universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Development of an UnaG-Based Biosensor for Sensitive Bilirubin Detection

verfasst von | submitted by

Isabella Kraus BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 220
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Joint-Masterstudium Evolutionary Genomics and Systems Biology
Betreut von | Supervisor:	Univ.-Prof. David Berry Privatdoz. PhD

Acknowledgements

First and foremost, I would like to express my sincere gratitude to the Center for Molecular Medicine (CeMM) for providing the opportunity to conduct this research within a world-class scientific environment. I am particularly grateful for the collaborative atmosphere and access to state-of-the-art resources that have been invaluable throughout this project.

I am especially grateful to Clarissa Campbell and her lab for their guidance and support, as well as for the opportunity to work in such an intellectually stimulating and encouraging research environment. I want to extend my heartfelt thanks to Maximilian Baumgartner-Wyatt for his close mentorship during this time—your knowledge, encouragement, and continuous support have been immensely valuable to me.

I also wish to thank David Berry for generously contributing his time and expertise.

I am profoundly grateful to my parents, Alex and Elke, for giving me the opportunity to study abroad and for supporting me in pursuing a master's degree. Their encouragement and the privilege of such opportunities have shaped my path and made this journey possible.

Finally, I would like to express my deepest appreciation to my partner, Florian Schmidt, for being my constant source of strength and encouragement. Your unwavering support means the world to me.

Abstract

The gut microbiome plays a critical role in host physiology, particularly by modulating immune responses and key metabolic pathways. While the microbial conversion of bilirubin to urobilinogen has been recently elucidated, the downstream reduction of urobilinogen to stercobilinogen remains mechanistically uncharacterised. This thesis aimed to identify candidate microbial enzymes responsible for this transformation and develop a scalable fluorescence-based biosensor for bilirubin detection. A comprehensive bioinformatic screening of the Oligo Mouse Microbiome (OMM21) consortium was conducted, combining BLASTP homology searches against known bilirubin reductases (BilR), enzymatic classification via EC numbers, and cross-strain comparative analyses. The study identified multiple high-confidence BilR homologues, particularly in *Flintibacter butyricus*, *Thomasclavelia ramosa*, and *Escherichia coli*, as well as a refined set of CH–CH acting oxidoreductases (EC 1.3.-.-) as plausible urobilinogen reductase (UroR) candidates.

Parallel laboratory work established an inducible *UnaG*-based fluorescence biosensor in *E. coli* for bilirubin detection. The system demonstrated high specificity and sensitivity in vitro, reliably detecting nanomolar bilirubin concentrations. However, validation in biological samples revealed challenges, including bilirubin instability, photodegradation, and interfering background fluorescence. Experiments identified light exposure, prolonged incubation, and antibiotic presence as major factors reducing fluorescence signal fidelity. Additionally, biological validation using BilR-expressing and β-glucuronidase-deficient strains yielded inconsistent results, possibly due to limited substrate uptake or incomplete metabolite conversion.

In summary, this work provides (i) a prioritised set of candidate UroR enzymes for further functional validation and (ii) a validated, optimised *UnaG*-based biosensor system with defined operational limitations. The findings lay the groundwork for future investigations into the microbiome's role in bilirubin and urobilinogen metabolism, with potential implications for gut immune homeostasis and metabolic health.

Kurzfassung

Das Menschliche Mikrobiom spielt eine entscheidende Rolle für die Physiologie des Wirts, insbesondere durch die Modulation von Immunreaktionen und zentralen Stoffwechselwegen. Während die mikrobielle Umwandlung von Bilirubin zu Urobilinogen kürzlich aufgeklärt wurde, bleibt die nachgeschaltete Reduktion von Urobilinogen zu Stercobilinogen mechanistisch bislang uncharakterisiert. Ziel dieser Arbeit war es, mikrobielle Enzymkandidaten zu identifizieren, die für diese Umwandlung verantwortlich sind, sowie ein skalierbares, fluoreszenzbasiertes Biosensorsystem zur Detektion von Bilirubin zu entwickeln. Dazu wurde ein umfassendes bioinformatisches Screening des Oligo Mouse Microbiome (OMM21)-Konsortiums durchgeführt, das BLASTP-Homologiesuchen gegen bekannte Bilirubin-Reduktasen (BilR), eine enzymatische Klassifikation über EC-Nummern sowie vergleichende Analysen zwischen verschiedenen Stämmen kombinierte. Die Studie identifizierte mehrere hochkonfidente BilR-Homologe, insbesondere in *Flintibacter butyricus*, *Thomasclavelia ramosa* und *Escherichia coli*, sowie eine eingegrenzte Auswahl an Oxidoreduktasen, die auf CH–CH-Gruppen wirken (EC 1.3.-.-), als plausible Kandidaten für Urobilinogen-Reduktasen (UroR).

Parallel dazu wurde ein induzierbares, auf *UnaG* basierendes Fluoreszenz-Biosensorsystem in *E. coli* zur Bilirubin-Detektion etabliert. Das System zeigte in vitro eine hohe Spezifität und Sensitivität und konnte zuverlässig Bilirubinkonzentrationen im Nanomolarbereich detektieren. Die Validierung in biologischen Proben offenbarte jedoch Herausforderungen, darunter die Instabilität von Bilirubin, photochemische Zersetzung sowie störende Hintergrundfluoreszenz. Weitere Experimente identifizierten Lichtexposition, verlängerte Inkubationszeiten und das Vorhandensein von Antibiotika als Hauptfaktoren, die die Genauigkeit des Fluoreszenzsignals beeinträchtigten. Zudem führten Validierungsexperimente mit BilR-exprimierenden und β-Glucuronidase-defizienten Stämmen zu inkonsistenten Ergebnissen, möglicherweise aufgrund limitierter Substrataufnahme oder unvollständiger Metabolit-Umwandlung.

Zusammenfassend liefert diese Arbeit (i) eine priorisierte Liste von UroR-Enzymkandidaten für weitere funktionelle Validierungen sowie (ii) ein validiertes und optimiertes *UnaG*-basiertes Biosensorsystem mit definierten anwendungsspezifischen Einschränkungen. Die Ergebnisse bilden eine Grundlage für zukünftige Untersuchungen zur Rolle des Mikrobioms im Bilirubin- und Urobilinogen-Stoffwechsel und eröffnen potenzielle Implikationen für die intestinale Immunhomöostase und die metabolische Gesundheit.

Contents

1. Introduction

The human microbiome plays a fundamental role in regulating host physiology, particularly through its bidirectional interactions with the immune system. These molecular interactions shape immune responses, while the immune system concurrently modulates the composition and function of the microbial community (1; 2). Despite the increasing recognition of this intricate system, several mechanistic aspects remain insufficiently characterised, necessitating further investigation into their implications for health and disease (1; 2; 3; 4).

One such underexplored interaction involves the compound bilirubin and its associated pathological state, hyperbilirubinaemia, which is characterised by elevated bilirubin levels in the bloodstream (5; 6). Hyperbilirubinaemia primarily arises from liver dysfunction and is implicated in conditions such as jaundice and kernicterus, a severe bilirubin-induced neurological disorder (4; 7; 8; 9; 10). While hyperbilirubinaemia itself is not typically associated with inflammatory bowel disease (IBD), emerging evidence suggests that altered bilirubin metabolism occurs in the gut during IBD (4; 10; 11). This shows that the microbial taxonomic alterations associated with IBD and other inflammatory diseases may have a direct impact on the bilirubin reduction activity of the gut microbiome (4; 11). Given that the microbiome is involved in both bilirubin metabolism and the regulation of host immune responses, and considering that bilirubin-derived metabolites modulate immune signalling pathways, investigating this axis offers a promising avenue to better understand microbial-host interactions in the context of intestinal immune homeostasis (3; 4; 12).

Bilirubin is a catabolic end product of haem degradation, predominantly arising from the physiological turnover of erythrocytes. The degradation process is initiated by the enzymatic oxidation of haem via haem oxygenase, resulting in the formation of biliverdin (Figure 1.1). Biliverdin is subsequently enzymatically reduced by biliverdin reductase into unconjugated bilirubin. The unconjugated bilirubin is then transported in the serum, bound to albumin. It is delivered to the liver, where it undergoes conjugation via bilirubin glucuronyl transferase, facilitating its excretion into the bile. Once in the intestine, bilirubin is deconjugated and converted to urobilinogen by gut microbiota expressing bilirubin reductase. Urobilinogen follows distinct metabolic fates: it may be spontaneously oxidised in the kidney to form urobilin, which is excreted in urine, or it may be further reduced to stercobilinogen and subsequently oxidised into stercobilin, which is excreted in faeces (4; 6). However, the specific enzyme(s) and microbial species responsible for catalysing the reduction of urobilinogen to stercobilinogen have not yet been identified, representing a significant gap in our understanding of this metabolic pathway (4).

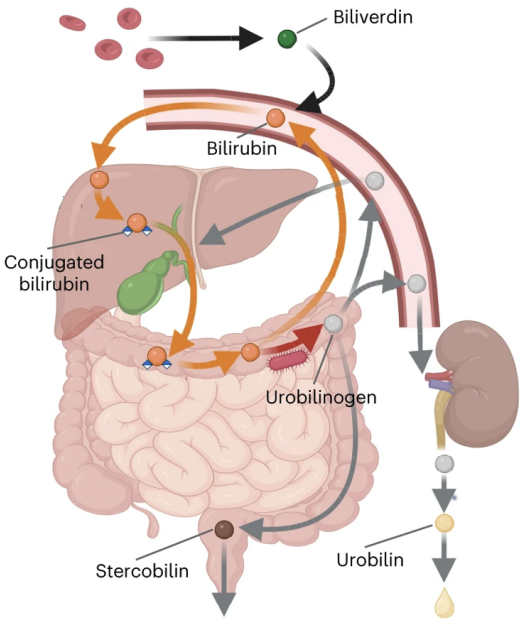

Figure 1.1.: Illustrated representation of the haem degradation and bilirubin pathway. Image generated using BioRender.com.

Recent advances in our understanding of microbial metabolism have revealed novel pathways by which gut microbiota interact with host-derived compounds. A key development in this area was the study conducted by Hall et al. (2024), which identified and characterised the microbial conversion of bilirubin into urobilinogen via a previously uncharacterised enzyme, *bilR*, a bilirubin reductase (4). In their study, fluorescence-based screening was applied to cultured gut bacterial strains in the presence of bilirubin to detect bilirubin reduction activity. Comparative genomic analyses between reducing and non-reducing strains were then employed to identify candidate reductase genes, ultimately leading to the identification of *bilR*—a gene homologous to 2,4-dienoyl-CoA reductase (EC: 1.3.1.34). This work established a foundation for understanding microbial involvement in bilirubin metabolism and its potential implications for host physiology (4).

Building upon this foundational work, the present thesis aims to further investigate the microbial mechanisms underlying bilirubin metabolism in the human gut. Drawing inspiration from the methodological framework established by Hall et al.(2024), this study combines fluorescence-based detection, bacterial culturing, and genomic screening to investigate the enzymatic conversion of urobilinogen to stercobilinogen. A key objective is to identify the microbial taxa and genes responsible for this reduction step, which remains uncharacterised in current literature.

Bioinformatically, the approach involves screening the defined Oligo Mouse Microbiome (OMM21) synthetic microbial community to assess the complexity of finding constituent species (13; 14; 15). Additionally, genomic analyses are employed to determine whether any OMM21 strains possess the *bilR* gene or its homologues, such as 2,4-dienoyl-CoA reductase, which has been previously implicated in bilirubin reduction (4). The OMM21 strains were specifically chosen not only for their relevance to the mouse gut microbiome but also for their culturable nature, which enables individual strain testing (13; 14; 15). This makes them particularly well-suited for follow-up investigations of candidate urobilinogen-reducing enzyme(s), both in vitro and in vivo, including future validation studies in cell culture systems and mouse models beyond the scope of the present project.

In parallel, laboratory-based experiments aim to develop a scalable, fluorescence-based biosensor for detecting bilirubin reduction. This biosensor is tested using co-cultures of bacteria known to reduce either conjugated or unconjugated bilirubin. To facilitate the detection of bilirubin reduction, this study employs the fluorescent protein UnaG, which fluoresces upon binding to unconjugated bilirubin (5; 16). A codon optimised plasmid, *pMAL-c5x_UnaG* (Figure 1.2), encoding *UnaG* was transformed into competent *Escherichia coli* (*E. coli*). In this construct, *UnaG* expression is controlled by the LacI repressor system, which inhibits transcription in the absence of the inducer isopropyl β-D-1-thiogalactopyranoside (IPTG). Upon IPTG addition, LacI repression is lifted, allowing RNA polymerase to transcribe the *UnaG* gene. This inducible system enables tight regulation of reporter gene expression. Furthermore, the plasmid contains an ampicillin resistance cassette, permitting the selective growth of transformed bacterial colonies.

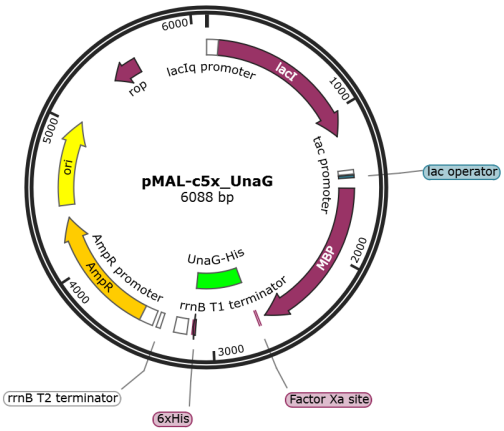

Figure 1.2.: Plasmid map of *pMAL-c5x_UnaG*, used for inducible expression of the UnaG fluorescent protein in *E. coli*. Image generated using SnapGene.

Structurally, urobilinogen and stercobilinogen are positional isomers that differ in the degree of reduction of their tetrapyrrole backbone (Figure 1.3). Stercobilinogen contains two additional hydrogen atoms across its methine (–CH=CH–) bridges compared to urobilinogen,

rendering it a more fully saturated molecule. The enzymatic conversion of urobilinogen to stercobilinogen thus involves the reduction of carbon–carbon double bonds within the conjugated pyrrolic system (17; 18). Based on this reaction type, the enzyme responsible is likely to fall within the EC 1 class of oxidoreductases, and more specifically, the EC 1.3 subclass—enzymes acting on CH–CH groups (19; 20).

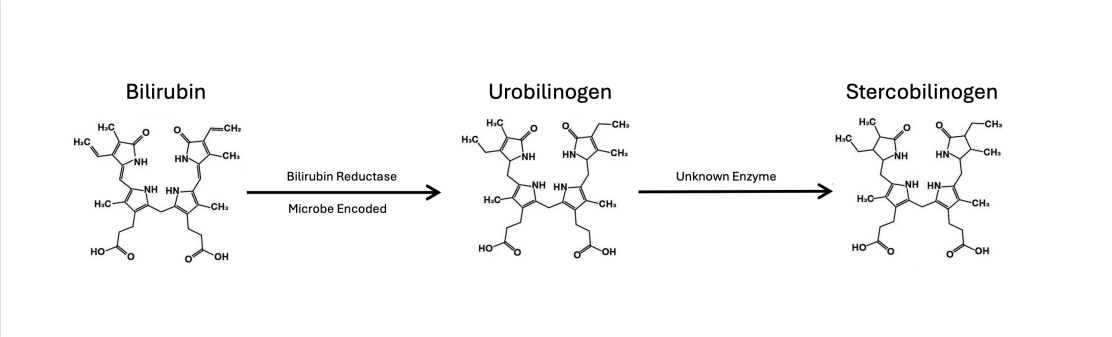

Figure 1.3.: Bilirubin degradation pathway showing the structural conversion of bilirubin to urobilinogen and stercobilinogen. Stercobilinogen is the more reduced isomer, containing two additional hydrogen atoms across the methine (–CH=CH–) bridges of the tetrapyrrole backbone.

Thus, the hypothesis is that specific gut microbes encode urobilinogen reductases (UroR) that catalyse the reduction of urobilinogen to stercobilinogen. These enzymes, yet to be characterised, are expected to function as CH–CH oxidoreductases within the broader context of microbial bilirubin metabolism.

A deeper understanding of this microbial transformation has broad biomedical relevance. Given the integral role of the microbiome in modulating host immunity and bilirubin turnover, elucidating the enzymatic basis of urobilinogen reduction may reveal new molecular links between microbial metabolism and host immune regulation. Ultimately, this line of investigation could inform the development of novel microbiome-targeted therapies for inflammatory and metabolic disorders associated with dysregulated bilirubin levels, including liver dysfunction, IBD and other related immunopathologies.

2. Methods and Materials

2.1. Bioinformatic Identification of Candidate Urobilinogen Reductases

As an initial step toward identifying candidate urobilinogen reductases (UroR) within the Oligo Mouse Microbiome (OMM21) consortium, a targeted homology search was conducted to assess the presence of known bilirubin reductases (BilR). Characterised forms of bilirubin reductase were used as query sequences in BLASTP searches against all OMM21 proteomes (21; 22). These included the short BilR variants from *Clostridioides difficile* and *Clostridium symbiosum*, as well as the long variant from *Ruminococcus gnavus*. Putative homologues identified through these searches were recorded for further analysis. This targeted search was particularly relevant, as BilR plays a central role in the metabolic pathway of interest. Identifying its presence in OMM21 strains is essential for evaluating their conversion potential. Since the overarching aim is to discover a novel urobilin reductase (UroR), confirming whether BilR is already encoded by these organisms provides critical context. Additionally, a literature review was performed to determine whether any members of the OMM21 consortium, or other gut-associated bacteria, have been previously reported to reduce urobilinogen.

