Kaunas University of Technology

Faculty of Chemical Technology

Characterisation of inducible gene expression systems and their application in biosensors for gallic and *p*-coumaric acid quantitative evaluation

Master's Final Degree Project

Ingrida Kutraitė

Project author

Prof. Naglis Malys

Supervisor

Kaunas, 2021

Kaunas University of Technology
Faculty of Chemical Technology

Characterisation of inducible gene expression systems and their application in biosensors for gallic and *p*-coumaric acid quantitative evaluation

Master's Final Degree Project

Industrial biotechnology (6211FX010)

Ingrida Kutraitė
Project author

Prof. Naglis Malys
Supervisor

Researcher dr. Eglė Valančienė
Reviewer

Kaunas, 2021

Kaunas University of Technology

Faculty of Chemical Technology

Ingrida Kutraitė

Characterisation of inducible gene expression systems and their application in biosensors for gallic and *p*-coumaric acid quantitative evaluation

Declaration of Academic Integrity

I confirm that the final project of mine, on the topic „Characterisation of inducible gene expression systems and their application in biosensors for gallic and *p*-coumaric acid quantitative evaluation“ is written completely by myself; all the provided data and research results are correct and have been obtained honestly. None of the parts of this thesis have been plagiarised from any printed, Internet-based or otherwise recorded sources. All direct and indirect quotations from external resources are indicated in the list of references. No monetary funds (unless required by Law) have been paid to anyone for any contribution to this project.

I fully and completely understand that any discovery of any manifestations/case/facts of dishonesty inevitably results in me incurring a penalty according to the procedure(s) effective at Kaunas University of Technology.

Ingrida Kutraitė

(name and surname)

Kutraitė, Ingrida. Characterisation of Inducible Gene Expression Systems and Their Application in Biosensors for Gallic and *p*-Coumaric Acid Quantitative Evaluation. Master's Final Degree Project. Supervisor prof. Naglis Malys. Faculty of Chemical Technology, Kaunas University of Technology.

Study field and area (study field group): Biotechnology, Technological sciences.

Keywords: gallic acid, *p*-coumaric acid, inducible gene expression system, biosensor, transcription factor.

Kaunas, 2021. 83.

Summary

Phenolic acids are secondary plant metabolites with many significant properties, such as antioxidant, antimicrobial, or antiviral. Not only are phenolic acids beneficial to human health, but they find applications in a variety of industries, such as food, pharmaceutical, dye, and cosmetics, and their demand is growing annually. In biotechnology, phenolic acids are either extracted from plant materials or produced by microbial fermentation. The biosynthesis and production of phenolic acids are monitored and quantified using various analytical methods such as chromatography, UV spectroscopy, as well as electrochemical methods. Despite many different advantages of these methods, they do not provide sufficient throughput levels. Whereas such simple and efficient tools as genetically encoded biosensors can be used for *in vivo* real-time measurement of phenolic acids. Moreover, genetically encoded biosensors outperform other analytical methods when it comes to the high-throughput analysis and provide possibilities to perform measurements at the single-cell level, screen phenolic acid-producing strains, engineer metabolism and enzymes, select metabolic pathways of various substrate decomposition, or monitor bacterial strain evolution.

The object of this project is gallic and *p*-coumaric acid-inducible gene expression systems, which can be applied as genetically encoded biosensors. Notably, no gallic acid-specific genetically encoded biosensor has yet been developed. Whereas *p*-coumaric acid is known to be an important precursor of flavonoids, lignans, and stilbenes and is also being used for the production of protocatechuic acid. Therefore, the development of biosensors specific to gallic and *p*-coumaric acids will provide novel tools for *in vivo* monitoring and quantification of these industrially important phenolic acids.

