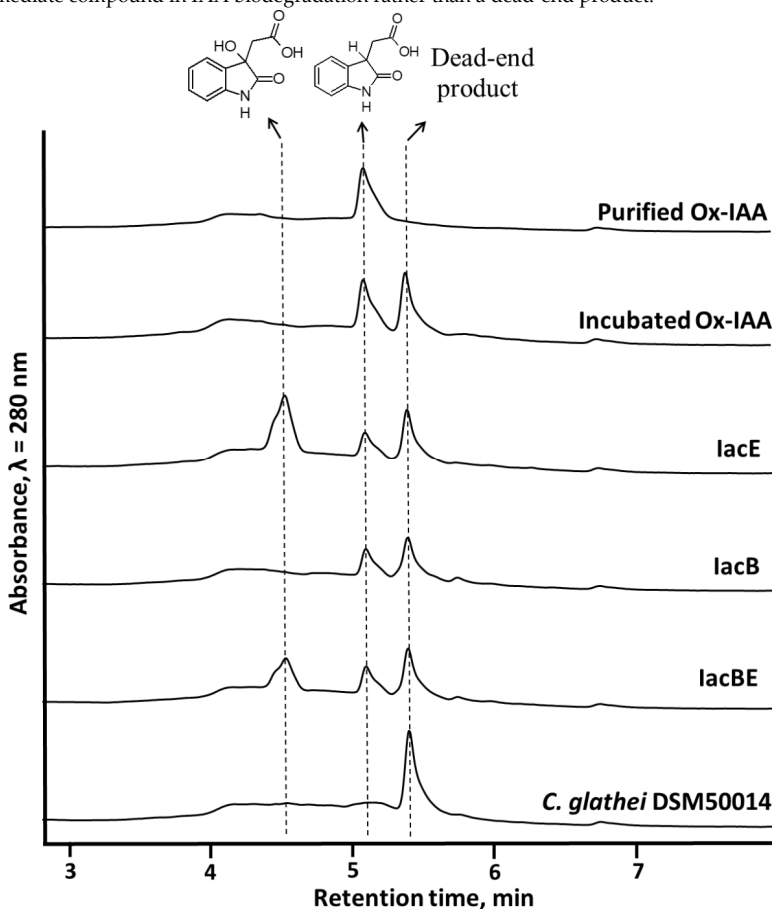


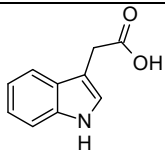
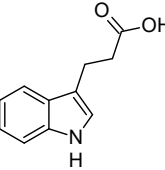
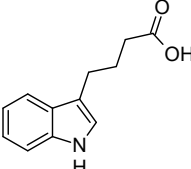

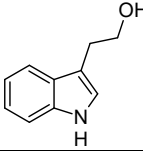
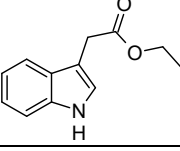
Figure 5. HPLC analysis of the bioconversion reaction mixture of 2-hydroxyindole-3-acetic acid (Ox-IAA) with *E. coli* cells expressing different combination of IacB and IacE proteins or *C. glathei* DSM50014. Bioconversion reactions were carried out at 30 °C overnight.

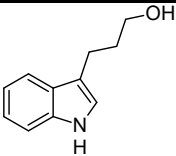
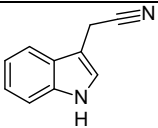
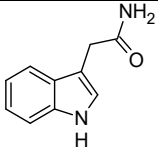
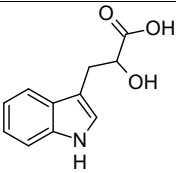
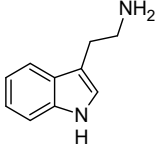
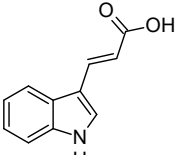
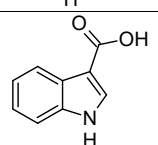
3.3. Substrate Scope of IacA and IacE Proteins

To test the substrate scope of IacA and IacE, different derivatives of indole harboring different groups at the third position were tested as substrates with *E. coli* cells co-expressing IacA and IacE.

Among 11 tested derivatives, tryptamine, indole-3-acrylic acid and indole-3-carboxylic acid were not oxidized by IacA. The rest of the compounds such as 3-(2-hydroxyethyl)indole, ethyl-3-indole-acetate, 3-(3-hydroxypropyl)indole, 3-indoleacetonitrile, indole-3-acetamide, indole-3-butyric acid, DL-indole-3-lactic acid and indole-3-propionic acid were oxidized to corresponding 2-oxo derivatives, albeit with different efficiency (Table 1, Supplementary Figure S5–S16). The two factors that govern the substrate scope of IacA could be the length of the carbon atom side chain, since the only substrate with less than two carbon atoms in the side chain (indole-3-carboxylic acid) was not oxidized, and the presence of the double bond in the side chain (indole-3-acrylic acid). Next, IacE was found to be capable of converting 3-(2-hydroxyethyl)indol-2-one, ethyl-(2-oxo-indol-3-yl)acetate, 3-(3-hydroxypropyl)indol-2-one, 2-oxo-3-indoleacetonitrile, 2-oxindole-3-butyric acid and 2-oxindole-3-propionic acid into corresponding DOAA homologs. Collectively, these results demonstrate a wide substrate scope of both IacA and IacE proteins.