Subsequently, genomic homology analyses were conducted to identify candidate UroR proteins within the OMM21 consortium. Raw sequencing reads for the OMM21 members were obtained as FASTA files from the CeMM institute. These organisms were selected due to their well-characterised, reproducible nature, ease of cultivation in our facility, and widespread use in murine microbiome experiments (13; 14; 15).

Protein prediction was performed using Prokka (version 1.14.6) (23). Enzyme Commission (EC) numbers were assigned to the Prokka-predicted proteins with ECPred (version 1.1) (24) and EggNOG-mapper (version 2.1.12) (25; 26) under default parameters. Both tools were used for the subsequent comparison of EC number assignments. Proteins annotated as oxidoreductases (EC 1.-.-.-), particularly those acting on CH–CH group donors (EC 1.3.-.-), were extracted for further investigation.

Plots and comparative analyses of ECPred and EggNOG-mapper annotations, as well as all wet lab data and visualisations, were carried out using R software (version 4.3.1).

2.2. Development of a Scalable Bilirubin Biosensor

2.2.1. High Efficiency Transformation Protocol

For the scalable bilirubin biosensor, a plasmid harbouring the *UnaG* gene was introduced into competent *Escherichia coli* via the Invitrogen Protocol C2987H/C2987I. Two 1 mL cryogenic tubes containing BL21 (DE3) competent *E.coli* (New England Biolabs) were thawed on ice for 10 minutes. One tube received 1 µL of pMAL-c5x-UnaG plasmid DNA (positive control; Addgene), while the second tube received 1 µL of sterile distilled water (negative control). Both tubes were incubated on ice for 30 minutes, subjected to a 10-second heat shock at 42 °C, and returned to ice for an additional 5 minutes. Thereafter, 950 µL of Super Optimal broth with Catabolite repression S.O.C (Invitrogen) was added to each tube, which was then incubated in a heat block at 37 °C and 300 rpm for 60 minutes. Ampicillin-containing agar plates were prewarmed at 37 °C, and aliquots of the transformed cultures (50 µL, 100 µL and 150 µL) were spread onto three plates for each control. All plates were incubated overnight at 37 °C.

Following overnight incubation, ten bacterial colonies from each plate were inoculated into Lysogeny Broth (LB) supplemented with 1000 µg mL^{-1} ampicillin and incubated at 37 °C with shaking at 170 rpm. No growth was observed in the negative control. The following morning, the optical densities (OD_{600}) of all cultures were measured, and the two cultures from each plate with the highest OD_{600} values were selected for plasmid extraction. *E. coli* cells were harvested by centrifugation, and plasmid DNA was extracted using the Monarch® Plasmid Miniprep Kit (New England Biolabs), following the manufacturer's protocol. DNA was eluted in nuclease-free water and quantified using a NanoDrop spectrophotometer. The two plasmid samples with the highest concentrations were selected for sequencing, which was performed by Microsynth.

Rather than sequencing the entire plasmid, only the region containing the gene of interest *UnaG* was targeted, spanning nucleotide positions 2686 to 3069. The two sequencing results (Plasmid_1 and Plasmid_2) were received as FASTA files. A single FASTA file was then compiled, containing the two experimental plasmid sequences (Plasmid_1 and Plasmid_2), the full reference plasmid, and a control plasmid sequence corresponding to the gene region of interest (positions 2686–3069). This FASTA file was submitted to the Clustal Omega multiple sequence alignment tool (EMBL-EBI, version 1) (27), with the alignment type set to "DNA". The resulting alignment was visualised using MView (version 2)(28) to validate successful introduction of the UnaG plasmid into *UnaG*.

2.2.2. Determination of Non-Toxic IPTG Levels

The subsequent steps were focused on assessing various isopropyl β-D-1-thiogalactopyranoside (IPTG; Thermoscientific) concentrations that had been reported as effective in other *E.coli* strains. This assessment was undertaken to ascertain whether the tested IPTG concentrations exerted toxic effects on *E. coli* and to identify the concentration that maximised bilirubin synthesis.

A literature review identified a working IPTG concentration range between 0.1mM and 1mM(16; 29; 30; 31). Following the literature review (16; 29; 30; 31), a dilution series was designed, starting at 4 mM (four times the highest reported concentration) and decreasing sequentially to 2 mM, 1 mM, 0.5 mM, 0.25 mM, 0.125 mM and 0.1 mM, alongside a negative non-inducer control. Two *E.coli* cultures, previously confirmed to harbour the plasmid through sequencing, were incubated overnight in LB medium containing ampicillin (1000 µg mL^{-1}) at 37 °C with shaking at 170 rpm. Cell densities were adjusted to an OD_{600} of 0.5 by dilution in prewarmed LB-ampicillin medium. A 96-well microplate was then prepared for growth-curve analysis. One column was designated as the blank, with 200 µL of LB-ampicillin medium. Negative-control wells were prepared by combining 100 µL of bacterial suspension with 100 µL of LB–ampicillin medium. IPTG dilutions were pipetted according to the predefined series, with immediate induction (0 h). A Breathe-Easy® sealing membrane (Sigma-Aldrich) was applied to the plate to prevent evaporation and cross-contamination, and absorbance at 600 nm was recorded overnight using a microplate reader (SpectraMax® i3x Multi-Mode Microplate Reader). Bacterial viability was monitored to assess potential IPTG toxicity.

2.2.3. Toxicity Testing of Bilirubin and Analogues on *E.coli* Growth

The succeeding objective was to determine a bilirubin concentration that conferred both high sensitivity and specificity while avoiding any toxic effects on *E.coli*. Additionally, it was necessary to verify that IPTG alone mediated induction of *UnaG* expression, rather than an alternative endogenous mechanism. A comprehensive literature survey was undertaken to collate bilirubin levels applied in analogous studies of *UnaG*–bilirubin interactions. Reported bilirubin concentrations were found to range from 17.1 µM to 0.095 nM (4; 16; 29; 32; 33; 34). Based on this survey, a broad range of bilirubin (Cayman Chemical) concentrations — 12 µM, 1.25 µM, and 0.012 µM were selected. Additionally, stercobilin and urobilin (Cayman Chemical) were included at a concentration of 12 µM to assess any potential toxicity to *E.coli*. *E.coli* stocks were inoculated in 5 mL LB medium supplemented with 1000 µg mL^{-1} ampicillin using a sterile loop, and cultures were incubated overnight at 37 °C with agitation at 170 rpm. The following morning, overnight cultures were diluted 1:100 into fresh LB-ampicillin medium and incubated under identical conditions to reach mid-logarithmic phase.

Three IPTG induction conditions were prepared by combining the culture with LB–ampicillin 1:1 containing either no inducer (control), 4 mM IPTG or 1 mM IPTG. The 4 mM IPTG solution was prepared by diluting 20 µL of a 10 mM IPTG stock into 480 µL LB–ampicillin medium. For the 1 mM condition, 5 µL of the same 10 mM stock was diluted into 495 µL LB-ampicillin medium. Each 6 mL induction mixture was aliquoted into 96-well microplates and incubated at 37 °C for 2–4 h to induce IPTG-mediated expression of the *UnaG* gene. Metabolite stocks, 15 mM bilirubin, urobilin, and stercobilin in DMSO (ThermoFisher) were diluted in 1000 µg mL^{-1} LB–ampicillin to yield final assay concentrations of 12 µM, 1.2 µM, and 0.012 µM. Following induction, these dilutions were added to the microplate wells, which were then sealed with Breathe-Easy® membranes. Absorbance at 600 nm was recorded overnight to monitor culture density and assess any metabolite-induced growth inhibition.

2.2.4. Functional Evaluation of the IPTG–*UnaG* Fluorescence System

Having confirmed that neither IPTG nor the metabolites exerted cytotoxic effects on *E. coli*, the IPTG–*UnaG* system was evaluated for functional performance. The assay was designed to quantify the bilirubin–*UnaG* interaction, thereby determining induction sensitivity and specificity at each metabolite concentration, and to ascertain whether IPTG induction triggered *UnaG* synthesis and if bilirubin binding to the expressed protein produced a measurable fluorescence signal. Parallel control experiments omitting IPTG or employing a *UnaG*-negative strain were conducted to exclude endogenous fluorescence; stercobilin and urobilin were assayed concurrently to assess background signals from related tetrapyrroles. Fluorescence measurements were performed with excitation at 480 nm and emission detection at 535 nm. These parameters were applied consistently in this and all subsequent assays to ensure comparability of results.

Overnight cultures of *E. coli* harbouring the pMAL-c5x-UnaG plasmid and of BL21 (DE3) cells lacking the plasmid were grown at 37 °C. The following morning, each culture was diluted 1:100 into fresh LB–ampicillin and incubated for 2 h, at which point OD_{600} reached 0.5. The cultures were used for downstream assays.
Four assay conditions were established:

(i) Plasmid-containing cells without IPTG

(ii) Cell-free medium with metabolites only

(iii) Plasmid-containing cells with 4 mM IPTG

(iv) Plasmid-free cells with 4 mM IPTG

Each condition was accompanied by a control comprising the corresponding medium or bacterial culture. Metabolites were added at final concentrations of 12 µM bilirubin, 1.2 µM bilirubin, 12 µM stercobilin and 12 µM urobilin, with no further incubation (0 h). Plates were sealed with a membrane, and fluorescence was immediately measured in a microplate reader.

2.2.5. Kinetic and Sensitivity Assessment of the IPTG–*UnaG* System

IPTG induction durations were optimised by subjecting mid-logarithmic *E. coli* cultures to 2 mM IPTG for 0, 2, 4 and 18 h. Following the addition of 2 mM IPTG, cultures were incubated at 37 °C with agitation at 170 rpm for the specified periods. Controls were included for each time point, and metabolite assays were performed by adding bilirubin (12 µM and 1.2 µM), stercobilin (12 µM) and urobilin (12 µM) to each culture before measurement. Subsequently, an additional induction time series was performed using incubation periods of 4 h and 6 h. Mid-logarithmic *E. coli* cultures were subjected to 2 mM IPTG and incubated at 37 °C with agitation at 170 rpm for the specified durations. To assess the dynamic range of the fluorescence response, bilirubin, stercobilin and urobilin were added immediately before measurement at final concentrations of 25 µM, 12.5 µM and 1.25 µM, respectively. The plate was sealed with a gas-permeable membrane, and fluorescence was recorded.

Finally, the lower limit of bilirubin detection was established. Mid-logarithmic *E. coli* cultures were induced with 2 mM IPTG and incubated at 37 °C with agitation at 170 rpm for 4 h and 6 h. Before measurement, bilirubin was added at final concentrations of 2.5 µM, 1.25 µM, 0.6 µM, 0.3 µM, 0.125 µM, 0.06 µM and 0.03 µM. The plate was sealed with a gas-permeable membrane, and fluorescence was recorded to define the assay's lower detection threshold.

2.2.6. Validation of the IPTG–*UnaG* System

To highlight the applicability of the system, *E. coli* strains with different metabolising capabilities were employed to validate the IPTG–*UnaG* system. The first two wild-type (BilR-WT) and catalytically dead (BilR-CD) variants were derived from BL21 (DE3) cells (New England Biolabs) transformed with the pSF-OXB15 vector (Sigma-Aldrich, Figure 2.1), which harbours the *Ruminococcus gnavus* bilirubin reductase (*BilR*) gene and a kanamycin resistance marker. In the BilR-WT strain, the *BilR* gene remained unaltered, whereas in the BilR-CD strain, point mutations were introduced to inactivate the enzyme. Three additional strains were used from the *E. coli* K-12 Keio collection (Horizon Discovery): the parental K-12 wild-type, which retains the native *gusA* gene, and two isogenic *gusA* knockout mutants.

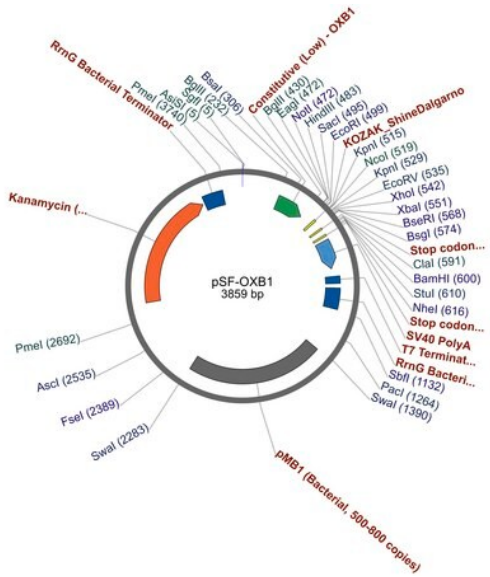

Figure 2.1.: Plasmid map of pSF-OXB15 vector, used for inducible expression of the *R. gnavus* *BilR* gene in the BilR-CD and BilR-WT strains. Image source: Sigma-Aldrich (35)

To highlight the applicability of the system, *E. coli* strains with different metabolising capabilities were employed to validate the IPTG–*UnaG* system. The first two wild-type (BilR-WT) and catalytically dead (BilR-CD) variants were derived from BL21 (DE3) cells (New

England Biolabs) transformed with the pSF-OXB15 vector (Sigma-Aldrich), which harbours the *Ruminococcus gnavus* bilirubin reductase (*BilR*) gene and a kanamycin resistance marker. In the BilR-WT strain, the *BilR* gene remained unaltered, whereas in the BilR-CD strain, point mutations were introduced to inactivate the enzyme. Three additional strains were used from the *E. coli* K-12 Keio collection (Horizon Discovery): the parental K-12 wild-type, which retains the native *gusA* gene, and two isogenic *gusA* knockout mutants. In the gut, β-glucuronidase (*gusA*) converts conjugated bilirubin to the unconjugated form, facilitating its reabsorption (36).

To prepare the strains for sequencing and further analysis, all five bacterial strains were cultured overnight at 37 °C with agitation (170 rpm) in LB medium. The BilR-WT and BilR-CD strains were grown in LB supplemented with kanamycin (1000 µg mL^{-1}), while the K-12 Keio strains were cultured in antibiotic-free LB. Following incubation, cells were harvested by centrifugation, and plasmid DNA was extracted from the resulting pellets using the Monarch® Plasmid Miniprep Kit (New England Biolabs), following the manufacturer's instructions. DNA was eluted in nuclease-free water and quantified using NanoDrop spectrophotometry. As the K-12 Keio strains did not harbour plasmids, NanoDrop analysis was used solely to verify the absence of detectable DNA. For the BilR-WT and BilR-CD strains, the plasmid preparations yielding the highest DNA concentrations were selected for sequencing, which was performed by Microsynth.

The entire plasmid was sequenced, and the resulting sequencing data for BilR-WT and BilR-CD were received as FASTA files. A single FASTA file was then compiled, containing both experimental plasmid sequences along with the full reference plasmid sequence. This file was submitted to the Clustal Omega multiple sequence alignment tool (EMBL-EBI, version 1) (27), with the alignment type set to "DNA". The alignment output was visualised using MView (version 2) (28).

To validate the IPTG–*UnaG* system in the context of bilirubin reduction, BilR-WT and BilR-CD strains were cultivated in LB medium supplemented with 1000 µg mL^{-1} kanamycin (overnight at 37 °C with agitation 170 rpm). The following morning, cultures were adjusted to an OD_{600} of 0.1–0.2 and aliquoted into sterile tubes. Four conditions were established in duplicate: BilR-WT ± exogenous bilirubin, BilR-CD ± exogenous bilirubin, and cell-free LB–kanamycin ± bilirubin as a medium control. Bilirubin was supplied to a final concentration of 0.6 µM. Tubes were incubated overnight at 37 °C, 170 rpm, and final OD_{600} readings were recorded to assess growth and any bilirubin-dependent effects. On the subsequent day, *E. coli* cultures harbouring the IPTG–*UnaG* construct were grown to mid-log phase and induced with IPTG (2 mM final) for 4 h at 37 °C, 170 rpm. At the end of induction, BilR-WT, BilR-CD and medium control samples were centrifuged, and the supernatants were sterile-filtered; 100 µL aliquots were dispensed into individual columns of a 96-well plate. Subsequently, 100 µL of induced UnaG–*E. coli* culture was added to each well. A bilirubin titration series (1.25 µM, 0.6 µM, 0.3 µM, 0.15 µM and 0.078 µM) was prepared by serial dilution in the bacteria suspension. Plates were sealed with a gas-permeable membrane, and fluorescence was measured immediately (excitation 480 nm, emission 535 nm).

The *E. coli* K-12 Keio parental strain and two isogenic *gusA* knockout mutants (mutant 1 and mutant 2) were revived by overnight incubation in LB medium at 37 °C. Cultures were diluted to an OD_{600} of 0.1–0.2 in LB and subdivided into duplicate aliquots supplemented with or without 25 µM bilirubin ditaurate (Santa Cruz Biotechnology); cell-free media controls (with and without bilirubin) were prepared in parallel. All tubes were incubated overnight at 37 °C with agitation (170 rpm).

Subsequently, *UnaG*-expressing *E. coli* cultures were induced with 2 mM IPTG for 4 h at 37 °C, 170 rpm. Bilirubin ditaurate-treated Keio culture supernatants were centrifuged (4,000 × g, 10 min) and sterile filtered. In a 96-well microplate, 100 µL of each supernatant was mixed with 100 µL of induced *UnaG* culture. A bilirubin calibration series (2.5 µM, 1.25 µM and 0.6 µM) was prepared by mixing induced *UnaG* cells 1:1 with corresponding bilirubin solutions. Plates were sealed with a gas-permeable membrane, and fluorescence (excitation 480 nm, emission 535 nm) was measured immediately.