Seven gallic acid- and four *p*-coumaric acid-inducible gene expression systems were predicted during the study based on the identified gene operons responsible for the catabolism of these phenolic acids. 19 and 14 inducible plasmid vectors specific for gallic and *p*-coumaric acids were constructed, respectively. The potential -35 and -10 promoter sequence motifs were identified in the intergenic region of the gallic acid-inducible gene expression system in *Pseudomonas putida* KT2440. However, no sufficiently conservative sequence region corresponding to the transcription factor GalR binding site was found. The *Pp*GalR/$P_{PP\_RS13150}$, *Pp*$P_{PP\_RS13150}$, *Pp*GalR/$P_{PP\_RS13170}$, *Pp*$P_{PP\_RS13170}$, *Pp*-A/$P_{PP\_RS13165}$, *Pp*GalR-C/$P_{PP\_RS13165}$ inducible systems were found suitable for the detection of gallic acid, and the *Bp*PadR/$P_{BPUM\_RS03685}$, *Bs*PadR-C/P_{BSU_34400}, *Bm*PadR/$P_{BMD\_RS02900, BMD\_RS01890}$ inducible systems were able to respond to the *p*-coumaric acid. The biosensor based on the *Pp*GalR/$P_{PP\_RS13150}$ system tested in *P. putida* KT2440 is specific to gallic acid in a concentration-dependent manner, and the *Bm*PadR-C/$P_{BMD\_RS02900,BMD\_RS01890}$ system-based biosensor tested in *E. coli* Top10 responds to *p*-coumaric acid in a concentration-dependent manner and can be used for the detection of ferulic acid.

Kutraitė, Ingrida. Indukuojamų genų ekspresijos sistemų charakterizavimas ir taikymas biosensoriuose galo ir *p*-kumaro rūgščių kiekybiniam įvertinimui. Magistro baigiamasis projektas. Vadovas prof. Naglis Malys. Kauno technologijos universitetas, cheminės technologijos fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Biotechnologijos, Technologijų mokslai.

Reikšminiai žodžiai: galo rūgštis, *p*-kumaro rūgštis, indukuojama genų ekspresijos sistema, biosensorius, transkripcijos faktorius.

Kaunas, 2021. 83 p.

Santrauka

Fenolinės rūgštys yra antriniai augalų metabolitai, pasižymintys daugeliu svarbių savybių, tokių kaip antioksidacinės, antimikrobinės ar antivirusinės. Jos yra ne tik naudingos žmogaus sveikatai, bet yra naudojamos įvairiose pramonės srityse, tokiose kaip maisto, farmacijos, dažų ir kosmetikos, ir jų poreikis kasmet vis auga. Biotechnologijoje fenolinės rūgštys yra išgaunamos ekstrakcijos būdu iš augalinių medžiagų arba gaminamos naudojant mikrobinę fermentaciją. Fenolinių rūgščių biosintezė ir gamyba yra stebima ir kiekybiškai įvertinama naudojant įvairius analitinius metodus, tokius kaip chromatografija, UV spektroskopija, taip pat elektrocheminius metodus. Nepaisant šių metodų privalumų, jie nesuteikia pakankamo našumo lygio. Tuo tarpu tokios paprastos ir veiksmingos priemonės kaip genetiškai užkoduoti biosensoriai gali būti naudojami fenolinių rūgščių matavimui *in vivo*. Be to, genetiškai užkoduoti biosensoriai pranoksta kitus analizės metodus kai norima atlikti analizę dideliu našumu ir suteikia galimybes atlikti matavimus vienos ląstelės lygmenyje, atrinkti fenolines rūgštis gaminančias padermes, vystyti metabolizmą ir fermentus, atrinkti fermentinius kelius vedančius prie įvairių substratų skaidymo ar stebėti bakterinių padermių evoliuciją.

Tyrimo objektas yra galo ir *p*-kumaro rūgštimis indukuojamos genų ekspresijos sistemos, kurios gali būti pritaikomos kaip genetiškai užkoduoti biosensoriai. Pažymėtina, kad dar nebuvo sukurtas galo rūgščiai specifiškas genetiškai užkoduotas biosensorius. *p*-Kumaro rūgštis yra žinoma kaip svarbus flavonoidų, lignanų ir stilbenų pirmtakas, ir yra naudojama protokatechuinės rūgšties gamybai. Todėl galo ir *p*-kumaro rūgštims specifiškų biosensorių sukūrimas suteiktų naujas priemones šių pramoniniu požiūriu svarbių fenolinių rūgščių stebėjimui ir kiekybiniam nustatymui *in vivo*.