Table 1. Substrate scope of IacA and IacE. Products of the IacA-catalyzed reaction—homologs of Ox-IAA, products of IacE-catalyzed reaction—homologs of DOAA. “+” indicates successful conversion, “–” indicates no conversion, NA—not analyzed. Bioconversion efficiency of IacA is presented as mean \pm SD.

Substrate	Structure	Activity	
		IacA (Bioconversion Efficiency, %)	IacE
IAA		+ (34 \pm 8)	+
IPA		+ (22 \pm 5)	+
IBA		+ (21 \pm 5)	+
Indole		+ (NA)	–
3-(2-hydroxyethyl)indole		+ (32 \pm 7)	+
Ethyl-3-indole-acetate		+ (19 \pm 6)	+

3-(3-hydroxypropyl)indole		+ (37 ± 8)	+
3-indoleacetonitrile		+ (>90)	+
Indole-3-acetamide		+ (28 ± 5)	-
Indole-3-lactic acid		+ (43 ± 11)	-
Tryptamine		-	NA
Indole-3-acrylic acid		-	NA
Indole-3-carboxylic acid		-	NA

3.4. Oxygen Incorporated into DOAA is Derived from Water

To gain insight into the reaction mechanism of IacAE-catalyzed IAA conversion into DOAA, H₂¹⁸O was used to trace the origin of oxygen atoms incorporated during IAA oxygenation. The molecular mass of Ox-IAA remained unchanged when *E. coli* cells expressing IacA were used for bioconversion of IAA in H₂¹⁸O environment (Figure 6A). However, DOAA with molecular mass of 209 was clearly observed in two cases: when *E. coli* cells expressing IacAE were used for bioconversion of IAA in H₂¹⁸O environment and with Ox-IAA and IacE-expressing *E. coli* cells in H₂¹⁸O environment (Figure 6B). Suspension of *E. coli* cells in H₂¹⁸O and solution of Ox-IAA in H₂¹⁶O were mixed in a volume ratio of 1:1 for the experiments described in Figure 6C, explaining the

presence of $[M+H]^+$ ions 208 and 210 with comparable intensities. Taken together, these findings demonstrated that the oxygen atom at C3 in DOAA had originated from water.

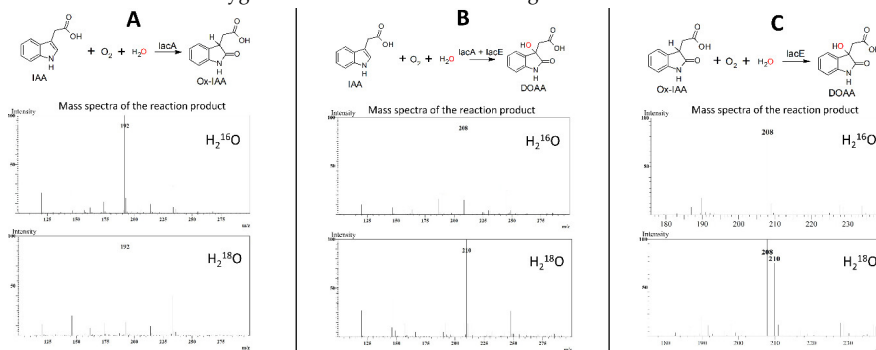


Figure 6. Mass spectra of Ox-IAA and DOAA obtained by using *E. coli* cells expressing IacA and IacE in H₂¹⁶O environment and H₂¹⁸O environment. (A)—Ox-IAA from IAA + IacA, (B)—DOAA from IAA + IacAE, (C)—DOAA from Ox-IAA + IacE. All mass spectra that are presented were recorded in positive ionization mode.

4. Discussion

4.1. Involvement and Role of IacB in IAA Biodegradation

Recently, *Cupriavidus pinatubonensis* JMP134 cells carrying recombinant IacB, IacE or IacBE proteins were tested for ability to consume the products of IAA conversion by IacA (most likely Ox-IAA), and strains carrying IacE and IacBE were confirmed to be able to produce DOAA, with the latter combination producing higher amounts of DOAA [23]. This led to the hypothesis that IacB might be an auxiliary or accessory protein for IacE [23]. In this report, similar amounts of DOAA were produced with IacE- and IacBE-expressing *E. coli* cells using Ox-IAA as substrate. Also, no DOAA production was observed with cells expressing IacB only, which was in agreement with the activity of IacB expressed in *C. pinatubonensis* JMP134. A conversion of Ox-IAA into DOAA with cells expressing IacE also dismissed the proposed second attack of IacA [23]. Also, bioconversion of IAA homologs into corresponding DOAA-like derivatives was achieved with IacAE proteins only, strongly suggesting that IacB does not participate in the first two reactions of IAA degradation. On the other hand, an endogenous protein with a similar function might be present in *E. coli* or *C. pinatubonensis*, complementing the absence of IacB in IacE-expressing cells. No proteins with significant sequence similarity to IacB could be detected in the genomes of *E. coli* and *C. pinatubonensis* by using blastp tool [44], and since no function can be deduced from the sequence of this protein, the role of IacB in IAA degradation remains unclear.