2.2.7. Investigation of Bilirubin Degradation, Antibiotic and Bilirubin Ditaurate Effects

Unexpectedly weak and inconsistent fluorescence signals during BilR-WT and BilR-CD validation prompted an investigation into bilirubin stability under assay conditions. It was hypothesised that prolonged incubation might induce photodegradation. To test this, LB–kanamycin (1000 µg mL^{-1}) tubes containing 6 µM bilirubin were incubated at 37 °C and 170 rpm for 2 h and 4 h, either exposed to ambient light or shielded by aluminium foil. In parallel, *UnaG–E. coli* cultures were induced with 2 mM IPTG and incubated for 4 h. Immediately prior to measurement, 100 µL of each bilirubin treatment was mixed with 100 µL of induced culture in a 96-well plate, which was then sealed with a gas-permeable membrane and read (excitation 480 nm, emission 535 nm).

A second hypothesis proposed that kanamycin might accelerate bilirubin degradation. To evaluate the effect of kanamycin on bilirubin integrity, *UnaG–E. coli* cultures were again induced with 2 mM IPTG and incubated for 4 h. A bilirubin dilution series (6 µM, 3 µM and 1.5 µM) was prepared in LB both with and without kanamycin (1000 µg mL^{-1}). Additionally, LB–kanamycin media were diluted 1:3 and 1:4 with antibiotic-free LB before bilirubin addition. Each dilution (100 µL) was mixed with 100 µL of induced culture, transferred to a 96-well plate, sealed, and fluorescence was recorded immediately (excitation 480 nm, emission 535 nm).

Since neither light exposure nor kanamycin significantly accelerated bilirubin degradation, the intrinsic stability of bilirubin in LB medium was subsequently assessed over time. Bilirubin was added to antibiotic-free LB at a final concentration of 25 µM and incubated (37 °C, 170 rpm) for 18 h, 6 h, 4 h, 2 h, 1 h and immediately before measurement (0 h). Concurrently, *UnaG*-expressing *E. coli* cultures were induced with 2 mM IPTG for 4 h under identical conditions. At each time point, 100 µL of the bilirubin-treated medium was combined with 100

µL of induced culture in a 96-well microplate. Calibration standards were prepared by mixing induced culture 1:1 with LB containing bilirubin at 2.5 µM, 1.25 µM, 0.6 µM and 0.3 µM. Plates were sealed with a gas-permeable membrane, and fluorescence was measured immediately.

Due to the unusually low and inconsistent fluorescence signals observed in the *E. coli* K-12 Keio experiments, it was hypothesised that bilirubin ditaurate (BDT) might be exerting an inhibitory effect on bacterial growth or metabolism, thereby contributing to the aberrant results. To investigate this, the *E. coli* K-12 Keio parental strain and the two isogenic *gusA* knockout mutants (mutant 1 and mutant 2), along with the *UnaG*-expressing *E. coli* strain, were revived by overnight incubation in LB medium at 37 °C with agitation.

The following morning, cultures were diluted to an OD_{600} of 0.1–0.2 in fresh LB medium and subdivided into duplicate aliquots, supplemented with either 25 µM or 12.5 µM bilirubin ditaurate (Santa Cruz Biotechnology). Cell-free media controls (without BDT) were prepared in parallel. Following incubation, the Keio culture supernatants were centrifuged at 4,000 $\times$ g for 10 minutes and sterile filtered.

In a 96-well microplate, 100 µL of each supernatant was combined with 100 µL of induced *UnaG* culture. The plates were sealed with a gas-permeable membrane to minimise evaporation and allow oxygen exchange. Absorbance at 600 nm was measured immediately to monitor any immediate impact on cell density or optical interference.

3. Results

3.1. Bioinformatic Identification of Candidate Urobilinogen Reductases

To assess the distribution of bilirubin reductase (BilR) homologues within the OMM21 strains, BLASTP searches were performed using both short and long BilR sequences as queries. The presence of homologous sequences was evaluated based on two primary criteria: an E-value threshold of $\leq 1\text{e}{-13}$, commonly considered indicative of statistically significant homology, and a query cover threshold of $\geq 50\%$ to ensure meaningful sequence alignment. The short BilR variant from *C. difficile* was identified in 13 of the 21 OMM21 strains.

Table 3.1.: Short BilR from *C. difficile*

Strain	Hit	Enzyme	Query Cover	Evalue
Acutalibacter muris	yes	FAD-dependent oxidoreductase	0.98	1E-36
Adlercreutzia mucosicola	yes	FAD-dependent oxidoreductase	0.96	2E-36
Akkermansia muciniphila	yes	NADH:flavin oxidoreductase	0.73	2E-19
Bacteroides caecimuris	no	–	–	–
Bifidobacterium animalis	no	–	–	–
Blautia pseudococcoides	yes	FAD-dependent oxidoreductase	0.95	1E-41
Clostridium innocuum	yes	NAD(P)/FAD-dependent oxidoreductase	0.99	1E-37
Enterocloster clostridioformis	yes	FAD-dependent oxidoreductase	0.98	1E-37
Enterococcus faecalis	yes	FAD-dependent oxidoreductase	0.95	1E-21
Escherichia coli	yes	NADPH 2,4-dienoyl-CoA reductase	1.00	4E-40
Extibacter muris	yes	NAD(P)/FAD-dependent oxidoreductase	0.99	1E-39
Flavonifractor plautii	yes	NADH:flavin oxidoreductase	0.99	1E-121
Flintibacter butyricus	yes	bilirubin reductase	1.00	1E-125
Ligilactobacillus murinus	no	–	–	–
Limosilactobacillus reuteri	yes	NAD(P)/FAD-dependent oxidoreductase	0.96	7E-25
Mucispirillum schaedleri	no	–	–	–
Muribaculum intestinale	no	–	–	–
Parabacteroides goldsteinii	no	–	–	–
Thomasclavelia ramosa	yes	Pyridine disulfide oxidoreductase	1.00	3E-69
Turicimonas muris	no	–	–	–
Xylanibacter rodentium	no	–	–	–

Most of the hits exhibited high query cover values (ranging from 0.73 to 1.00) and highly significant E-values (as low as 1E-125), indicating strong sequence homology (3.1). Notably, *F. butyricus* and *E. coli* demonstrated the highest sequence identity, each with a query cover of

1.00. These hits were annotated as either bilirubin reductase or its homologue, 2,4-dienoyl-CoA reductase. In contrast, eight strains, including *B. caecimuris*, *M. intestinale*, and *X. rodentium*, did not yield any significant hits under the defined criteria. Using the short BilR from *C. symbiosum* as the query, homologues were again found in 15 strains (3.2). The majority displayed high query cover (up to 0.99) and E-values below the threshold, reinforcing the robustness of the hits. Notably, in *F. butyricus*, a perfect match was observed with a query cover of 1.00 and an E-value of 0.0. *E. coli* was the only strain in this search to return a hit for the bilirubin reductase homologue 2,4-dienoyl-CoA reductase, albeit with a higher E-value of 1.0. The annotation of this hit as bilirubin reductase further supports its identification as a functional BilR homologue. However, some discrepancies were observed relative to the *C. difficile* BilR results. For instance, *M. intestinale*, which had no hit in the *C. difficile*-based screen, returned a match annotated as "Old Yellow Enzymes" with a query cover of 0.99. Furthermore, *M. schaedleri* and *L. murinus*, initially classified as lacking BilR, showed a high query cover (0.98 and 0.96) in this screen.

Table 3.2.: Short BilR from *C. symbiosum*

Strain	Hit	Enzyme	Query Cover	Evalue
Acutalibacter muris	yes	FAD-dependent oxidoreductase	0.64	1E-27
Adlercreutzia mucosicola	yes	FAD-dependent oxidoreductase	0.97	3E-22
Akkermansia muciniphila	yes	NADH:flavin oxidoreductase	0.94	1E-12
Bacteroides caecimuris	no	–	–	–
Bifidobacterium animalis	no	–	–	–
Blautia pseudococcoides	yes	FAD-dependent oxidoreductase	0.94	6E-34
Clostridium innocuum	yes	FAD-dependent oxidoreductase	0.96	3E-33
Enterocloster clostridioformis	yes	FAD-dependent oxidoreductase	0.96	6E-35
Enterococcus faecalis	yes	FAD-dependent oxidoreductase	0.95	3E-26
Escherichia coli	yes	NADH:flavin oxidoreductase	0.94	2E-24
Extibacter muris	no	–	–	–
Flavonifractor plautii	no	–	–	–
Flintibacter butyricus	yes	bilirubin reductase	1.00	0.0
Ligilactobacillus murinus	yes	NADH:flavin oxidoreductase	0.92	4E-13
Limosilactobacillus reuteri	yes	NAD(P)/FAD-dependent oxidoreductase	0.95	9E-21
Mucispirillum schaedleri	yes	FAD-dependent oxidoreductase	0.98	1E-32
Muribaculum intestinale	yes	Old Yellow Enzymes	0.99	0.0
Parabacteroides goldsteinii	yes	NADH:flavin oxidoreductase	0.89	4E-17
Thomasclavelia ramosa	yes	NADH:flavin oxidoreductase	0.92	2E-55
Turicimonas muris	no	–	–	–
Xylanibacter rodentium	no	–	–	–

In a third analysis using the long BilR variant from *R. gnavus*, 14 strains were identified as containing putative homologues (Table 3.3). Query cover values ranged from 0.51 to 1.00, and E-values were consistently low (e.g., $1E^{-104}$ for *E. muris*), indicating strong sequence conservation. The highest homology was seen in *C. innocuum*, *E. clostridioformis*, and *E. muris*, each exhibiting complete query coverage (1.00) and highly significant E-values. Interestingly, in both short-form queries, *F. butyricus* yielded a perfect hit annotated as bilirubin reductase,

with full query coverage and an E-value of 0.0, strongly suggesting the presence of a canonical BilR. In contrast, in Table 3.3, a different type of oxidoreductase was identified in this strain. Instead, *T. ramosa* showed a perfect match for a long-form bilirubin reductase, again with complete query coverage and an E-value of 0.0, indicating a highly confident and functionally relevant homologue. Notably, *E. coli* also returned a significant hit in this search, implying that it possesses homologues for both the short and long BilR variants.

Table 3.3.: Long BilR (Multiple Species)

Strain	Hit	Enzyme	Query Cover	Evalue
Acutalibacter muris	yes	FAD-dependent oxidoreductase	0.92	5E-63
Adlercreutzia mucosicola	yes	FAD-dependent oxidoreductase	0.99	8E-56
Akkermansia muciniphila	yes	NADH:flavin oxidoreductase	0.51	6E-19
Bacteroides caecimuris	no	–	–	–
Bifidobacterium animalis	yes	NADH:flavin oxidoreductase	0.52	8E-33
Blautia pseudococcoides	yes	NAD(P)/FAD-dependent oxidoreductase	0.99	5E-88
Clostridium innocuum	yes	NAD(P)/FAD-dependent oxidoreductase	1.00	8E-100
Enterocloster clostridioformis	yes	FAD-dependent oxidoreductase	1.00	1E-83
Enterococcus faecalis	yes	FAD-dependent oxidoreductase	0.98	2E-78
Escherichia coli	yes	NADPH 2,4-dienoyl-CoA reductase	0.99	1E-78
Extibacter muris	yes	NAD(P)/FAD-dependent oxidoreductase	1.00	2E-104
Flavonifractor plautii	yes	FAD-dependent oxidoreductase	1.00	4E-77
Flintibacter butyricus	yes	FAD-dependent oxidoreductase	0.99	1E-89
Ligilactobacillus murinus	no	–	–	–
Limosilactobacillus reuteri	yes	NAD(P)/FAD-dependent oxidoreductase	1.00	1E-81
Mucispirillum schaedleri	no	–	–	–
Muribaculum intestinale	no	–	–	–
Parabacteroides goldsteinii	no	–	–	–
Thomasclavelia ramosa	yes	bilirubin reductase, long form	1.00	0.0
Turicimonas muris	no	–	–	–
Xylanibacter rodentium	no	–	–	–

Across all three searches, the presence of potential BilR homologues was consistent in several strains, including *A. muris*, *A. mucosicola*, *E.Coli*, and *B. pseudococcoides*, which consistently yielded significant hits. The results suggest that a large subset of the OMM21 community potentially encodes BilR or BilR-like enzymes, warranting further investigation into their enzymatic activities. These findings provide a foundation for further functional validation of candidate urobilinogen reductases in the community.

Comparative genomic analyses of the 21-member OMM consortium were performed to identify potential UroR candidates, focusing on proteins annotated with EC numbers corresponding to oxidoreductases (EC 1.-.-.-), and more specifically, to those acting on CH–CH group donors (EC 1.3.-.-). The annotation pipeline employed both ECPred and EggNOG-mapper tools to ensure comprehensive coverage and validation across the dataset.
Figure 3.1 illustrates the total number of proteins annotated with an EC number across the 21 OMM strains. Annotation using ECPred yielded a substantially higher number of

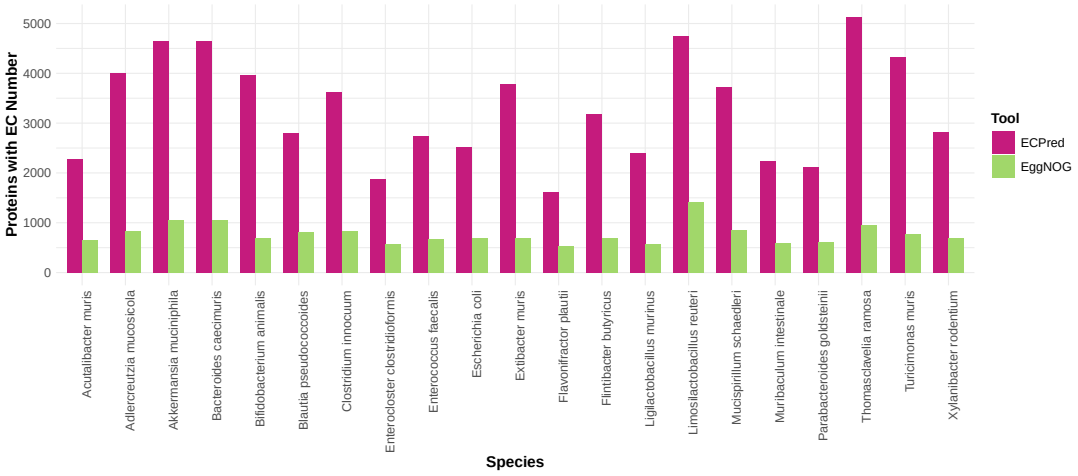

Figure 3.1.: Comparison of total proteins annotated with EC numbers across OMM species using ECPred and EggNOG-mapper.

EC-annotated proteins in all species compared to EggNOG-mapper. This difference reflects the broader annotation capabilities of ECPred, likely due to its orthology-based methodology and extensive reference database. For instance, in *T. ramosa*, ECPred assigned over 5,000 proteins with EC numbers, while EggNOG-mapper yielded fewer than 1,000.

To refine the list of potential UroR candidates, proteins annotated as Class 1 oxidoreductases (EC 1.-.-.-) were extracted and quantified. As shown in Figure 3.2, ECPred consistently identified a greater number of Class 1 enzymes compared to EggNOG-mapper across all species. The filtering for oxidoreductases substantially reduced the pool of non-relevant proteins, resulting in approximately a tenfold decrease in candidate numbers for most species. Nevertheless, the remaining set still included over 100 candidate proteins per species, which remains impractically large when aiming to identify a single, specific UroR enzyme.

Further filtering revealed the subset of EC 1.3.-.- enzymes—those acting on the CH–CH group of donors. Figure 3.3 displays the number of proteins annotated within this specific class. Again, ECPred consistently reported a higher number of relevant enzymes than EggNOG-mapper across most species. The refined candidate pool was further reduced, with fewer than 10 potential enzymes identified per species in the majority of cases. This level of reduction rendered experimental follow-up substantially more feasible, aligning with the objective of isolating a single, specific UroR enzyme.

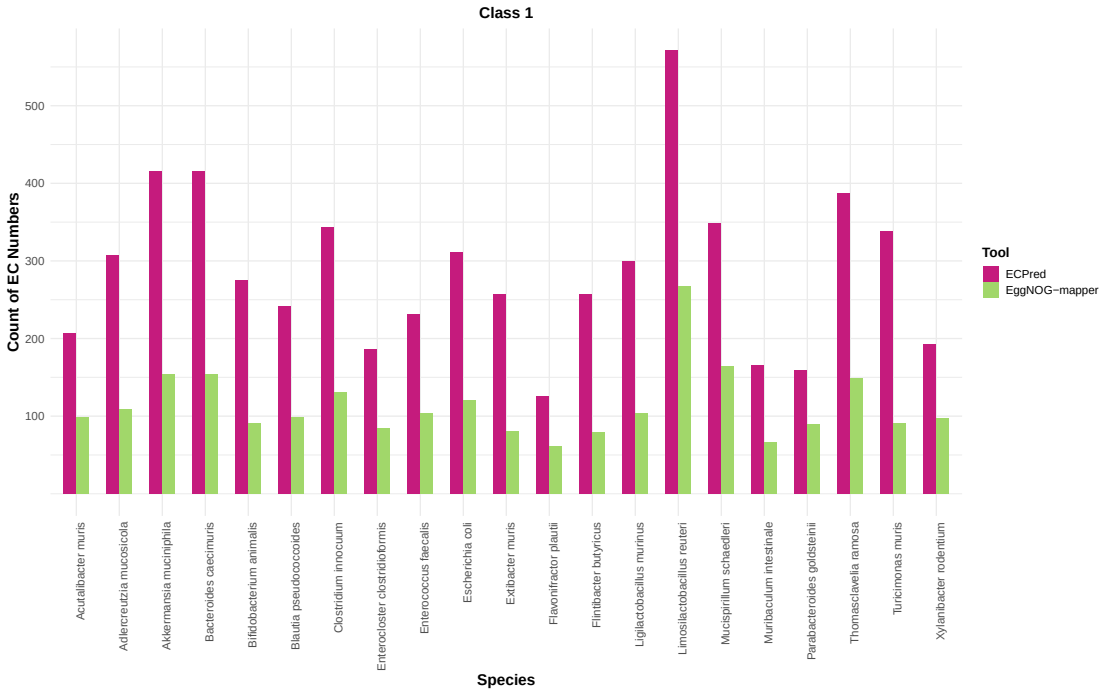

Figure 3.2.: Number of EC 1.-.-.- class oxidoreductases identified in each OMM strain by ECPred and EggNOG-mapper.