Remiantis identifikuotais katabolizmo genų operonais, numatytos septynios galo rūgštimi ir keturios *p*-kumaro rūgštimi indukuojamos genų ekspresijos sistemos. Sukonstruota 19 ir 14 indukuojamų plazmidinių vektorių, specifiškų galo ir *p*-kumaro rūgštims, atitinkamai. *Pseudomonas putida* KT2440 mikroorganizme galo rūgšties indukuojamos sistemos tarpgeniniame regione identifikuotos potencialios -35 ir -10 promotorinės sritys. Tačiau nerasta pakankamai konservatyvios sekos srities, atitinkančios transkripcijos faktoriaus GalR prisijungimo vietą. *Pp*GalR/P$_{PP_RS13150}$, *Pp*P$_{PP_RS13150}$, *Pp*GalR/P$_{PP_RS13170}$, *Pp*P$_{PP_RS13170}$, *Pp*-A/P$_{PP_RS13165}$, *Pp*GalR-C/P$_{PP_RS13165}$ sistemų pagrindu sukurti biosensoriai buvo identifikuoti kaip tinkami galo rūgšties nustatymui, o indukuojamų sistemų *Bp*PadR/P$_{BPUM_RS03685}$, *Bs*PadR-C/P$_{BSU_34400}$, *Bm*PadR/P$_{BMD_RS02900,BMD_RS01890}$ pagrindu sukurti biosensoriai nustatyti kaip tinkami *p*-kumaro rūgšties nustatymui. *Pp*GalR/P$_{PP_RS13150}$ sistemos pagrindu sukurtas biosensorius, testuotas *P. putida* KT2440, yra specifiškas galo rūgščiai priklausomai nuo koncentracijos, o *Bm*PadR-C/P$_{BMD_RS02900,BMD_RS01890}$ sistemos pagrindu sukurtas biosensorius, testuotas *E. coli* Top10, reaguoja į *p*-kumaro rūgštį priklausomai nuo koncentracijos ir gali būti naudojamas ferulo rūgšties nustatymui

Table of contents

List of abbreviations and terms **8**
Introduction **9**
1. Literature review **10**
1.1. Phenolic acids 10
1.1.1. Functions, sources, and applications of phenolic acids 10
1.1.2. Structure and synthesis of phenolic acids 10
1.2. Use of inducible gene expression systems in biosensors 12
1.2.1. Inducible gene expression systems 13
1.2.2. Transcription factor-based biosensors 13
1.2.3. Applications of transcription factor-based biosensors 15
1.3. Transcription factor-based biosensors for detection of phenolic acids 16
1.3.1. Transcription factor-based biosensors of *p*-coumaric acid 17
1.3.2. Transcription factor-based biosensors of caffeic acid 18
1.3.3. Transcription factor-based biosensors of salicylic acid 19
1.3.4. Transcription factor-based biosensors of vanillic acid 20
1.3.5. Transcription factor-based biosensors of protocatechuic acid 20
1.3.6. Transcription factor-based biosensors of ferulic acid 21
1.4. Justification of the project aim and objectives 21
2. Materials and methods **22**
2.1. Equipment, tools, and materials 22
2.2. Strains and plasmids 23
2.3. Methods 23
2.3.1. Bioinformatics 23
2.3.2. Cell cultivation 24
2.3.3. Microbial genomic DNA purification 24
2.3.4. Plasmid DNA purification 24
2.3.5. DNA amplification 25
2.3.6. Agarose gel electrophoresis 27
2.3.7. Extraction of DNA fragments from agarose gel 27
2.3.8. Plasmid construction 28
2.3.8.1. Restriction digestion 28
2.3.8.2. Ligation 28
2.3.8.3. HiFi DNA assembly 29
2.3.9. Preparation of competent cells 29
2.3.10. Transformation and screening 30
2.3.10.1. Transformation 30
2.3.10.2. Colony PCR 31
2.3.10.3. Secondary restriction digestion 31
2.3.11. Replacement of antibiotic resistance gene 32
2.3.12. Preservation of suitable biosensors for long-term storage 32
2.3.12.1. Preservation of biosensors to MicroBank 32
2.3.12.2. Preservation of biosensors with glycerol 32
2.3.13. Screening of inducible systems 33