4.2. The Role of Other Iac Proteins and Analogies with Indole Biodegradation

A recently described mechanism for aerobic indole degradation [41,45] shows some similarities with an aerobic degradation of IAA. First of all, the two molecules share the same scaffold and also perform functions as signaling molecules, albeit in different organisms [1,10]. Most microorganisms with reported indole or IAA degradation capability belong to Proteobacteria, namely *Pseudomonas*, *Acinetobacter*, *Burkholderia* and related genera. The composition of genes involved in indole and IAA degradation (*iif* and *iac*, respectively) and their functions also appear to be similar. Both IifC and IacA are flavin-dependent oxygenases requiring a flavin reductase (IifD and IacG, respectively). However, these enzymes belong to different classes of flavin-dependent oxygenases: indole monooxygenase IifC is a member of the group E monooxygenases [30] while IacA possesses an ACAD fold and belongs to the group D of flavin monooxygenases [46]. Remarkably, IacA enzymes from IAA-degrading organisms form a separate branch from other group D flavin-dependent oxygenases

(Figure S22) and could possibly represent a new group of epoxidation-catalyzing enzymes in this group. It has been proposed that *lifC* converts indole to an unstable epoxide. Then, indole-2,3-epoxide rapidly hydrolyzes to a diol, which is a substrate for the *lifB* dehydrogenase [27,41]. In the absence of the dehydrogenase, the epoxide spontaneously loses the water molecule, and the formed 3-indoxyl dimerizes into indigo [41]. However, in the case of IAA and *lacA*, a mechanism of oxidation of the IAA is not clear. Based on experiments using $H_2^{18}O$, it can be proposed that: i) *lacA* forms epoxide (Figure 7, reaction I), the latter in a non-enzymatic acid-catalyzed process, due to which the nucleophile attacks the more substituted carbon because it is this carbon that holds a greater degree of positive charge, is hydrolyzed to a diol, which, after spontaneous dehydration, forms Ox-IAA (in the case of a basic epoxide ring opening Ox-IAA would contain ^{18}O originated from water, since the reaction occurs by an S_N2 mechanism, and the less substituted carbon is the site of nucleophilic attack); ii) the primary product of *lacA*-catalyzed reaction is 2-hydroxy-IAA (Figure 7, reaction II), which tautomerizes into Ox-IAA. Also, it should be stressed that it would be impossible to distinguish between the two pathways if $^{18}O_2$ is applied instead of atmospheric oxygen.

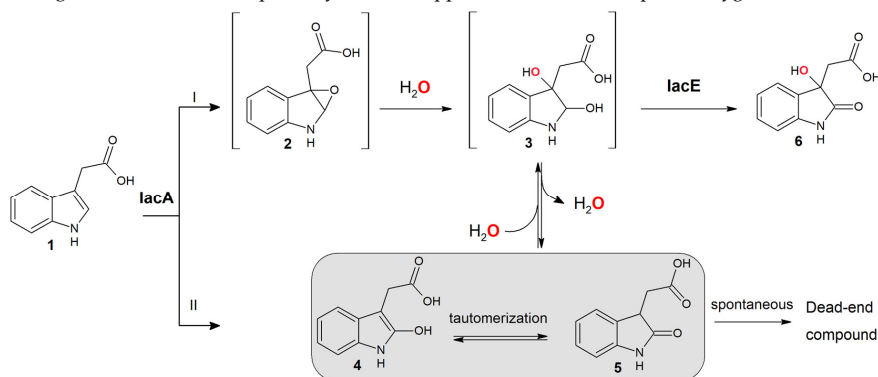


Figure 7. Proposed pathways of *lacAE*-mediated conversion of IAA into DOAA. 1—IAA, 2—2,3-epoxyIAA, 3—2,3-dihydroxyindoline-3-acetic acid, 4—2-hydroxyIAA, 5—Ox-IAA, 6—DOAA. ^{18}O atom is marked in red.

Comparing to catabolism of indole, a more complicated situation is observed at the next step of IAA degradation. The second reaction of indole and IAA biodegradation is catalyzed by a short chain dehydrogenase/reductase *lifB* and *lacE*, respectively. Since *lifB* performs the oxidation of 2,3-dihydroxyindoline during indole biodegradation [30,41], an analogous function might be hypothesized for *lacE* as well. In such a case, *lacA* produces an epoxide, similarly to *lifC*[30], then, after spontaneous hydrolysis, *lacE* oxidizes a hydroxy group at C2 position of the formed indoline derivative to yield DOAA (Figure 7). Hence, the observed Ox-IAA should be a dead-end product of the *lacA*-catalyzed reaction (as 3-indoxyl/indigo in the case of *lifC*). However, Ox-IAA was fully consumed by IAA-induced *C. glathei* cells as well as by *E. coli* cells carrying *lacE* protein or by a cell-free extract prepared from *E. coli* cells producing *lacE* protein and supplemented with NADH. However, all attempts to register a conversion of Ox-IAA in the presence of $NAD(P)^+$ or $NAD(P)H$ by the purified recombinant *lacE* protein have been unsuccessful. In addition, a spontaneous transformation of Ox-IAA to the unidentified product (Figure 5, retention time 5.4 min) is observed even in the absence of any enzymes. A similar process takes place with all Ox-IAA homologs containing a carboxylic group (i.e., 2-hydroxyindole-3-propionic acid and 2-hydroxyindole-3-butyric acid), suggesting an instability of such compounds under experimental conditions. However, the most remarkable feature of the *lacE*-mediated reaction is that the oxygen atom from water has been introduced into the product – DOAA (Figure 6).

lacE possesses a Rossmann fold domain, strongly indicating a redox function of this enzyme. Also, according to the SDRED database [47], *lacE* belongs to the HFAM1 family of the classical short-chain dehydrogenases/reductases. All major structural motifs of this family can be identified in the

sequence of IacE (Supplementary Figure S21): four active site residues (Asn111, Ser139, Tyr153, and Lys157), NNAG motif, stabilizing the central β -sheet (Asn96, Asn97, Ala98, Gly99), a glycine-rich motif for the binding of NAD(P)H (9ThrGlyAlaAlaArgGlyLeuGly16) and PG motif (Pro193, Gly194). IacE also clusters with other NAD(P)-dependent short chain dehydrogenases/reductases in a phylogenetic tree (Supplementary Figure S23). Thus, it is unlikely that IacE uses H₂O directly to produce DOAA. A more plausible mechanism would be the spontaneous addition of water to Ox-IAA forming a diol derivative, which could then be oxidized by IacE to produce a stable DOAA (Figure 7).