Interestingly, a comparative analysis of EC 1.3.-.- annotated proteins across species revealed a total of 15 overlapping enzymes between *B. caecimuris* and *A. muciniphila* (Table 3.4). These overlapping enzymes indicate that both species share identical EC number annotations, and further investigation confirmed that the corresponding proteins were assigned to the same bacterial species in both datasets. This degree of interspecies overlap within such an enzymatic class was unexpected. The shared proteins include several functionally annotated reductases such as cinnamate reductase, enoyl-[acyl-carrier-protein] reductase, and dihydroorotate dehydrogenase, all of which are involved in redox-related metabolic processes. The presence of multiple tRNA-dihydrouridine synthases and glutamate synthase subunits further suggests a conserved role in fundamental redox and biosynthetic pathways. The identification of shared candidates not only highlights potential evolutionary or functional convergence but also strengthens their candidacy as UroR enzymes by virtue of their cross-species conservation. Moreover, the reduced candidate space, further narrowed by this overlap, offers a practical advantage, enabling prioritisation of enzymes that are both conserved and functionally plausible for experimental validation.

To validate the differences observed between the annotation tools, a comparative analysis was performed to identify overlapping protein annotations between ECPred and EggNOG-mapper. This comparison confirmed that ECPred annotated a greater number of proteins within the EC 1.3.-.- class. For subsequent analysis, only proteins with an assigned EC number from either

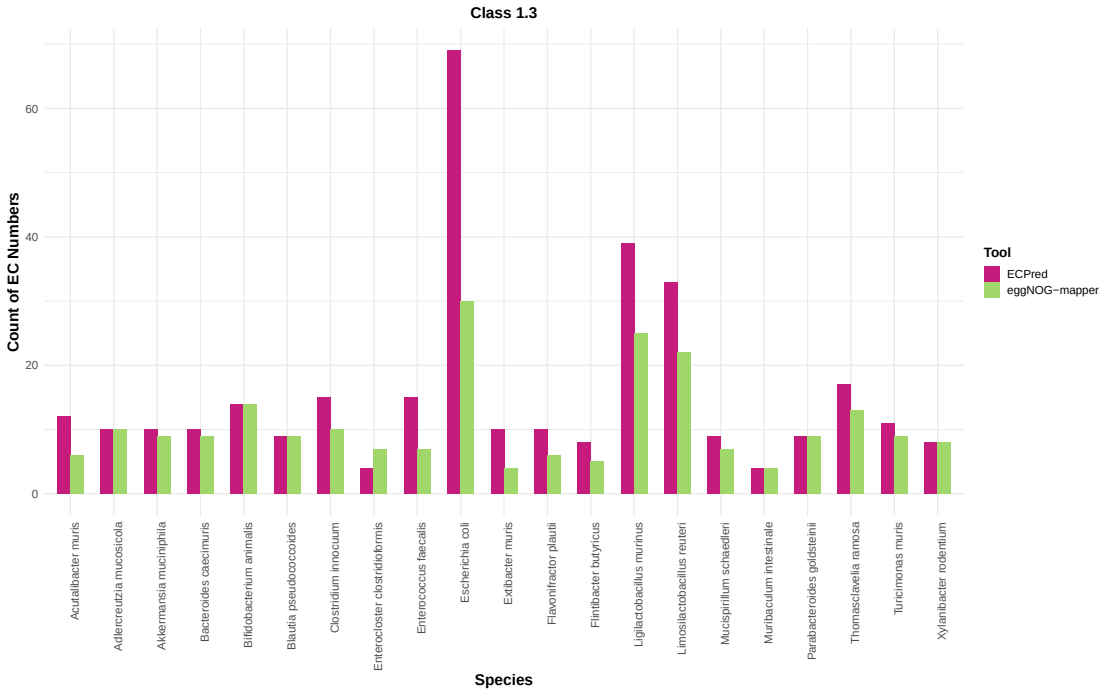

Figure 3.3.: EC 1.3.-.- class annotations, oxidoreductases acting on CH–CH group donors, across the OMM consortium.

tool were considered. Entries without EC annotations were excluded, as they could not be confidently associated with enzymatic function. Given the objective of identifying functionally characterised redox enzymes, retaining only annotated proteins ensured a focused and biologically meaningful candidate set for UroR identification.

Interestingly, it was found that ECPred consistently assigned more EC numbers than EggNOG-mapper across the majority of species. This result was unexpected, given the relative age and underlying databases of the two tools. ECPred, developed in 2017, is based on a dataset comprising 5,254 EC numbers (24). However, it utilises an outdated version of BLAST (BLAST+ 2.7.1, released in October 2017), raising questions about the completeness and currency of its annotation database. In contrast, EggNOG-mapper, released in 2023, integrates normalised functional annotations from multiple sources (including KEGG, PFAM, and SMART) and spans around 5,000 EC numbers based on KEGG/UniProt mappings (25). Based on the more recent development and broader reference base of EggNOG-mapper, it had been anticipated that it would annotate a larger number of proteins with EC numbers. However, the results suggest the opposite. This discrepancy could be attributed either to ECPred's underlying BLAST database containing a greater number of EC-associated sequences or to EggNOG-mapper applying stricter criteria for EC number assignment, resulting in a more conservative output.

Table 3.4.: List of overlapping enzymes between *B. caecimuris* and *A. muciniphila*

Protein ID	Annotation
DDAOKNCB_00053	Cinnamate reductase
DDAOKNCB_00840	3-dehydro-bile acid delta(4,6)-reductase
DDAOKNCB_01203	Glutamate synthase [NADPH] small chain
DDAOKNCB_01516	tRNA-dihydrouridine(16) synthase
DDAOKNCB_01876	Dihydroorotate dehydrogenase B (NAD(+))
DDAOKNCB_01999	Enoyl-[acyl-carrier-protein] reductase [NADH]
DDAOKNCB_02444	Ferredoxin–NADP reductase
DDAOKNCB_02815	UDP-N-acetylenolpyruvoylglucosamine reductase
DDAOKNCB_03489	tRNA-dihydrouridine synthase B
DDAOKNCB_03696	Glutamate synthase [NADPH] small chain
DDAOKNCB_00789	Precorrin-2 dehydrogenase
DDAOKNCB_00790	Porphobilinogen deaminase
DDAOKNCB_01253	Hypothetical protein
DDAOKNCB_01378	Cinnamate reductase
DDAOKNCB_03014	Hypothetical protein

Potential UroR enzymes within the OMM21 consortium were successfully identified and filtered, resulting in fewer than 10 candidate proteins in most bacterial species. This substantial reduction in complexity renders downstream experimental validation and functional characterisation considerably more feasible, enabling a focused search for a true UroR enzyme. Together, the results from sequence homology searches and EC-based annotation analyses enabled a systematic narrowing of candidate urobilinogen reductases across the OMM21 consortium. The integration of BLASTP-based screening with enzyme classification filtering identified several strains harbouring high-confidence homologues of known BilR sequences—most notably *F. butyricus* and *T. ramosa*, both of which displayed perfect matches with canonical BilR enzymes. Subsequent EC-based filtering further reduced the candidate pool, particularly within the EC 1.3.-.- subclass, to a level amenable to downstream experimental investigation. Comparative annotation revealed tool-specific biases: ECPred consistently annotated more proteins with EC numbers than EggNOG-mapper, likely due to differences in database currency and stringency of assignment criteria. Notably, the discovery of 15 overlapping enzymes between *B. caecimuris* and *A. muciniphila* points to conserved redox-related functions across phylogenetically distant strains. Collectively, these findings provide a focused and plausible set of UroR candidates, establishing a solid foundation for future experimental validation and mechanistic characterisation.

To complement the bioinformatic analyses, a literature review was conducted to determine whether any OMM21 strains or other gut-associated bacteria have previously been reported to reduce urobilinogen. However, this search did not yield any relevant findings. No published evidence was identified linking the investigated strains, or other known gut commensals, to enzymatic activity involved in urobilinogen reduction.

3.2. Development of a Scalable Bilirubin Biosensor

3.2.1. High Efficiency Transformation Protocol

To verify successful transformation and confirm the presence of the pMAL-c5x-*UnaG* plasmid in the engineered *E. coli* strains, plasmid DNA was purified and submitted for sequencing. The resulting sequences were aligned using Clustal Omega (EMBL-EBI, version 1) and visualised with MView (version 2) (27; 28).

MView automatically calculates both percent coverage and percent identity for each aligned sequence relative to the reference. Here, the full plasmid sequence served as the reference. Percent coverage indicates the proportion of the reference sequence aligned with each query sequence, while percent identity reflects the nucleotide similarity within the aligned region (28).

Both experimental plasmid sequences (Plasmid 1 and Plasmid 2) showed high coverage (99.8%) of the reference plasmid, confirming successful sequencing across nearly the entire plasmid length. Plasmid 1 exhibited 87.4% identity with the reference sequence, while Plasmid 2 showed 84.0% identity. The control plasmid sequence, which contained only the gene region of interest, aligned with 99.8% coverage but showed a much lower identity (16.1%), as expected due to its limited sequence content.

The alignment results are summarised in Table 3.5. Both experimental plasmid sequences (Plasmid 1 and Plasmid 2) demonstrated high coverage (99.8%) of the reference sequence, confirming that the targeted region of the plasmid (positions 2686–3069) was successfully sequenced. Plasmid 1 exhibited 87.4% identity with the reference, while Plasmid 2 showed 84.0% identity. These values are consistent with partial sequencing, as only a small region of the full plasmid was analysed. The control plasmid, which included only the gene of interest (positions 2686–3069), also achieved 99.8% coverage but showed a lower identity (16.1%), which was expected due to its limited sequence length and content relative to the full reference plasmid.

These results confirmed the presence and integrity of the *UnaG*-containing plasmid in both experimental strains, validating the success of the transformation protocol and the suitability of the constructs for downstream applications.

Table 3.5.: Summary of *UnaG* plasmid sequence alignment results (Clustal Omega, visualised with MView).

Plasmid	Percent Coverage	Percent Identity
Full Plasmid	100.00%	100.00%
Control Plasmid	99.8%	16.1%
Plasmid 1	99.8%	87.4%
Plasmid 2	99.8%	84.0%

3.2.2. Determination of IPTG Levels

The effect of various IPTG concentrations on *E. coli* viability was assessed to identify a non-toxic inducer concentration for subsequent experiments and to determine the most appropriate concentration for continued use. Concentrations ranging from 0.1 mM to 4 mM IPTG were tested, based on values reported in the literature, alongside a non-induced control for baseline comparison.

Growth curves measured over 24 hours revealed no substantial inhibitory effect across the tested IPTG concentrations. As shown in Figure 3.4, absorbance values at 600 nm remained comparable to the control between 12 and 24 hours, stabilising within the range of 0.3 to 0.4. This indicated that none of the IPTG concentrations exerted observable toxic effects on bacterial growth.

Moreover, no significant differences in growth were observed between the two *E. coli* cultures used in the assay, suggesting consistent responses to IPTG induction across both cultures. This confirmed that the selected concentration range did not introduce strain-dependent variability and could be applied broadly without compromising viability.

Based on these observations, IPTG concentrations between 1 mM and 4 mM were identified as particularly suitable for subsequent experiments. Both concentrations were well tolerated by the bacterial cultures and were considered likely to ensure robust induction of *UnaG* expression throughout the population.

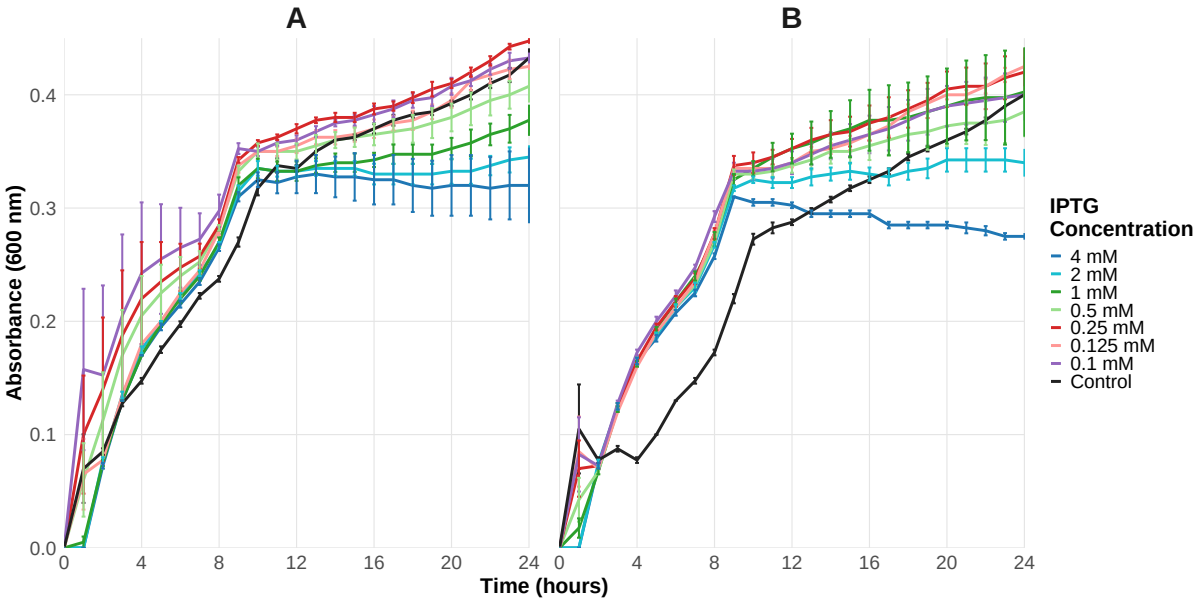

Figure 3.4.: Growth curves of *E. coli* cultures exposed to varying IPTG concentrations over 24 hours. Absorbance at 600 nm was used to monitor cell density.

3.2.3. Testing of Bilirubin and Analogues on *E.coli* Growth

Following the identification of 1 mM and 4 mM as suitable IPTG concentrations that did not adversely affect *E. coli* viability, the range and potential toxicity of the metabolites bilirubin, urobilin, and stercobilin were evaluated. Bilirubin was tested at three concentrations (12 µM, 1.2 µM, and 0.012 µM), alongside 12 µM of urobilin and stercobilin, in both IPTG-induced and non-induced *E. coli* cultures. A control condition containing no IPTG and no *E. coli* was included to serve as a background reference, enabling correction for absorbance contributions from the media, metabolites, or solvent, and to confirm the absence of contamination during assay preparation.

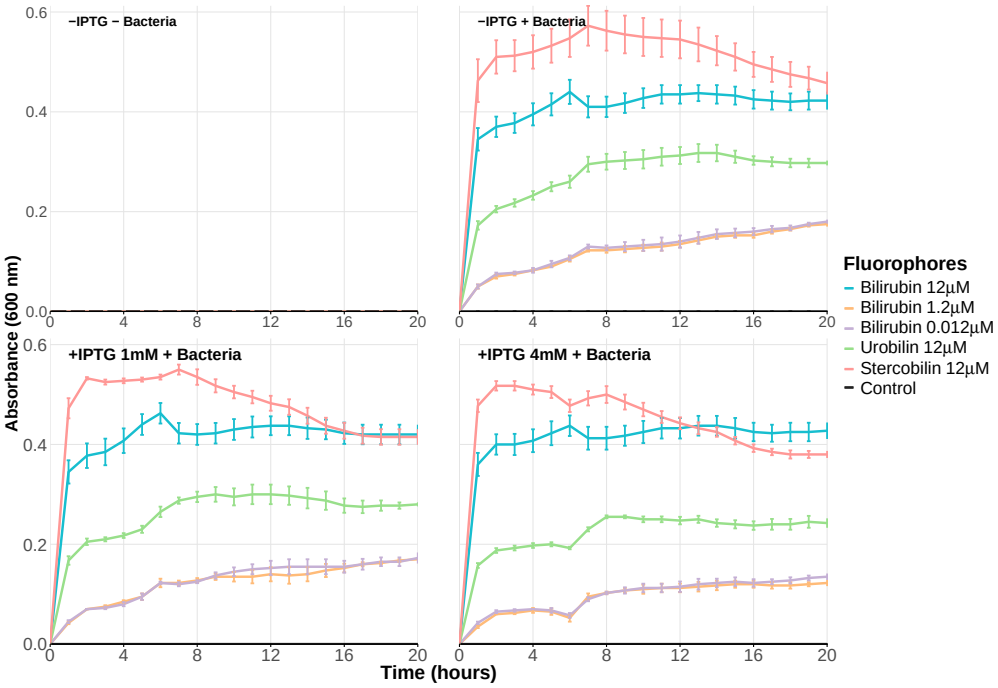

Figure 3.5.: Growth curves of *E. coli* cultures exposed to bilirubin, stercobilin, and urobilin under IPTG-induced and non-induced conditions. Absorbance at 600 nm was monitored over 20 hours to assess potential toxicity.

As shown in Figure 3.5, none of the tested metabolite conditions resulted in a significant reduction in bacterial growth relative to the non-induced control. Across all conditions, absorbance at 600 nm increased progressively over time, reaching similar levels by the end of the 20-hour incubation period. The subsequent plateau or slight decline observed in some growth curves is likely attributable to nutrient depletion and overgrowth, as the bacterial cultures reached high densities and exhausted the available medium. These findings indicate that neither bilirubin, at any of the tested concentrations, nor the other metabolites (urobilin and stercobilin) exerted detectable toxic effects on *E. coli* under the given experimental conditions.

In addition, the results confirmed that the presence of IPTG, whether at 1 mM or 4 mM, did not modulate the toxicity profile of the metabolites. Growth curves from IPTG-induced conditions were indistinguishable from those without IPTG, suggesting that IPTG-mediated expression of UnaG did not exacerbate sensitivity to any of the metabolites.

These findings are consistent with concentrations previously reported in the literature. Prior studies employing bilirubin concentrations ranging from 17.1 µM to 0.095 nMin *E. coli*-based UnaG assays did not report adverse effects on bacterial viability (4; 16; 29; 32; 33; 34). The present results corroborate these findings, confirming the suitability of these metabolites for further use in live-cell fluorescence assays.

3.2.4. Functional Evaluation of the IPTG–*UnaG* Fluorescence System

Having established a concentration range in which both IPTG and the selected metabolites (bilirubin, stercobilin, and urobilin) were non-toxic to *E. coli* and mutually compatible, the next objective was to evaluate the functional induction of fluorescence via the IPTG–*UnaG* system. This evaluation aimed to determine whether the addition of IPTG could effectively induce *UnaG* expression, thereby producing a fluorescence signal upon binding to bilirubin. In parallel, fluorescence from the metabolites alone, as well as potential background signals from *E. coli* cultures, were measured to confirm the specificity and reliability of the observed signal.

As illustrated in Figure 3.6, fluorescence emission was detected exclusively in *E. coli* cultures harbouring the *UnaG*-encoding plasmid when induced with either 1 mM or 4 mM IPTG in the presence of bilirubin. No detectable fluorescence was observed in cultures containing the same components but lacking the plasmid, thereby confirming the specificity of signal generation through IPTG-induced *UnaG* expression. This also served as independent functional confirmation that the recombinant *E. coli* strain retained the correct plasmid.

Furthermore, controls containing only metabolites (bilirubin, stercobilin, or urobilin) in the absence of IPTG and/or plasmid exhibited negligible background fluorescence, supporting the specificity and inducibility of the system. Consistent with earlier absorbance-based viability data, stercobilin and urobilin did not interfere with bacterial metabolism and similarly showed minimal fluorescence signal, confirming their lack of cross-reactivity under the tested conditions.

These results, in combination with the previous viability assays, demonstrate that fluorescence observed arises specifically from the IPTG-mediated induction of *UnaG* and its interaction with bilirubin, validating the effectiveness of the IPTG–*UnaG* system for further use in live-cell fluorescence assays.

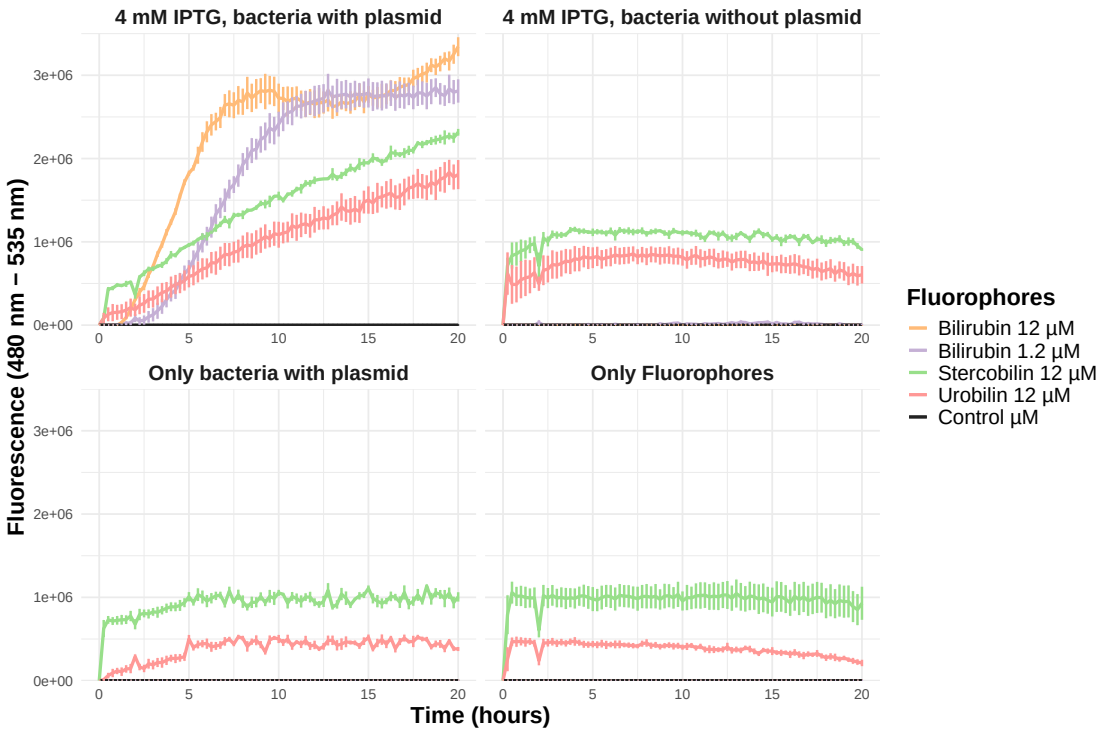

Figure 3.6.: Fluorescence signal generated by *E. coli* cultures under various induction and metabolite conditions. Fluorescence (480 nm excitation / 535 nm emission) was recorded over time to confirm IPTG-mediated *UnaG* expression and functional interaction with bilirubin.

3.2.5. Kinetic and Sensitivity Assessment of the IPTG–*UnaG* System

To optimise IPTG induction and maximise fluorescence from the *UnaG*–bilirubin interaction, different induction times were assessed. Cultures were induced with 2 mM IPTG and incubated for 0, 2, 4, or 18 hours before metabolite addition. The aim was to determine the minimum induction time required for sufficient *UnaG* expression and effective fluorescence generation upon bilirubin binding.

As shown in Figure 3.7, no measurable fluorescence signal was observed at the 0 hour induction point, indicating that immediate addition of bilirubin did not allow sufficient time for *UnaG* synthesis and interaction. In contrast, both 2 hour and 4 hour induction periods produced strong fluorescence signals, with the 4 hour condition yielding the highest intensity. This suggests enhanced *UnaG* expression and/or binding efficiency at longer induction times. Notably, no fluorescence was observed at the 18 hour time point, which may be attributed to degradation of *UnaG*, bilirubin instability, or metabolic breakdown during prolonged incubation.

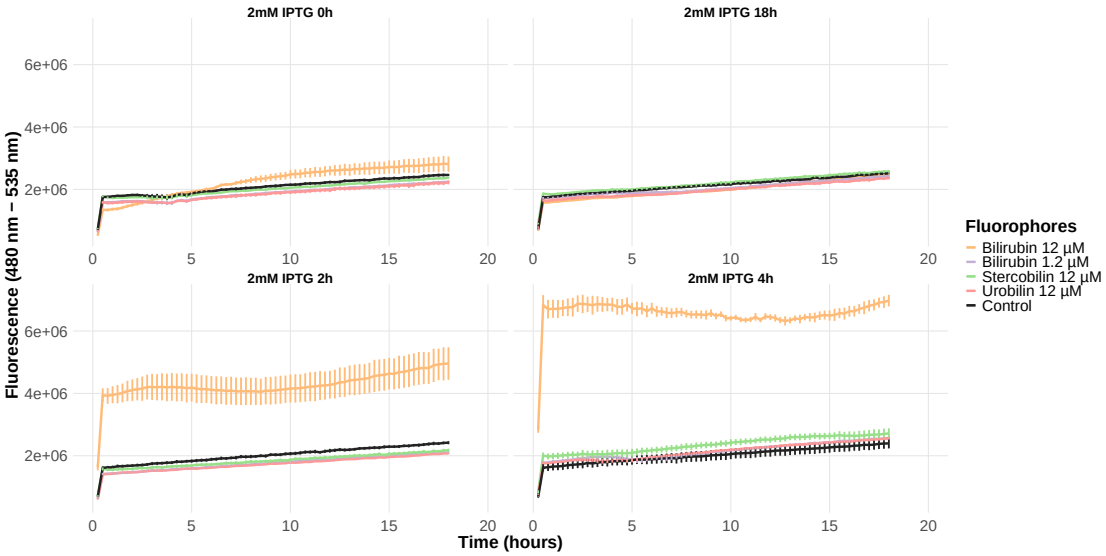

Figure 3.7.: Fluorescence intensity following IPTG induction of *UnaG* expression for 0, 2, 4, and 18 hours. Bilirubin was added after the indicated induction periods. Fluorescence was monitored over time.

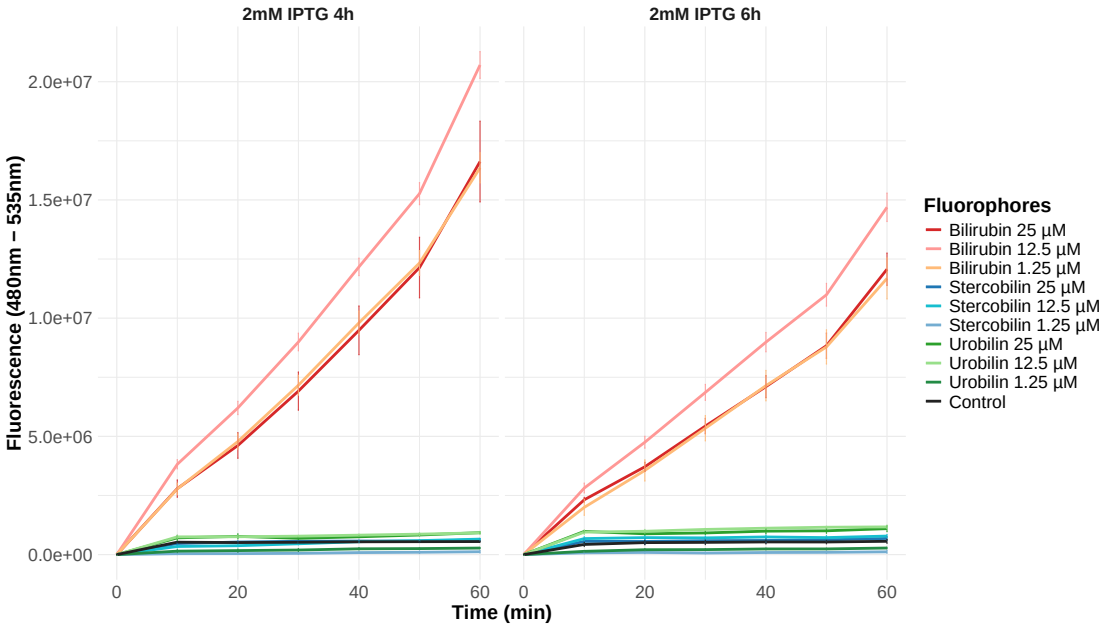

Figure 3.8.: Comparison of fluorescence following 4 hour and 6 hour IPTG induction. Fluorescence was measured at varying concentrations of bilirubin, stercobilin, and urobilin (1.25–25 µM) to assess detection range and background interference.

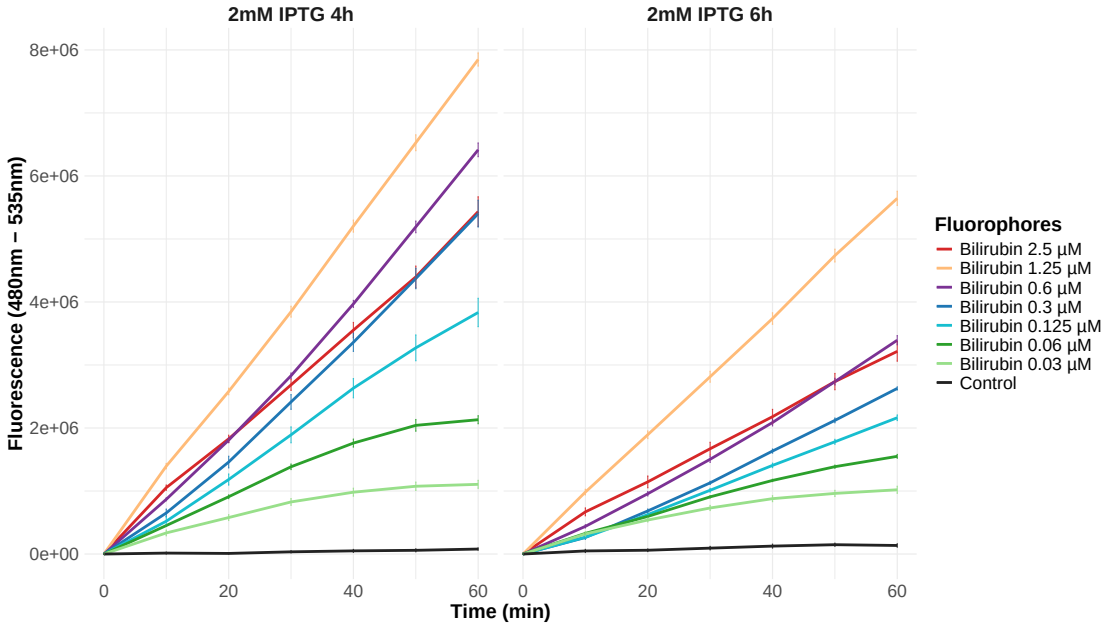

Figure 3.9.: Lower detection limit assessment for bilirubin using the *UnaG* fluorescence system. Bilirubin concentrations from 2.5 µM to 0.03 µM were tested under 4 hour and 6 hour IPTG induction conditions.

Based on these results, a 4 hour induction period was selected as optimal for achieving maximal fluorescence. To assess moderate extension of induction without reaching the detrimental 18 hour duration, 4 hour and 6 hour IPTG induction times were compared (Figure 3.8). Additionally, this experiment evaluated the upper and lower metabolite detection limits and assessed whether higher concentrations of stercobilin and urobilin would contribute to background fluorescence, as initially found to be negligible in Figure 3.6.

Consistent with previous observations (18h induction, Figure 3.7)), the 6 hour induction resulted in a marked reduction in fluorescence compared to the 4 hour condition (Figure 3.8), supporting the hypothesis that extended induction negatively affects the fluorescence signal. The 4 hour induction again produced the strongest signal, reaffirming it as the optimal duration for reliable *UnaG*–bilirubin fluorescence. Moreover, bilirubin remained highly detectable at 25 µM, exceeding the highest literature-reported concentration (17.1 µM, (33)), and was still clearly measurable at 1.25 µM. In contrast, stercobilin and urobilin, even at elevated concentrations up to 25 µM, generated only minimal fluorescence, confirming their negligible interference and validating the specificity of the *UnaG*–bilirubin system.

To assess the lower detection limit, an additional experiment was performed using bilirubin concentrations ranging from 2.5 µM down to 0.03 µM, again under both 4 hour and 6 hour IPTG induction conditions (Figure 3.9). Fluorescence was detectable at concentrations as low

as 0.03 µM under the 4 hour induction condition, whereas the 6 hour induction produced lower overall signal intensities across all concentrations. These findings demonstrated that the system is sufficiently sensitive to detect very low levels of bilirubin–*UnaG* interaction, and further confirmed that 4 hours is the optimal induction period for achieving both maximal signal strength and sensitivity.

3.2.6. Validation of the IPTG–*UnaG* System

To verify the presence and integrity of the pSF-OXB15-*BilR* plasmid in the engineered *E. coli* strains, plasmid DNA was extracted from the BilR-WT and BilR-CD cultures and submitted for full plasmid sequencing. Three additional strains from the *E. coli* K-12 Keio collection—comprising the wild-type parental strain and two isogenic *gusA* knockouts—were processed in parallel. As the K-12 Keio strains do not contain plasmids, NanoDrop analysis was used solely to confirm the absence of detectable plasmid DNA. The parental Keio strain yielded a low DNA concentration of 5.32 ng μL^{-1}, consistent with the expectation for plasmid-free cells.

Sequencing results for the BilR-WT and BilR-CD strains were received as FastQ files. These files were aligned with the full reference plasmid sequence using Clustal Omega (EMBL-EBI, version 1), and the alignment was visualised using MView (version 2) (27; 28). Percent coverage (the proportion of the reference sequence aligned) and percent identity (the nucleotide similarity across the aligned region) were automatically calculated by MView.

The BilR-CD strain showed full alignment with the BilR-CD reference sequence (100.0% coverage and 100.0% identity), confirming the successful transformation and correct plasmid sequence. When aligned to the full plasmid reference, the experimental BilR-CD plasmid sequence displayed 90.8% coverage and 36.2% identity, indicating partial alignment and divergence in sequence, as expected for a catalytically inactivated variant.