A further conversion of DOAA in *C. glathei* cells might be related to indole catabolism where the conversion of 3-hydroxyindolin-2-one—a structural homolog of DOAA—is performed by IifA, a putative cofactor-independent oxygenase, composed of two domains [41]. Interestingly, both IacI from the proposed IAA degradation pathway in *C. glathei* and the C-terminal domain of IifA contain a SnoaL-fold domain. Therefore, it was hypothesized that IacI, possibly together with IacB, could perform the consecutive conversion of DOAA. However, neither bioconversion with recombinant *E. coli* cells co-expressing IacBI (Supplementary Figure S16) nor in vitro conversion with soluble fractions of *E. coli* lysates containing IacB and IacI (Supplementary Figure S17) resulted in the transformation of DOAA. One explanation might be that *E. coli* is not a suitable host for obtaining active Iac proteins other than IacAE, since all biotransformation reactions of DOAA were achieved with recombinant producers closely related to the original IAA-degrading organism [23,24,28]. So far, the metabolic gap between DOAA and catechol during IAA biodegradation as well as functions of other Iac proteins remains unfilled.

4.3. Perspectives of the Bioconversion of IAA homologs

The growth of IAA-degrading microorganisms, namely *P. putida* 1290, has been tested on IAA derivatives (IPA, IBA, indole-3-acetaldehyde, indole-3-acrylic acid, indole-3-lactic acid, indole-3-pyruvic acid, naphthalene acetic acid) but no growth was observed, except for indole-3-acetaldehyde [21]. In spite of that, this report describes the conversion of some of these compounds to DOAA-like derivatives by using an *E. coli*-based bioconversion platform carrying IacAE protein combination. This implies that either IacAE proteins from different organisms have different substrate scope, or other Iac proteins of the degradation pathway do not accept DOAA-like derivatives as substrates thus preventing the complete assimilation of IAA homologs for growth. Furthermore, IacA was found to be able to oxidize eight out of 11 tested IAA derivatives, while IacE further converted 6 out of 8 Ox-IAA-like derivatives into DOAA homologs, suggesting a wide substrate scope, which could only be limited by the availability of substrates. It should be stressed that 2,3-dihydroxyindoline-3-acetic acid (Figures 3 and 7) has two chiral centers at C2 and C3, hence this compound can exist in various enantiomeric and diastereomeric forms. To determine which one is a true substrate of IacE and which is an absolute stereostructure of the DOAA, additional studies have to be carried out.

Both IPA and IBA are present in different environments. Yet, no organisms or enzymes with IPA- or IBA-degradation capability have been reported. By showing that the IacAE enzyme system can convert IPA and IBA to corresponding DOAA-like derivatives, this report suggests that IacAE could be a central part of such hypothetical pathway. From the practical point of view, these enzymes are attractive for the engineering of the artificial metabolic cascades converting 3-substituted carboxy derivatives of indole. Such products could be of special importance since *N*-heterocyclic compounds with DOAA scaffold exhibit important biological activities. For example, convolutamydines (4,6-dibromo-3-hydroxyoxindoles), specifically a convolutamydine A, promoted the development of normal cell characteristics in a tumor cell line HL-60 [48]. DOAA is also a building block of plethora of biologically-active compounds: proteasome inhibitor TMC-95 [49], inhibitors of tubulin polymerization celogentins[50], medicinal plant-derived alkaloids paratunamides [51], and arundaphine [52]. Also, as IPA was shown to possess antitubercular activity [18], it could represent a perfect scaffold for a lead compound during the target-based drug optimization [53] and IacA or IacAE systems could be useful for the development of such compounds with improved antitubercular activity. Some of these compounds have been chemically synthesized, usually via enol

reactions [54], which often require the use of organic solvents, acids, and extreme conditions. Enzymatic synthesis of DOAA derivatives by using IacAE offers several advantages, including mild reaction conditions and, probably, enantioselectivity.

5. Conclusions

In this report, the ability of *Caballeronia glathei* strain DSM50014 to use indole-3-acetic acid as a sole carbon and energy source was demonstrated. The first enzyme in this process, a flavin-dependent oxygenase IacA was shown to convert IAA into 2-hydroxyindole-3-acetic acid without the incorporation of oxygen from H₂¹⁸O, supporting the role of IacA as a hydroxylase rather than an epoxidase. The conversion of Ox-IAA into the central intermediate 3-hydroxy-2-oxindole-3-acetic acid was shown to be catalyzed by IacE and without the requirement for IacB. Furthermore, the oxygen atom incorporated into DOAA was derived from water. Also, IacA and IacE were shown to convert a wide range of indole derivatives, including biologically active compounds indole-3-propionic acid and indole-3-butyric acid, into corresponding DOAA-like derivatives.