Similarly, the BilR-WT strain showed 100.0% coverage and 100.0% identity with the BilR-WT reference sequence, confirming complete fidelity. The corresponding experimental plasmid exhibited 85.9% coverage and 65.6% identity when aligned to the full plasmid, again consistent with expected variation and partial sequencing success. These results confirm that the correct plasmids were present and distinguishable in the BilR-WT and BilR-CD strains, supporting their validity for use in downstream IPTG–*UnaG* functional assays.

To validate the functionality and specificity of the IPTG–*UnaG* system for bilirubin detection in a biological context, fluorescence assays were performed using *E. coli* strains harbouring either the wild-type *BilR* gene (BilR-WT) or a catalytically inactive version (BilR-CD) containing point mutations. The experimental hypothesis was that, following overnight incubation with bilirubin, the BilR-WT strain would metabolise bilirubin via active BilR, leading to reduced fluorescence relative to the input concentration. In contrast, the BilR-CD strain, serving as a negative control, was expected to show fluorescence corresponding to the full amount of bilirubin added, as no active consumption was anticipated.

Table 3.6.: Summary of *BilR* plasmid sequence alignment results (Clustal Omega, visualised with MView).

Strain	Percent Coverage	Percent Identity
BilR-CD (reference alignment)	100.0%	100.0%
BilR-CD (plasmid alignment)	90.8%	36.2%
BilR-WT (reference alignment)	100.0%	100.0%
BilR-WT (plasmid alignment)	85.9%	65.6%

However, the results presented in Figure 3.10 (and replicated in the Supplementary Information A, Figure A.1) did not align with the initial expectations. Fluorescence across all experimental conditions was markedly low, despite the addition of 0.6 µM bilirubin, and substantially below the signal observed in the corresponding bilirubin standard curve.

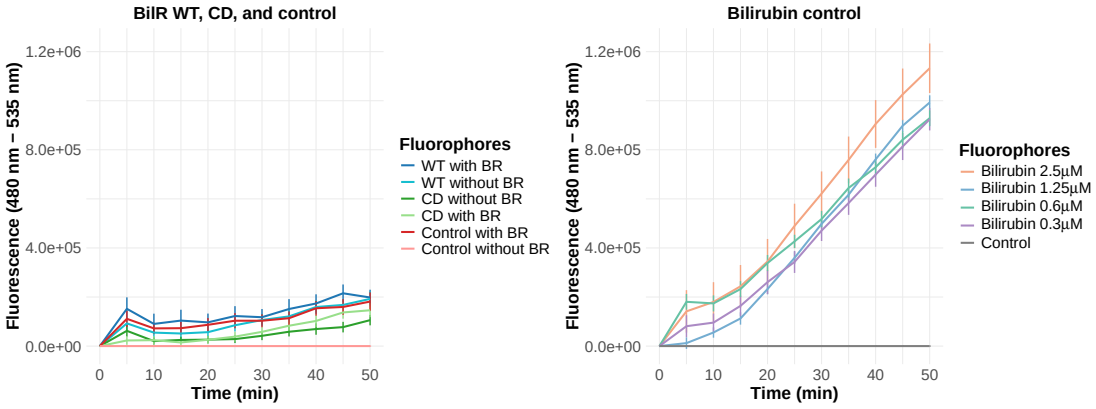

Figure 3.10.: Fluorescence intensity measurements of *E. coli* strains containing either wild-type (BilR-WT) or catalytically dead (BilR-CD) versions of the *BilR* gene, with or without added bilirubin (BR). A control with and without BR and a bilirubin concentration reference series are also shown.

The BilR-WT strain incubated with bilirubin displayed a fluorescence signal nearly identical to the same strain incubated without bilirubin. Unexpectedly, it also showed marginally higher fluorescence than its catalytically inactive counterpart (BilR-CD), which is the opposite of the anticipated trend.

Furthermore, the BilR-CD strain with bilirubin exhibited fluorescence levels well below the expected profile, while its counterpart without bilirubin also showed measurable fluorescence, despite no metabolite being added. Similarly, both BilR-WT and BilR-CD strains without added bilirubin produced detectable fluorescence signals, contrary to the expectation of no signal in the absence of the metabolite. The control condition with bilirubin also showed

only minimal fluorescence, far below what would be anticipated based on the bilirubin-only calibration.

Similarly, both BilR-WT and BilR-CD strains without added bilirubin exhibited measurable fluorescence, contrary to the expectation of no signal in the absence of the metabolite. The control with bilirubin also showed minimal signal, well below the levels observed in the bilirubin-only standard curve.

Collectively, these results indicate substantial inconsistencies. Fluorescence was observed in conditions lacking bilirubin, while an insufficient signal was detected under conditions expected to produce strong fluorescence. These discrepancies may be attributed to experimental variability, poor uptake or retention of bilirubin, inefficient binding to *UnaG*, cellular background interference, or other unknown factors that may suppress or distort the fluorescence output.

To further test the IPTG–*UnaG* system in a biological context, strains from the *E. coli* K-12 Keio collection was employed to validate functionality using conjugated bilirubin, bilirubin ditaurate (BDT). In this assay, it was hypothesised that BDT would be converted into unconjugated bilirubin by GusA, a β-glucuronidase present in wild-type *E. coli*, enabling detection via *UnaG*-mediated fluorescence. The parental Keio strain, which expresses GusA, was therefore expected to produce a fluorescence signal upon BDT addition. In contrast, the two GusA knockout mutants (mutant 1 and mutant 2) were expected to show no fluorescence due to their inability to convert BDT into detectable bilirubin.

As shown in Figure 3.11 and also repeated in Figure A.2, fluorescence levels were generally low across all samples. The Keio parental strain incubated with BDT did not exhibit a notable increase in fluorescence compared to its non-BDT control. Unexpectedly, both GusA mutants, regardless of BDT treatment, showed higher fluorescence than the parental strain. Moreover, the mutants without BDT addition also displayed elevated fluorescence, which contradicts the assumption that no signal should be generated in the absence of the metabolite.

Despite these irregularities, the positive controls behaved as expected, demonstrating a clear fluorescence signal. The control with unconjugated bilirubin also produced measurable fluorescence, albeit at a slightly reduced level—likely due to the 18-hour incubation period.

These results suggest that while the IPTG–*UnaG* system can detect bilirubin under ideal conditions, its performance in biological samples may be hindered by experimental variability, nonspecific background fluorescence, or incomplete conversion of BDT.

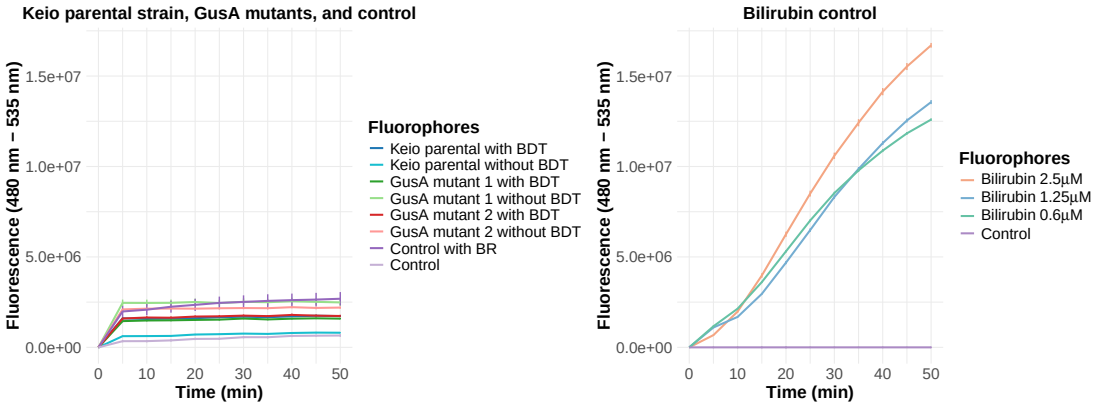

Figure 3.11.: Fluorescence measurements of Keio parental and GusA knockout strains with or without bilirubin ditaurate (BDT). Unexpectedly, both mutants exhibited higher fluorescence than the wild-type strain, regardless of BDT treatment.

3.2.7. Investigation of Bilirubin Degradation, Antibiotic and Bilirubin Ditaurate Effects

Due to the unexpectedly low fluorescence signals observed in the validation experiments, it was hypothesised that bilirubin degradation, known to be accelerated by light exposure due to its photosensitivity (37; 38), could be contributing to the weak fluorescence output. Additionally, it was considered that the presence of the antibiotic kanamycin may exacerbate this degradation.

To test the effect of light exposure on bilirubin stability, samples were incubated for 2 hours and 4 hours under either covered or uncovered conditions (Figure 3.12). A clear reduction in fluorescence was observed in uncovered samples for both time points, confirming that light exposure contributes to bilirubin degradation. However, this reduction was not sufficient to explain the severely diminished fluorescence previously observed in the BilR-CD and BilR-WT assays (Figure 3.10).

To assess the effect of kanamycin, additional bilirubin controls were incubated with the antibiotic in media, without cells. Following incubation, *UnaG*-expressing *E. coli*—which lacks a kanamycin resistance cassette—was added. As shown in Figure 3.12, samples with kanamycin exhibited lower fluorescence compared to bilirubin-only controls, suggesting that kanamycin may indeed contribute to degradation. Although the effect was moderate, it warranted further investigation.

A subsequent experiment was performed to explore whether diluting kanamycin could mitigate this effect (Figure 3.13). Bilirubin was tested at multiple concentrations (6, 3, and 1.5 µM) in the presence of undiluted and diluted kanamycin (1:3 and 1:4 in LB medium). Among these, the 1:3 dilution at 6 µM bilirubin yielded a notably stronger fluorescence

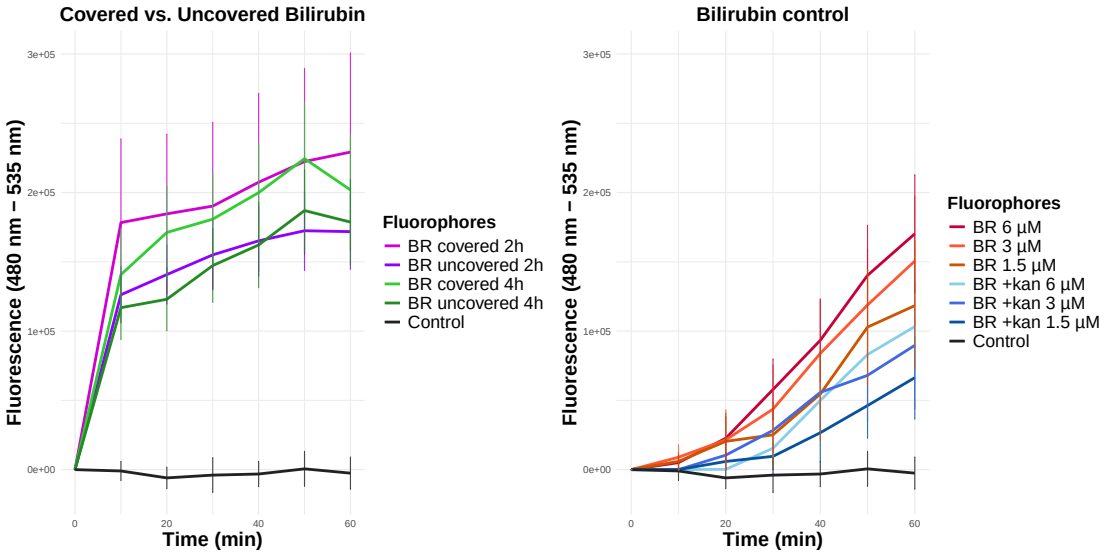

Figure 3.12.: Effect of light exposure and kanamycin on bilirubin fluorescence. Fluorescence was compared between covered and uncovered samples at 2 h and 4 h, and in the presence or absence of kanamycin across various bilirubin concentrations.

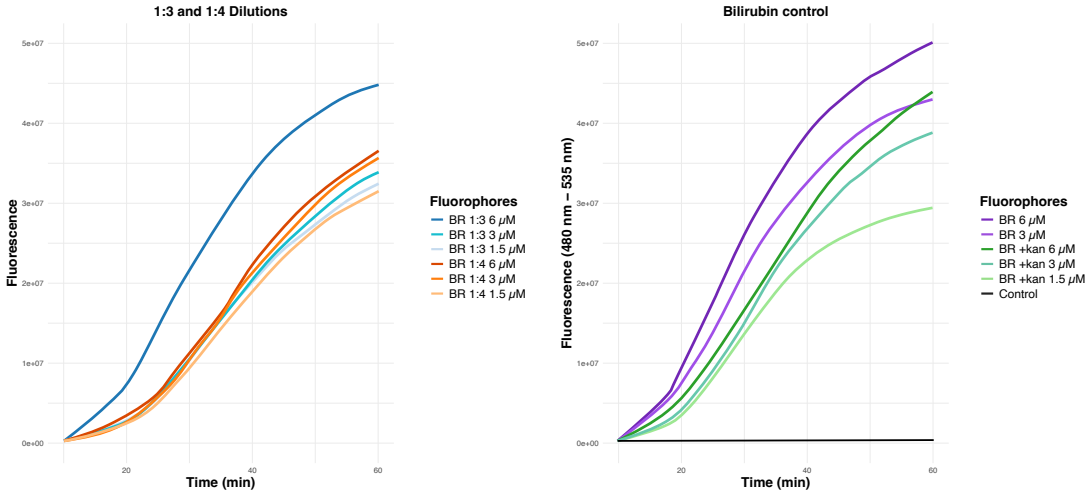

Figure 3.13.: Fluorescence intensity of bilirubin samples in the presence of diluted kanamycin (1:3 and 1:4), compared to undiluted kanamycin and bilirubin-only controls.

signal, indicating reduced degradation. However, the remaining dilutions produced results comparable to the kanamycin-positive control, suggesting that while dilution may help under certain conditions, it does not fully account for the signal loss observed in earlier experiments.

Another variable considered was incubation time. Given that the BilR validation experiments were conducted after overnight incubation, it was hypothesised that prolonged exposure may significantly reduce bilirubin concentration. To test this, bilirubin fluorescence was measured after 1, 2, 4, 6, and 18 hours of incubation (Figure 3.14). A marked decline in fluorescence was observed as early as 2 hours, with only minimal signal remaining after 18 hours. These findings suggest that prolonged incubation results in extensive bilirubin degradation, likely explaining part of the signal loss in the BilR-CD and BilR-WT assays.

While this does not explain all anomalies—particularly the presence of fluorescence in samples where no bilirubin was added—it strongly indicates that bilirubin degradation over extended incubation times is a significant contributing factor to the observed results. Minimising light exposure and reducing incubation time are therefore crucial for preserving signal integrity in future applications of the IPTG–*UnaG* system.

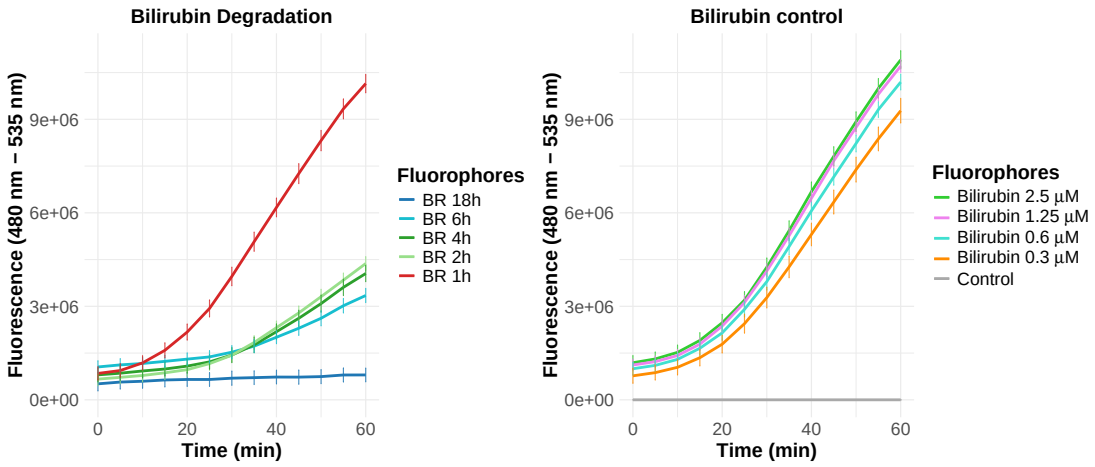

Figure 3.14.: Bilirubin degradation over time. Fluorescence was monitored after 1, 2, 4, 6, and 18 hours of incubation. A strong time-dependent decrease in fluorescence is evident, with minimal signal remaining at 18 hours.

To assess whether bilirubin ditaurate (BDT) exerted any toxic or inhibitory effects on *E. coli*, an experiment was conducted using both the *UnaG*-expressing strain and the K-12 Keio strains. It was hypothesised that BDT might impair cell viability or metabolism, potentially accounting for the low and inconsistent fluorescence signals previously observed (Figure 3.11).

However, as shown in Figure 3.15, BDT did not exhibit any observable toxic effect on either *UnaG* or Keio *E. coli*. On the contrary, cultures supplemented with 25 µM BDT appeared to be slightly more viable than those treated with 12.5 µM, suggesting no detrimental impact at either concentration. These findings indicate that BDT is unlikely to be the cause of the weak and anomalous fluorescence signals observed in previous experiments.

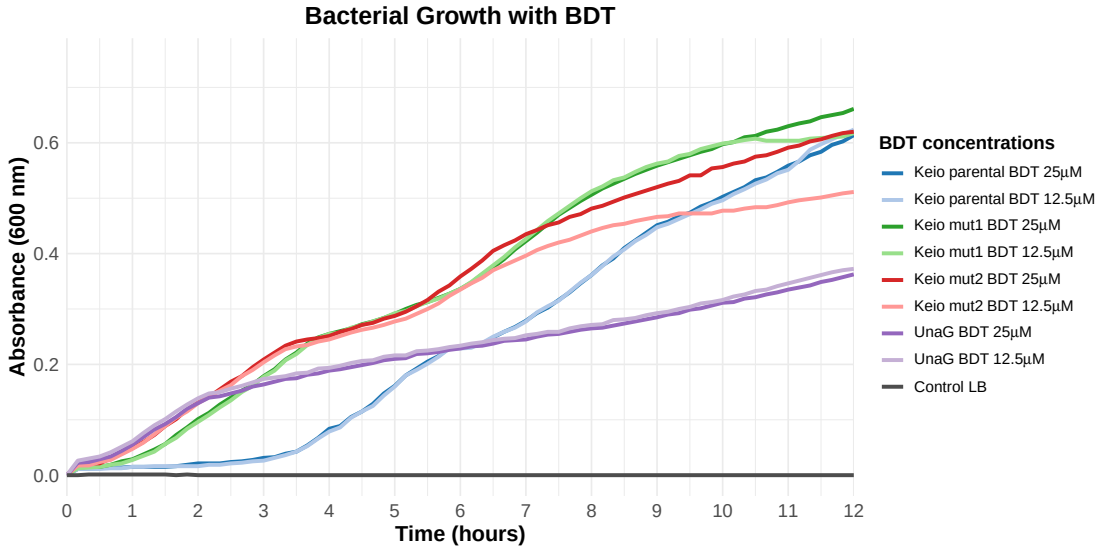

Figure 3.15.: Growth assessment of *E. coli* strains in the presence of bilirubin ditaurate (BDT). Cultures were exposed to either 12.5 µM or 25 µM BDT. No growth inhibition was observed at either concentration, indicating that BDT is non-toxic under these conditions.

4. Discussion

4.1. Bioinformatic Identification of Candidate Urobilinogen Reductases

This study presents a comprehensive bioinformatic approach to identify candidate urobilinogen reductases (UroR) within the Oligo Mouse Microbiome (OMM21) consortium. Through targeted sequence homology searches, functional annotation pipelines, and systematic filtering of oxidoreductase classes, a refined list of plausible UroR candidates has been established. The integration of multiple methodological layers—including literature review, BLASTP screening, and EC-based classification—provides both breadth and specificity to the search for enzymes involved in urobilinogen reduction.

The identification of homologues to known bilirubin reductases (BilR) across the OMM21 strains highlights the widespread presence of reductive capabilities in the consortium. Notably, perfect matches for both short and long BilR variants in *Flintibacter butyricus*, *Escherichia Coli* and *Thomasclavelia ramosa* suggest strong evolutionary conservation of these reductases. These strains emerge as leading candidates for further functional validation, potentially representing previously uncharacterised sources of UroR activity.

Discrepancies between the short and long BilR homology searches underscore the importance of using multiple reference sequences to capture sequence divergence across species. Some strains lacking hits in one search yielded strong homologues in another, indicating possible variations in the structural domains or functional motifs of BilR-like enzymes across taxa. This observation further justifies the complementary use of both short- and long-form query sequences.

The EC-based annotation pipeline enables an orthogonal line of evidence, independent of direct sequence homology. Class 1 oxidoreductases (EC 1.-.-.-) and particularly the CH–CH acting subclass (EC 1.3.-.-) were targeted due to their relevance in redox transformations. The combination of ECPred and EggNOG-mapper ensured robustness, albeit with striking disparities in annotation coverage. Contrary to expectations, ECPred—a tool based on an older dataset—consistently identified more EC-annotated proteins than the more recent EggNOG-mapper. This counterintuitive result may stem from differences in annotation stringency or underlying database breadth and should be considered when interpreting comparative outputs.

A particularly valuable outcome of the annotation-based screening was the identification of 15 overlapping oxidoreductases between *Bacteroides caecimuris* and *Akkermansia muciniphila*. This convergence among phylogenetically distant strains suggests potential conservation of function and strengthens the candidacy of these shared enzymes. However, closer inspection revealed that 13 of these enzymes are well-characterised with established functions, while the remaining two are annotated as hypothetical proteins with unknown roles. This implies that, if either species harbours a candidate UroR enzyme, it is most likely among the hypothetical proteins or corresponds to a protein not captured by either ECPred or eggNOG-mapper (39; 40; 41; 42; 43; 44; 45; 46; 47).

Importantly, the literature review conducted in parallel did not yield any prior reports of urobilinogen reduction by OMM21 members or other gut commensals. This absence of existing evidence further emphasises the novelty of the present study and validates the rationale for a bioinformatics-first discovery approach. The results presented here mark the first step toward establishing a functional role for gut microbes in urobilinogen reduction, a pathway of potential relevance for host-microbiome metabolic interactions.

The main limitation of this work lies in the reliance on in silico annotation and homology inference, which cannot definitively confirm enzymatic activity. While the combination of homology and EC classification provides strong predictive power, only functional assays can validate the reductive capabilities of the candidate proteins. Additionally, annotation discrepancies between tools underscore the need for cautious interpretation of in silico predictions.

Future research should prioritise experimental validation of the high-confidence candidates identified herein, particularly those with perfect or near-perfect query coverage and low E-values. Expression studies in heterologous systems or activity-based assays using synthesised urobilinogen substrates would directly confirm enzymatic function. In parallel, metatranscriptomic or proteomic profiling of OMM21 strains during co-culture or host colonisation could offer insight into the regulation and relevance of these enzymes in vivo.

In summary, this study establishes a robust and reproducible framework for identifying candidate urobilinogen reductases using bioinformatic tools. The findings not only provide a refined list of functional candidates for experimental follow-up but also pave the way for uncovering novel metabolic interactions between gut microbes and host bile pigment metabolism.

4.2. Development of a Scalable Bilirubin Biosensor

A functional and scalable biosensor system was developed in *Escherichia coli* to enable bilirubin detection through fluorescence generated by IPTG-inducible expression of the *UnaG* protein. The work combines molecular cloning, transformation, fluorescence assay development, and biological validation to characterise the system's specificity, sensitivity, and robustness under various conditions.

The transformation of *E. coli* with the *pMAL-c5x-UnaG* plasmid is confirmed through sequencing and alignment, which demonstrates high coverage and identity with the reference plasmid. The successful incorporation of the *UnaG* gene validates the foundational requirement for the biosensor platform and supports the integrity of downstream experiments.

The investigation of IPTG concentration reveals that inducer levels up to 4 mM do not exert toxic effects on *E. coli* growth, in agreement with prior literature. These results confirm that IPTG concentrations within this range are suitable for robust induction of *UnaG* without compromising cell viability. Similarly, bilirubin and its analogues, urobilin and stercobilin, are found to be non-toxic across a range of physiologically relevant concentrations, thus supporting their compatibility with live-cell assays.

Functional testing confirms that fluorescence is specifically induced in the presence of IPTG and bilirubin in strains harbouring the *UnaG* plasmid, but not in control conditions lacking either component. These findings validate the IPTG–*UnaG* system as a specific and reliable method for detecting bilirubin in live *E. coli* cultures. Importantly, stercobilin and urobilin do not elicit significant fluorescence, indicating low background interference and high target specificity of the system.

Kinetic analysis identifies 4 hours as the optimal IPTG induction time, yielding the highest fluorescence signal. Prolonged induction (e.g., 6–18 hours) leads to reduced signal intensity, possibly due to protein degradation or bilirubin instability. This finding highlights the importance of optimising induction time to maintain a balance between protein expression and fluorophore stability. Sensitivity tests demonstrate that the system can reliably detect low bilirubin concentrations, exceeding the sensitivity of similar systems reported in the literature.

However, the system encounters limitations when applied to more complex biological validation scenarios. Validation using strains expressing wild-type and catalytically inactive *BilR* reveals inconsistent and unexpectedly low fluorescence signals. These anomalies likely arise from a combination of factors, including bilirubin degradation over extended incubations, poor bilirubin uptake or retention by cells, and potential interference from cellular components or media.

Subsequent experiments investigating bilirubin degradation confirm its sensitivity to both light exposure and prolonged incubation. Notably, fluorescence intensity declines

markedly when incubating bilirubin for 4–6 hours prior to measurement and is nearly absent after 18 hours, emphasising the need for rapid measurement post-incubation. The presence of kanamycin further exacerbates bilirubin degradation, and while dilution of the antibiotic offers some mitigation, it does not completely restore fluorescence. Thus, antibiotic use and light exposure represent key variables that must be tightly controlled to ensure reliable assay performance.

These observations are consistent with the known photolability and oxidative instability of bilirubin. Under light exposure, the principal photoproduct formed is biliverdin (48; 49). In the context of neonatal phototherapy, bilirubin undergoes both structural isomerisation—forming lumirubin and various ZE/EZ configurational isomers, as well as slower photo-oxidation into colourless derivatives (50). Among these, monoglucuronylated lumirubins are particularly water-soluble and easily excreted (49; 51; 52), reducing the availability of bilirubin for *UnaG* binding.

In addition to isomerisation, chemical analyses confirm the occurrence of photooxidation in vivo, with products such as propentdyopents and methylvinylmaleimide-derived acids identified in the urine of neonates receiving phototherapy (53). These compounds were not found in untreated infants, indicating a direct role of light in promoting oxidative degradation. While most research has focused on photochemical transformations, bilirubin is also known to undergo auto-oxidation in air, particularly under aerobic and illuminated conditions, again yielding biliverdin and other breakdown products (48).

Together, these findings imply that both light-induced and spontaneous oxidation significantly reduce the stability and availability of bilirubin in vivo and in vitro conditions. This degradation likely underlies the diminished fluorescence observed in the time points and reinforces the need for carefully timed and protected assay workflows to preserve the substrate's integrity.

Validation experiments using *E. coli* K-12 strains with or without the *gusA* gene produced inconsistent and counter-intuitive results. Fluorescence signals were unexpectedly higher in knockout mutants than in the parental strain, and background fluorescence was detectable even in the absence of bilirubin. One possible explanation is the biochemical nature of bilirubin ditaurate (BDT), the substrate used in these assays. As the *gusA* gene encodes β-glucuronidase but not a sulfatase or taurase, it is unable to cleave the taurate moieties from BDT. This raises the possibility that BDT is not converted into unconjugated bilirubin within the bacterial cells, which may explain the lack of fluorescence activation in the *UnaG* system (54; 55).

It is further postulated that the peptide-like bond of BDT is resistant to hydrolysis by intestinal enzymes, including those expressed by *E. coli*. As a result, any microbial reduction of BDT could lead to the formation of a urobilinoid that remains conjugated to taurine. This would result in a more polar, taurin-conjugated metabolite that is unlikely to bind *UnaG* or

fluoresce, thus evading detection by the biosensor. While indirect evidence in the literature supports the formation of such polar urobilinoid conjugates from BDT, direct validation remains elusive (54; 55).

Given these uncertainties, it may be more appropriate to validate the *gusA* knockout strain using a conventional colourimetric β-glucuronidase assay, rather than relying solely on the *UnaG* biosensor. Furthermore, it appears that BDT-based validation may only be reliable in cell-free in vitro systems where enzymatic cleavage and fluorophore binding can be tightly controlled. These considerations underscore the need to refine substrate choice and validation methods when applying the biosensor in complex biological contexts.

In conclusion, while the IPTG–*UnaG* system performs reliably in controlled settings and demonstrates high sensitivity and specificity for bilirubin, its application in biologically complex samples presents challenges. Bilirubin degradation, antibiotic interactions, and background interference are key factors limiting consistent signal output. Nevertheless, the system provides a solid framework for bilirubin detection and may be further refined through improvements in fluorophore stability, uptake efficiency, and assay timing. Future work should focus on minimising degradation, validating uptake dynamics, and integrating rapid measurement protocols to enable more accurate in vivo and ex vivo applications.

Next steps of the experimental strategy (once the validation of the IPTG–*UnaG* system is complete) could involve attempting to reverse the urobilinogen to stercobilinogen pathway by converting stercobilinogen back to urobilinogen, and ultimately to bilirubin. If successful, the resultant metabolites could be analysed using the *UnaG*-System or liquid chromatography–mass spectrometry (LC–MS). The genomic profiles of the OMM21 strains could then be cross-referenced with their observed metabolic activities to identify potential candidates capable of catalysing this reverse conversion. Particular focus should be placed on strains harbouring *bilR* or homologous genes, to assess their involvement in the urobilinogen-to-stercobilinogen step.

List of Tables

List of Figures

A. Supplementary Information

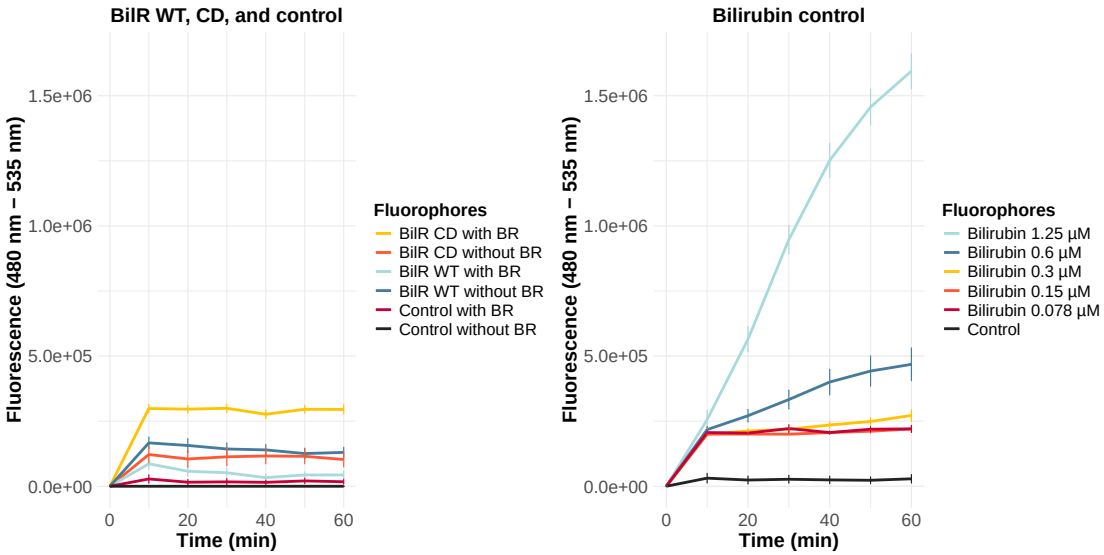

Figure A.1.: Replication of the validation experiment. Fluorescence intensities for BilR-WT and BilR-CD strains with or without bilirubin, including a control and a bilirubin standard curve.

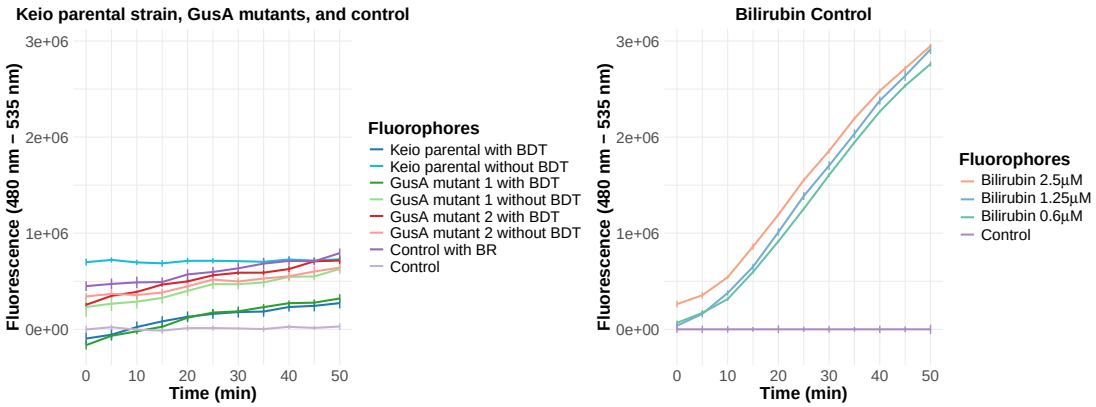

Figure A.2.: Replicate of the GusA–BDT validation experiment. Fluorescence data for Keio parental strain, GusA mutants, and controls.

Bibliography

[1] D. Zheng, T. Liwinski, and E. Elinav, "Interaction between microbiota and immunity in health and disease," *Cell Research*, vol. 30, pp. 492–506, June 2020. Publisher: Springer Science and Business Media LLC.

[2] S. Kim, C. Ndwandwe, H. Devotta, L. Kareem, L. Yao, and L. O'Mahony, "Role of the microbiome in regulation of the immune system," *Allergology International*, vol. 74, pp. 187–196, Apr. 2025. Publisher: Elsevier BV.

[3] J. L. Round and S. K. Mazmanian, "The gut microbiota shapes intestinal immune responses during health and disease," *Nature Reviews Immunology*, vol. 9, pp. 313–323, May 2009.

[4] B. Hall, S. Levy, K. Dufault-Thompson, G. Arp, A. Zhong, G. M. Ndjite, A. Weiss, D. Braccia, C. Jenkins, M. R. Grant, S. Abeysinghe, Y. Yang, M. D. Jermain, C. H. Wu, B. Ma, and X. Jiang, "BilR is a gut microbial enzyme that reduces bilirubin to urobilinogen," *Nature Microbiology*, vol. 9, pp. 173–184, Jan. 2024. Publisher: Nature Publishing Group.