Supplementary Materials: The following are available online at www.mdpi.com/2218-273X/10/4/663/s1, Table S1: Oligonucleotides used in this study, Table S2: Characteristics of bacterial strains used in this study, Figure S1: Growth of *Caballeronia glathei* DSM50014 on M9 minimal medium, Figure S2: SDS-PAGE of IacA and IacE co-expression in *E. coli*, Figure S3: SDS-PAGE of IacB expressed in *E. coli*, Figure S4: SDS-PAGE of IacB and IacE expressed in *E. coli*, Figure S5: HPLC chromatograms of bioconversion products, Figure S6: HPLC chromatograms of bioconversion products, Figure S7: HPLC chromatograms of bioconversion products, Figure S8: HPLC chromatograms of bioconversion products, Figure S9: HPLC chromatograms of bioconversion products, Figure S10: HPLC chromatograms of bioconversion products, Figure S11: HPLC chromatograms of bioconversion products, Figure S12: HPLC chromatograms of bioconversion products, Figure S13: HPLC chromatograms of bioconversion products, Figure S14: HPLC chromatograms of bioconversion products, Figure S15: HPLC chromatograms of bioconversion products, Figure S16: HPLC chromatograms of bioconversion products, Figure S17: HPLC chromatograms of bioconversion products, Figure S18: HPLC chromatograms of bioconversion products, Figure S19: ¹³C NMR spectrum of DOAA, Figure S20: ¹H NMR spectrum of DOAA, Figure S21: conservative sequence motifs in the sequence of IacE. Figure S22: Phylogenetic tree of IacA. Amino acid sequences of group D flavin-dependent oxygenase were picked according to [55]. Sequences were aligned by using the ClustalW algorithm, a maximum-likelihood tree was constructed by using MEGA7 software [56] with 1000 bootstrap replications. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test are shown. Figure S23: Phylogenetic tree of IacE. Amino acid sequences of short-chain dehydrogenases/reductases (SDR) were picked from each SDR family according to [57]. Sequences were aligned by using the ClustalW algorithm, a maximum-likelihood tree was constructed by using MEGA7 software [56] with 1000 bootstrap replications.

Author Contributions: Conceptualization, M.S., J.V. and R.M.; Methodology, M.S., J.V. and R.M.; Investigation M.S., R.S., J.V.; Writing – Original Draft Preparation, M.S.; Writing–Review & Editing, M.S., R.S., J.V. and R.M.; Visualization, M.S., R.S., J.V. and R.M.; Supervision, R.M.; Funding Acquisition, R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Social Fund under the No 09.3.3-LMT-K-712 “Development of Competences of Scientists, other Researchers and Students through Practical Research Activities” measure, Grant No. 09.3.3.-LMT-K-712-16-0259. This work was partially supported by research grant (agreement No. S-SEN-20-9) from Research Council of Lithuania.

Acknowledgments: We thank Rita Bukšnaitienė (Department of Organic Chemistry, Faculty of Chemistry and Geosciences, Vilnius University) and Daiva Tauraitė for assistance with NMR analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, J.-H.; Wood, T.K.; Lee, J. Roles of Indole as an Interspecies and Interkingdom Signaling Molecule. *Trends in Microbiology* **2015**, *23*, 707–718, doi:10.1016/j.tim.2015.08.001.
2. Li, G.; Young, K.D. Indole production by the tryptophanase TnaA in *Escherichia coli* is determined by the amount of exogenous tryptophan. *Microbiology* **2013**, *159*, 402–410, doi:10.1099/mic.0.064139-0.

3. Kim, J.; Park, W. Indole: a signaling molecule or a mere metabolic byproduct that alters bacterial physiology at a high concentration? *J. Microbiol.* **2015**, *53*, 421–428, doi:10.1007/s12275-015-5273-3.
4. Zarkan, A.; Caño-Muñiz, S.; Zhu, J.; Nahas, K.A.; Cama, J.; Keyser, U.F.; Summers, D.K. Indole Pulse Signalling Regulates the Cytoplasmic pH of *E. coli* in a Memory-Like Manner. *Sci. Rep.* **2019**, *9*, 1–10, doi:10.1038/s41598-019-40560-3.
5. Nikaido, E.; Giraud, E.; Baucheron, S.; Yamasaki, S.; Wiedemann, A.; Okamoto, K.; Takagi, T.; Yamaguchi, A.; Cloeckert, A.; Nishino, K. Effects of indole on drug resistance and virulence of *Salmonella enterica* serovar Typhimurium revealed by genome-wide analyses. *Gut Pathog.* **2012**, *4*, 5, doi:10.1186/1757-4749-4-5.
6. Kumar, A.; Sperandio, V. Indole Signaling at the Host-Microbiota-Pathogen Interface. *mBio* **2019**, *10*, doi:10.1128/mBio.01031-19.
7. Lee, J.-H.; Kim, Y.-G.; Kim, M.; Kim, E.; Choi, H.; Kim, Y.; Lee, J. Indole-associated predator-prey interactions between the nematode *Caenorhabditis elegans* and bacteria. *Environ. Microbiol.* **2017**, *19*, 1776–1790, doi:10.1111/1462-2920.13649.
8. Huc, T.; Konop, M.; Onyszkiewicz, M.; Podsadni, P.; Szczepańska, A.; Turlo, J.; Ufnal, M. Colonic indole, gut bacteria metabolite of tryptophan, increases portal blood pressure in rats. *Am. J. Physiol. Regul., Integr. Comp. Physiol.* **2018**, *315*, R646–R655, doi:10.1152/ajpregu.00111.2018.
9. Jaglin, M.; Rhimi, M.; Philippe, C.; Pons, N.; Bruneau, A.; Goustard, B.; Daugé, V.; Maguin, E.; Naudon, L.; Rabot, S. Indole, a Signaling Molecule Produced by the Gut Microbiota, Negatively Impacts Emotional Behaviors in Rats. *Front Neurosci.* **2018**, *12*, doi:10.3389/fnins.2018.00216.
10. Leyser, O. Auxin Signaling. *Plant Physiol.* **2018**, *176*, 465–479, doi:10.1104/pp.17.00765.
11. Woodward, A.W.; Bartel, B. Auxin: Regulation, Action, and Interaction. *Ann. Bot.* **2005**, *95*, 707–735, doi:10.1093/aob/mci083.
12. Enders, T.A.; Strader, L.C. Auxin activity: Past, present, and future. *Am. J. Bot.* **2015**, *102*, 180–196, doi:10.3732/ajb.1400285.
13. Tuomainen, M.; Lindström, J.; Lehtonen, M.; Auriola, S.; Pihlajamäki, J.; Peltonen, M.; Tuomilehto, J.; Uusitupa, M.; Mello, V.D. de; Hanhineva, K. Associations of serum indolepropionic acid, a gut microbiota metabolite, with type 2 diabetes and low-grade inflammation in high-risk individuals. *Nutr. Diabetes* **2018**, *8*, 1–5, doi:10.1038/s41387-018-0046-9.
14. de Mello, V.D.; Paananen, J.; Lindström, J.; Lankinen, M.A.; Shi, L.; Kuusisto, J.; Pihlajamäki, J.; Auriola, S.; Lehtonen, M.; Rolandsson, O.; et al. Indolepropionic acid and novel lipid metabolites are associated with a lower risk of type 2 diabetes in the Finnish Diabetes Prevention Study. *Sci. Rep.* **2017**, *7*, 46337, doi:10.1038/srep46337.
15. Zhao, Z.-H.; Xin, F.-Z.; Xue, Y.; Hu, Z.; Han, Y.; Ma, F.; Zhou, D.; Liu, X.-L.; Cui, A.; Liu, Z.; et al. Indole-3-propionic acid inhibits gut dysbiosis and endotoxin leakage to attenuate steatohepatitis in rats. *Exp. Mol. Med.* **2019**, *51*, 1–14, doi:10.1038/s12276-019-0304-5.
16. Sun, C.-Y.; Lin, C.-J.; Pan, H.-C.; Lee, C.-C.; Lu, S.-C.; Hsieh, Y.-T.; Huang, S.-Y.; Huang, H.-Y. Clinical association between the metabolite of healthy gut microbiota, 3-indolepropionic acid and chronic kidney disease. *Clinical Nutrition* **2019**, *38*, 2945–2948, doi:10.1016/j.clnu.2018.11.029.
17. Konopelski, P.; Konop, M.; Gawrys-Kopczynska, M.; Podsadni, P.; Szczepanska, A.; Ufnal, M. Indole-3-Propionic Acid, a Tryptophan-Derived Bacterial Metabolite, Reduces Weight Gain in Rats. *Nutrients* **2019**, *11*, 591, doi:10.3390/nu11030591.
18. Negatu, D.A.; Liu, J.J.; Zimmerman, M.; Kaya, F.; Dartois, V.; Aldrich, C.C.; Gengenbacher, M.; Dick, T. Whole-Cell Screen of Fragment Library Identifies Gut Microbiota Metabolite Indole Propionic Acid as Antitubercular. *Antimicrob. Agents Chemother.* **2018**, *62*, doi:10.1128/AAC.01571-17.
19. Negatu, D.A.; Yamada, Y.; Xi, Y.; Go, M.L.; Zimmerman, M.; Ganapathy, U.; Dartois, V.; Gengenbacher, M.; Dick, T. Gut Microbiota Metabolite Indole Propionic Acid Targets Tryptophan Biosynthesis in *Mycobacterium tuberculosis*. *mBio* **2019**, *10*, doi:10.1128/mBio.02781-18.
20. Ma, Q.; Zhang, X.; Qu, Y. Biodegradation and Biotransformation of Indole: Advances and Perspectives. *Front Microbiol.* **2018**, *9*, doi:10.3389/fmicb.2018.02625.
21. Leveau, J.H.J.; Lindow, S.E. Utilization of the Plant Hormone Indole-3-Acetic Acid for Growth by *Pseudomonas putida* Strain 1290. *Appl. Environ. Microbiol.* **2005**, *71*, 2365–2371, doi:10.1128/AEM.71.5.2365-2371.2005.