[5] L. Vitek, J. Terry D Hinds, D. E. Stec, and C. Tiribelli, "The physiology of bilirubin: health and disease equilibrium," *Trends in molecular medicine*, vol. 29, p. 315, Feb. 2023.

[6] J. F. Creeden, D. M. Gordon, D. E. Stec, and T. D. Hinds, "Bilirubin as a metabolic hormone: the physiological relevance of low levels," *American Journal of Physiology - Endocrinology and Metabolism*, vol. 320, pp. E191–E207, Feb. 2021.

[7] A. L. Edwinson, L. Yang, S. Peters, N. Hanning, P. Jeraldo, P. Jagtap, J. B. Simpson, T.-Y. Yang, P. Kumar, S. Mehta, A. Nair, M. Breen-Lyles, L. Chikkamenahalli, R. P. Graham, B. De Winter, R. Patel, S. Dasari, P. Kashyap, T. Griffin, J. Chen, G. Farrugia, M. R. Redinbo, and M. Grover, "Gut microbial β-glucuronidases regulate host luminal proteases and are depleted in irritable bowel syndrome," *Nature Microbiology*, vol. 7, pp. 680–694, Apr. 2022.

[8] S. Zhou, Z. Wang, F. He, H. Qiu, Y. Wang, H. Wang, J. Zhou, J. Zhou, G. Cheng, W. Zhou, R. Xu, and M. Wang, "Association of serum bilirubin in newborns affected by jaundice with gut microbiota dysbiosis," *The Journal of Nutritional Biochemistry*, vol. 63, pp. 54–61, Jan. 2019.

[9] C.-H. Huang, W.-J. Jeng, Y.-P. Ho, W. Teng, W.-T. Chen, Y.-C. Chen, S.-M. Lin, C.-T. Chiu, I.-S. Sheen, and C.-Y. Lin, "Increased regulatory T cells in patients with liver cirrhosis correlated with hyperbilirubinemia and predict bacterial complications," *Journal of Gastroenterology and Hepatology*, vol. 30, no. 4, pp. 775–783, 2015. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jgh.12781.

[10] S. Zoroddu, B. Di Lorenzo, P. Paliogiannis, A. A. Mangoni, C. Carru, and A. Zinellu, "The association between bilirubin concentrations and inflammatory bowel disease: Insights from a systematic review and meta-analysis," *European Journal of Clinical Investigation*, vol. 54, no. 11, p. e14281, 2024. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/eci.14281.

[11] K. M. Schieffer, S. M. Bruffy, R. Rauscher, W. A. Koltun, G. S. Yochum, and C. J. Gallagher, "Reduced total serum bilirubin levels are associated with ulcerative colitis," *PLOS ONE*, vol. 12, p. e0179267, June 2017. Publisher: Public Library of Science (PLoS).

[12] Y. Liu, P. Li, J. Lu, W. Xiong, J. Oger, W. Tetzlaff, and M. Cynader, "Bilirubin possesses powerful immunomodulatory activity and suppresses experimental autoimmune encephalomyelitis," *Journal of Immunology (Baltimore, Md.: 1950)*, vol. 181, pp. 1887–1897, Aug. 2008.

[13] C. Eberl, D. Ring, P. C. Münch, M. Beutler, M. Basic, E. C. Slack, M. Schwarzer, D. Srutkova, A. Lange, J. S. Frick, A. Bleich, and B. Stecher, "Reproducible Colonization of Germ-Free Mice With the Oligo-Mouse-Microbiota in Different Animal Facilities," *Frontiers in Microbiology*, vol. 10, p. 2999, Jan. 2020.

[14] D. Garzetti, S. Brugiroux, B. Bunk, R. Pukall, K. D. McCoy, A. J. Macpherson, and B. Stecher, "High-Quality Whole-Genome Sequences of the Oligo-Mouse-Microbiota Bacterial Community," *Genome Announcements*, vol. 5, pp. e00758–17, Oct. 2017.

[15] A. S. Weiss, A. Burrichter, and B. Stecher, "Das Oligo-MM-Modell in der Darmmikrobiomforschung," *BIOspektrum*, vol. 29, pp. 18–21, Feb. 2023.

[16] J. Kwon, J.-S. Park, M. Kang, S. Choi, J. Park, G. T. Kim, C. Lee, S. Cha, H.-W. Rhee, and S.-H. Shim, "Bright ligand-activatable fluorescent protein for high-quality multicolor live-cell super-resolution microscopy," *Nature Communications*, vol. 11, p. 273, Jan. 2020. Publisher: Nature Publishing Group.

[17] A. Imam, S. Kumar Suman, P. K. Kanaujia, and A. Ray, "Biological machinery for polycyclic aromatic hydrocarbons degradation: A review," *Bioresource Technology*, vol. 343, p. 126121, Jan. 2022.

[18] E. J. Toone, *Advances in Enzymology and Related Areas of Molecular Biology, Volume 77*. John Wiley & Sons, Aug. 2011. Google-Books-ID: Pmh8yToeuloC.

[19] C. H. Gray and D. C. Nicholson, "627. The chemistry of the bile pigments. The structures of stercobilin and d-urobilin," *Journal of the Chemical Society (Resumed)*, p. 3085, 1958. Publisher: Royal Society of Chemistry (RSC).

[20] "stercobilinogen (CHEBI:36479)."

[21] S. Altschul, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Research*, vol. 25, pp. 3389–3402, Sept. 1997.

[22] S. F. Altschul, J. C. Wootton, E. M. Gertz, R. Agarwala, A. Morgulis, A. A. Schäffer, and Y. Yu, "Protein database searches using compositionally adjusted substitution matrices," *The FEBS Journal*, vol. 272, pp. 5101–5109, Oct. 2005.

[23] T. Seemann, "Prokka: rapid prokaryotic genome annotation," *Bioinformatics*, vol. 30, pp. 2068–2069, July 2014.

[24] A. Dalkiran, A. S. Rifaioglu, M. J. Martin, R. Cetin-Atalay, V. Atalay, and T. Doğan, "ECPred: a tool for the prediction of the enzymatic functions of protein sequences based on the EC nomenclature," *BMC Bioinformatics*, vol. 19, p. 334, Dec. 2018.

[25] C. P. Cantalapiedra, A. Hernández-Plaza, I. Letunic, P. Bork, and J. Huerta-Cepas, "eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale," *Molecular Biology and Evolution*, vol. 38, pp. 5825–5829, Dec. 2021.

[26] J. Huerta-Cepas, D. Szklarczyk, D. Heller, A. Hernández-Plaza, S. K. Forslund, H. Cook, D. R. Mende, I. Letunic, T. Rattei, L. Jensen, C. von Mering, and P. Bork, "eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses," *Nucleic Acids Research*, vol. 47, pp. D309–D314, Jan. 2019.

[27] F. Madeira, N. Madhusoodanan, J. Lee, A. Eusebi, A. Niewielska, A. R. N. Tivey, R. Lopez, and S. Butcher, "The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024," *Nucleic Acids Research*, vol. 52, pp. W521–W525, July 2024.

[28] N. P. Brown, C. Leroy, and C. Sander, "MView: a web-compatible database search or multiple alignment viewer.," *Bioinformatics*, vol. 14, pp. 380–381, Jan. 1998.

[29] A. Kumagai, R. Ando, H. Miyatake, P. Greimel, T. Kobayashi, Y. Hirabayashi, T. Shimogori, and A. Miyawaki, "A bilirubin-inducible fluorescent protein from eel muscle," *Cell*, vol. 153, pp. 1602–1611, June 2013.

[30] S. Dube, K. Nanda, R. Rani, N. J. Kaur, J. K. Nagpal, D. J. Upadhyay, I. A. Cliffe, K. S. Saini, and K. P. Purnapatre, "UDP-N-acetylglucosamine enolpyruvyl transferase from Pseudomonas aeruginosa," *World Journal of Microbiology and Biotechnology*, vol. 26, pp. 1623–1629, Sept. 2010. Company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 9 Publisher: Springer Netherlands.

[31] J. Molina-López, F. Sanschagrin, and R. C. Levesque, "A peptide inhibitor of MurA UDP-*N*-acetylglucosamine enolpyruvyl transferase: The first committed step in peptidoglycan biosynthesis," *Peptides*, vol. 27, pp. 3115–3121, Dec. 2006.

[32] T. Lambert, "UnaG at FPbase."

[33] S. Iwatani, H. Nakamura, D. Kurokawa, K. Yamana, K. Nishida, S. Fukushima, T. Koda, N. Nishimura, H. Nishio, K. Iijima, A. Miyawaki, and I. Morioka, "Fluorescent protein-based

detection of unconjugated bilirubin in newborn serum," *Scientific Reports*, vol. 6, p. 28489, June 2016. Publisher: Nature Publishing Group.

[34] S. Iwatani, K. Yamana, H. Nakamura, K. Nishida, T. Morisawa, M. Mizobuchi, K. Osawa, K. Iijima, and I. Morioka, "A Novel Method for Measuring Serum Unbound Bilirubin Levels Using Glucose Oxidase–Peroxidase and Bilirubin-Inducible Fluorescent Protein (UnaG): No Influence of Direct Bilirubin," *International Journal of Molecular Sciences*, vol. 21, p. 6778, Sept. 2020.

[35] "PSF-OXB1 - WEAK STRENGTH BACTERIAL PROMOTER PLASMID plasmid vector for molecular cloning."

[36] A. Oshima, J. W. Kyle, R. D. Miller, J. W. Hoffmann, P. P. Powell, J. H. Grubb, W. S. Sly, M. Tropak, K. S. Guise, and R. A. Gravel, "Cloning, sequencing, and expression of cDNA for human beta-glucuronidase.," *Proceedings of the National Academy of Sciences*, vol. 84, pp. 685–689, Feb. 1987. Publisher: Proceedings of the National Academy of Sciences.

[37] A. G. Sofronescu, T. Loebs, and Y. Zhu, "Effects of temperature and light on the stability of bilirubin in plasma samples," *Clinica Chimica Acta*, vol. 413, pp. 463–466, Feb. 2012.

[38] N. N. Rehak, S. A. Cecco, and G. L. Hortin, "Photolysis of bilirubin in serum specimens exposed to room lighting," *Clinica Chimica Acta*, vol. 387, pp. 181–183, Jan. 2008.

[39] S. Dickert, A. J. Pierik, D. Linder, and W. Buckel, "The involvement of coenzyme A esters in the dehydration of (*R*)-phenyllactate to (*E*)-cinnamate by *Clostridium sporogenes*," *European Journal of Biochemistry*, vol. 267, pp. 3874–3884, June 2000. Publisher: Wiley.

[40] S. C. Harris, S. Devendran, J. M. Alves, S. M. Mythen, P. B. Hylemon, and J. M. Ridlon, "Identification of a gene encoding a flavoprotein involved in bile acid metabolism by the human gut bacterium Clostridium scindens ATCC 35704," *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, vol. 1863, pp. 276–283, Mar. 2018. Publisher: Elsevier BV.

[41] T. Lu, Y.-Q. Wu, and H.-Y. Song, "The Nucleotide Sequence of gltD Gene Encoding the Small Subunit of Rhodobacter sphaeroides Glutamate Synthase," *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao Acta Biochimica Et Biophysica Sinica*, vol. 29, no. 3, pp. 294–302, 1997.

[42] F. R. Blattner, G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao, "The Complete Genome Sequence of *Escherichia coli* K-12," *Science*, vol. 277, pp. 1453–1462, Sept. 1997. Publisher: American Association for the Advancement of Science (AAAS).

[43] H. Bergler, G. Hogenauer, and F. Turnowsky, "Sequences of the envM gene and of two mutated alleles in Escherichia coli," *Journal of General Microbiology*, vol. 138, pp. 2093–2100, Oct. 1992. Publisher: Microbiology Society.

[44] P. M. Dombrosky, M. B. Schmid, and K. D. Young, "Sequence divergence of the murB and rrfB genes from Escherichia coli and Salmonella typhimurium," *Archives of Microbiology*, vol. 161, pp. 501–507, June 1994. Publisher: Springer Science and Business Media LLC.

[45] O. Ninnemann, C. Koch, and R. Kahmann, "The E.coli fis promoter is subject to stringent control and autoregulation.," *The EMBO Journal*, vol. 11, pp. 1075–1083, Mar. 1992. Publisher: Springer Science and Business Media LLC.

[46] F. Kunst, N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessières, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S.-K. Choi, J.-J. Codani, I. F. Connerton, N. J. Cummings, R. A. Daniel, F. Denizot, K. M. Devine, A. Düsterhöft, S. D. Ehrlich, P. T. Emmerson, K. D. Entian, J. Errington, C. Fabret, E. Ferrari, D. Foulger, C. Fritz, M. Fujita, Y. Fujita, S. Fuma, A. Galizzi, N. Galleron, S.-Y. Ghim, P. Glaser, A. Goffeau, E. J. Golightly, G. Grandi, G. Guiseppi, B. J. Guy, K. Haga, J. Haiech, C. R. Harwood, A. Hénaut, H. Hilbert, S. Holsappel, S. Hosono, M.-F. Hullo, M. Itaya, L. Jones, B. Joris, D. Karamata, Y. Kasahara, M. Klaerr-Blanchard, C. Klein, Y. Kobayashi, P. Koetter, G. Koningstein, S. Krogh, M. Kumano, K. Kurita, A. Lapidus, S. Lardinois, J. Lauber, V. Lazarevic, S.-M. Lee, A. Levine, H. Liu, S. Masuda, C. Mauël, C. Médigue, N. Medina, R. P. Mellado, M. Mizuno, D. Moestl, S. Nakai, M. Noback, D. Noone, M. O'Reilly, K. Ogawa, A. Ogiwara, B. Oudega, S.-H. Park, V. Parro, T. M. Pohl, D. Portetelle, S. Porwollik, A. M. Prescott, E. Presecan, P. Pujic, B. Purnelle, G. Rapoport, M. Rey, S. Reynolds, M. Rieger, C. Rivolta, E. Rocha, B. Roche, M. Rose, Y. Sadaie, T. Sato, E. Scanlan, S. Schleich, R. Schroeter, F. Scoffone, J. Sekiguchi, A. Sekowska, S. J. Seror, P. Serror, B.-S. Shin, B. Soldo, A. Sorokin, E. Tacconi, T. Takagi, H. Takahashi, K. Takemaru, M. Takeuchi, A. Tamakoshi, T. Tanaka, P. Terpstra, A. Tognoni, V. Tosato, S. Uchiyama, M. Vandenbol, F. Vannier, A. Vassarotti, A. Viari, R. Wambutt, E. Wedler, H. Wedler, T. Weitzenegger, P. Winters, A. Wipat, H. Yamamoto, K. Yamane, K. Yasumoto, K. Yata, K. Yoshida, H.-F. Yoshikawa, E. Zumstein, H. Yoshikawa, and A. Danchin, "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis," *Nature*, vol. 390, pp. 249–256, Nov. 1997. Publisher: Springer Science and Business Media LLC.

[47] N. Raich, P. Romeo, A. Dubart, A. Beaupain, M. Cohen-Solal, and M. Goossens, "Molecular cloning and complete primary sequence of human erythrocyte porphobilinogen deaminase," *Nucleic Acids Research*, vol. 14, no. 15, pp. 5955–5968, 1986. Publisher: Oxford University Press (OUP).

[48] A. F. McDonagh and D. A. Lightner, "Photooxidation of Bilirubin to Biliverdin and Bilirubin Structure," *Journal of Chemical Education*, vol. 85, p. 199, Feb. 2008.

[49] Y. Uchida, Y. Takahashi, Y. Morimoto, P. Greimel, A. Tosaki, A. Kumagai, T. Nishikubo, and A. Miyawaki, "Noninvasive monitoring of bilirubin photoisomer excretion during phototherapy," *Scientific Reports*, vol. 12, p. 11798, July 2022.

[50] G. Agati, F. Fusi, S. Pratesi, P. Galvan, and G. Donzelli, "Bilirubin photoisomerization

products in serum and urine from a Crigler—Najjar type I patient treated by phototherapy," *Journal of Photochemistry and Photobiology B: Biology*, vol. 47, pp. 181–189, Dec. 1998.

[51] J. F. Ennever, M. Sobel, A. F. Mcdonagh, and W. T. Speck, "Phototherapy for Neonatal Jaundice: in Vitro Comparison of Light Sources," *Pediatric Research*, vol. 18, pp. 667–670, July 1984.

[52] J. Jasprova, M. Dal Ben, E. Vianello, I. Goncharova, M. Urbanova, K. Vyroubalova, S. Gazzin, C. Tiribelli, M. Sticha, M. Cerna, and L. Vitek, "The Biological Effects of Bilirubin Photoisomers," *PLOS ONE*, vol. 11, p. e0148126, Feb. 2016.

[53] D. A. Lightner, W. P. Linnane, and C. E. Ahlfors, "Bilirubin Photooxidation Products in the Urine of Jaundiced Neonates Receiving Phototherapy," *Pediatric Research*, vol. 18, pp. 696–700, Aug. 1984.

[54] R. Koníčková, A. Jirásková, J. Zelenka, L. Lešetický, M. Štícha, and L. Vítek, "Reduction of bilirubin ditaurate by the intestinal bacterium Clostridium perfringens.," *Acta Biochimica Polonica*, vol. 59, Apr. 2012.

[55] L. Vitek, F. Majer, L. Muchova, J. Zelenka, A. Jiraskova, P. Branny, J. Malina, and K. Ubik, "Identification of bilirubin reduction products formed by Clostridium perfringens isolated from human neonatal fecal flora," *Journal of Chromatography B*, vol. 833, pp. 149–157, Apr. 2006.