22. Lin, G.-H.; Chen, H.-P.; Huang, J.-H.; Liu, T.-T.; Lin, T.-K.; Wang, S.-J.; Tseng, C.-H.; Shu, H.-Y. Identification and characterization of an indigo-producing oxygenase involved in indole 3-acetic acid utilization by *Acinetobacter baumannii*. *Antonie van Leeuwenhoek* **2012**, *101*, 881–890, doi:10.1007/s10482-012-9704-4.
23. Donoso, R.; Leiva-Novoa, P.; Zúñiga, A.; Timmermann, T.; Recabarren-Gajardo, G.; González, B. Biochemical and Genetic Bases of Indole-3-Acetic Acid (Auxin Phytohormone) Degradation by the Plant-Growth-Promoting Rhizobacterium *Paraburkholderia phytofirmans* PsJN. *Appl. Environ. Microbiol.* **2016**, *83*, doi:10.1128/AEM.01991-16.
24. Greenhut, I.V.; Slezak, B.L.; Leveau, J.H.J. iac Gene Expression in the Indole-3-Acetic Acid-Degrading Soil Bacterium *Enterobacter soli* LF7. *Appl. Environ. Microbiol.* **2018**, *84*, doi:10.1128/AEM.01057-18.
25. Epstein, E.; Ludwig-Müller, J. Indole-3-butyric acid in plants: occurrence, synthesis, metabolism and transport. *Physiologia Plantarum* **1993**, *88*, 382–389, doi:10.1111/j.1399-3054.1993.tb05513.x.
26. Strader, L.C.; Bartel, B. Transport and metabolism of the endogenous auxin precursor indole-3-butyric acid. *Mol. Plant.* **2011**, *4*, 477–486, doi:10.1093/mp/ssf006.
27. Zolman, B.K.; Yoder, A.; Bartel, B. Genetic analysis of indole-3-butyric acid responses in *Arabidopsis thaliana* reveals four mutant classes. *Genetics* **2000**, *156*, 1323–1337.
28. Leveau, J.H.J.; Gerards, S. Discovery of a bacterial gene cluster for catabolism of the plant hormone indole 3-acetic acid. *FEMS Microbiol. Ecol.* **2008**, *65*, 238–250, doi:10.1111/j.1574-6941.2008.00436.x.
29. Scott, J.C.; Greenhut, I.V.; Leveau, J.H.J. Functional Characterization of the Bacterial iac Genes for Degradation of the Plant Hormone Indole-3-Acetic Acid. *J. Chem. Ecol.* **2013**, *39*, 942–951, doi:10.1007/s10886-013-0324-x.
30. Heine, T.; Großmann, C.; Hofmann, S.; Tischler, D. Indigoid dyes by group E monooxygenases: mechanism and biocatalysis. *Biol. Chem.* **2019**, *400*, 939–950, doi:10.1515/hsz-2019-0109.
31. Kavanagh, K.L.; Jörnvall, H.; Persson, B.; Oppermann, U. Medium- and short-chain dehydrogenase/reductase gene and protein families. *Cell Mol. Life Sci.* **2008**, *65*, 3895–3906, doi:10.1007/s00018-008-8588-y.
32. Liu, D.; Wei, Y.; Liu, X.; Zhou, Y.; Jiang, L.; Yin, J.; Wang, F.; Hu, Y.; Urs, A.N.N.; Liu, Y.; et al. Indoleacetate decarboxylase is a glyceryl radical enzyme catalysing the formation of malodorant skatole. *Nat. Commun.* **2018**, *9*, 1–8, doi:10.1038/s41467-018-06627-x.
33. Zolg, W.; Ottow, J.C. *Pseudomonas glathei* sp. nov., a new nitrogen-scavenging rod isolated from acid lateritic relicts in Germany. *J. Comp. Neurol.* **1975**, *164*, 287–299.
34. Viillard, V.; Poirier, I.; Cournoyer, B.; Haurat, J.; Wiebkin, S.; Ophel-Keller, K.; Balandreau, J. *Burkholderia graminis* sp. nov., a rhizospheric *Burkholderia* species, and reassessment of [*Pseudomonas*] *phenazinium*, [*Pseudomonas*] *pyrocinia* and [*Pseudomonas*] *glathei* as *Burkholderia*. *Int. J. Syst. Bacteriol.* **1998**, *48*, 549–563, doi:10.1099/00207713-48-2-549.
35. Sawana, A.; Adeolu, M.; Gupta, R.S. Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring environmental species. *Front Genet.* **2014**, *5*, doi:10.3389/fgene.2014.00429.
36. Dobritsa, A.P.; Samadpour, M. Transfer of eleven species of the genus *Burkholderia* to the genus *Paraburkholderia* and proposal of *Caballeronia* gen. nov. to accommodate twelve species of the genera *Burkholderia* and *Paraburkholderia*. *Int. J. Syst. Evol. Microbiol.* **2016**, *66*, 2836–2846, doi:10.1099/ijsem.0.001065.
37. Mahenthalingam, E.; Baldwin, A.; Dowson, C.G. *Burkholderia cepacia* complex bacteria: opportunistic pathogens with important natural biology. *J. Appl. Microbiol.* **2008**, *104*, 1539–1551, doi:10.1111/j.1365-2672.2007.03706.x.
38. Stopnisek, N.; Zühlke, D.; Carlier, A.; Barberán, A.; Fierer, N.; Becher, D.; Riedel, K.; Eberl, L.; Weisskopf, L. Molecular mechanisms underlying the close association between soil *Burkholderia* and fungi. *ISME J.* **2016**, *10*, 253–264, doi:10.1038/ismej.2015.73.
39. Liu, X.-Y.; Li, C.-X.; Luo, X.-J.; Lai, Q.-L.; Xu, J.-H. *Burkholderia jiangsuensis* sp. nov., a methyl parathion degrading bacterium, isolated from methyl parathion contaminated soil. *Int. J. Syst. Evol. Microbiol.* **2014**, *64*, 3247–3253, doi:10.1099/ijse.0.064444-0.
40. Sambrook, J. *Molecular cloning: a laboratory manual*; Third edition. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 2001.

41. Sadauskas, M.; Vaitekūnas, J.; Gasparavičiūtė, R.; Meškys, R. Genetic and Biochemical Characterization of Indole Biodegradation in *Acinetobacter* sp. Strain O153. *Appl. Environ. Microbiol.* **2017**, *AEM.01453-17*, doi:10.1128/AEM.01453-17.
42. Petkevičius, V.; Vaitekūnas, J.; Tauraitė, D.; Stankevičiūtė, J.; Šarlauskas, J.; Čėnas, N.; Meškys, R. A Biocatalytic Synthesis of Heteroaromatic N-Oxides by Whole Cells of *Escherichia coli* Expressing the Multicomponent, Soluble Di-Iron Monooxygenase (SDIMO) PmlABCDEF. *Adv. Synth. Catal.* **2019**, *361*, 2456–2465, doi:10.1002/adsc.201801491.
43. Sadauskas, M.; Statkevičiūtė, R.; Vaitekūnas, J.; Petkevičius, V.; Časaitė, V.; Gasparavičiūtė, R.; Meškys, R. Enzymatic synthesis of novel water-soluble indigoid compounds. *Dyes Pigments* **2020**, *173*, 107882, doi:10.1016/j.dyepig.2019.107882.
44. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410, doi:10.1016/S0022-2836(05)80360-2.
45. Qu, Y.; Ma, Q.; Liu, Z.; Wang, W.; Tang, H.; Zhou, J.; Xu, P. Unveiling the biotransformation mechanism of indole in a *Cupriavidus* sp. strain. *Mol. Microbiol.* **2017**, *106*, 905–918, doi:10.1111/mmi.13852.
46. Huijbers, M.M.E.; Montersino, S.; Westphal, A.H.; Tischler, D.; van Berkel, W.J.H. Flavin dependent monooxygenases. *Arch. Biochem. Biophys.* **2014**, *544*, 2–17, doi:10.1016/j.abb.2013.12.005.
47. Gräff, M.; Buchholz, P.C.F.; Stockinger, P.; Bommarius, B.; Bommarius, A.S.; Pleiss, J. The Short-chain Dehydrogenase/Reductase Engineering Database (SDRED): A classification and analysis system for a highly diverse enzyme family. *Proteins* **2019**, *87*, 443–451, doi:10.1002/prot.25666.
48. Kamano, Y.; Zhang, H.; Ichihara, Y.; Kizu, H.; Komiyama, K.; Pettit, G.R. Convolutamidine A, a novel bioactive hydroxyoxindole alkaloid from marine bryozoan *Amathia convoluta*. *Tetrahedron Lett.* **1995**, *36*, 2783–2784, doi:10.1016/0040-4039(95)00395-5.
49. Kohno, J.; Koguchi, Y.; Nishio, M.; Nakao, K.; Kuroda, M.; Shimizu, R.; Ohnuki, T.; Komatsubara, S. Structures of TMC-95A–D: Novel Proteasome Inhibitors from *Apiospora montagnei* Sacc. TC 1093. *J. Org. Chem.* **2000**, *65*, 990–995, doi:10.1021/jo991375+.
50. Suzuki, H.; Morita, H.; Shiro, M.; Kobayashi, J. Celogentin K, a new cyclic peptide from the seeds of *Celosia argentea* and X-ray structure of moroidin. *Tetrahedron* **2004**, *60*, 2489–2495, doi:10.1016/j.tet.2004.01.053.
51. Kagata, T.; Saito, S.; Shigemori, H.; Ohsaki, A.; Ishiyama, H.; Kubota, T.; Kobayashi, J. Paratunamides A–D, Oxindole Alkaloids from *Cinnamodendron axillare*. *J. Nat. Prod.* **2006**, *69*, 1517–1521, doi:10.1021/np0602968.
52. Khuzhaev, V.U.; Zhalolov, I.; Turgunov, K.K.; Tashkhodzhaev, B.; Levkovich, M.G.; Aripova, S.F.; Shashkov, A.S. Alkaloids from *Arundo donax*. XVII. Structure of the Dimeric Indole Alkaloid Arundaphine. *Chem. Nat. Comp.* **2004**, *40*, 269–272, doi:10.1023/B:CONC.0000039139.30391.2c.
53. Kaufmann, S.H.E. Indole Propionic Acid: a Small Molecule Links between Gut Microbiota and Tuberculosis. *Antimicrob. Agents Chemother.* **2018**, *62*, doi:10.1128/AAC.00389-18.
54. Xue, F.; Zhang, S.; Liu, L.; Duan, W.; Wang, W. Organocatalytic Enantioselective Cross-Aldol Reactions of Aldehydes with Isatins: Formation of Two Contiguous Quaternary Centered 3-Substituted 3-Hydroxyindol-2-ones. *Chem. Asian J.* **2009**, *4*, 1664–1667, doi:10.1002/asia.200900243.
55. Heine, T.; Van Berkel, W.J.H.; Gassner, G.; Van Pée, K.-H.; Tischler, D. Two-Component FAD-Dependent Monooxygenases: Current Knowledge and Biotechnological Opportunities. *Biology* **2018**, *7*, 42, doi:10.3390/biology7030042.
56. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874, doi:10.1093/molbev/msw054.
57. Persson, B.; Bray, J.E.; Bruford, E.; Dellaporta, S.L.; Favia, A.D.; Gonzalez Duarte, R.; Jörnvall, H.; Kallberg, Y.; Kavanagh, K.L.; Kedishvili, N.; et al. The SDR (Short-Chain Dehydrogenase/Reductase and Related Enzymes) Nomenclature Initiative. *Chem. Biol. Interact.* **2009**, *178*, 94–98, doi:10.1016/j.cbi.2008.10.040.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

NOTES

Vilnius University Press
9 Saulėtekio Ave., Building III, LT-10222 Vilnius
Email: info@leidykla.vu.lt, www.leidykla.vu.lt
Print run copies 12