2.3.14. Mathematical modelling 33
3. Results and discussion 35
3.1. Identification of gallic acid-inducible system 35
3.1.1. Sequence alignment of *galR*, *galB*, and *galC* homologues 36
3.1.2. Sequence alignment of intergenic regions 41
3.1.3. Construction of gallic acid-inducible plasmid vectors 43
3.1.4. Selection of the solvent for investigation of gallic acid-inducible biosensors 46
3.1.5. Screening of gallic acid-inducible gene expression systems 48
3.1.6. The specificity of gallic acid-inducible biosensor 49
3.1.7. The sensitivity of gallic acid-inducible biosensor 50
3.2. Identification of *p*-coumaric acid-inducible system 52
3.2.1. Construcion of *p*-coumaric acid-inducible plasmid vectors 52
3.2.2. Screening of *p*-coumaric acid-inducible gene expression systems 54
3.2.3. The specificity of *p*-coumaric acid-inducible biosensor 55
3.2.4. The sensitivity of *p*-coumaric acid-inducible biosensor 57
4. Recommendations part 58
Conclusions 61
List of references 62
Appendices 69
Appendix 1. Schematic view of constructs assembly using ligation reaction 69
Appendix 2. Schematic view of constructs assembly using HiFi assembly method 70
Appendix 3. Schematic view of antibiotic resistance gene replacement 71
Appendix 5. *glsB*/*galT* intergenic regions 75
Appendix 6. Gallic acid-inducible plasmids designed and constructed in this study 77
Appendix 7. *p*-Coumaric acid-inducible plasmids designed and constructed in this study 81

List of abbreviations and terms

ADB – agarose dissolving buffer;
BLAST – basic local alignment search tool;
DNA – deoxyribonucleic acid;
FACS – fluorescence-activated cell sorting;
GFP – green fluorescent protein;
GlsB – glutaminase B;
IGR – intergenic region;
YFP – yellow fluorescent protein;
mRNA – messenger RNA;
NCBI – national center for biotechnology information;
pDNA – plasmid DNA;
PCR – polymerase chain reaction;
RBS – ribosome binding site;
Rcf – rotation centrifugal force;
Rease – restriction endonuclease;
RFP – red fluorescent protein;
RNA – ribonucleic acid;
RNAP – RNA polymerase;
Rpm – revolutions per minute;
TF – transcription factor.

Terms:

Plasmid – a double-stranded, circular, extrachromosomal DNA molecule inside the cell, which is physically separated from chromosomal DNA and can replicate independently.

Vector – DNA molecule carrying a target genetic material into another cell. In this project, plasmids were assembled and transferred into the host microorganisms, therefore terms plasmid, vector, and construct are used alternatively.

Introduction

Phenolic acids are widely found in nature, principally in plants and fungi. They belong to phenolics, constituting one of the largest groups of these chemical compounds [1], [2]. Phenolic acids are characterised as great antioxidants, antimicrobial substances, modifying agents of cell signalling pathways, and possess a high therapeutic potential [1]. They are known as essential building blocks and are used in various industries, including food, medical, agricultural, and cosmetics [3], providing great interest for extraction and production, and the use of these compounds increases annually [1]. The importance of the detection of phenolic acids to improve production is, therefore, considerable.

There is a diversity of analytical methods for monitoring of production and detection of specific molecules, such as chromatography, UV spectroscopy, electrospray ionization combined with mass spectrometry [4], H^1- and C^{13}-based NMR techniques [5], or electrochemical methods [6]. Although providing great effectiveness and accuracy, they are limited by low throughput, restricting the application for investigating millions of molecules or in demand to screen at the single-cell level [5]. Developing simple and rapidly operating biotests such as biosensors could increase the efficiency of the production of compounds of interest. Biosensors function by simply combining two elements, of which many were combined to generate small modules with specified behaviours, including switches [7], pulse generators [8], and oscillators [9].

Currently, a lot of attempts in the field of synthetic biology are put in biosensors, adapted in a novel way. The merging of biological systems with engineering provides an ability to manipulate the operating of microorganisms on a genetic level [10]. The use of inducible gene expression systems might contribute to it, as they manage to provide the arrangement of genes in the desired way, and the detectable signal will be received comparatively simply. Such systems are an essential component of metabolic engineering, enabling strict control of gene expression and leading to various applications, such as detection of molecules of interest [11], identification of producing strains [12], investigation of biomaterials [13], or even use in gene therapy [14]. However, the number of compounds that can be detected by genetically encoded biosensors is still limited [15]. To increase the diversity, molecules of interest are investigated and applied in such a field [16], [17]. In this final project, gallic and *p*-coumaric acid-inducible gene expression systems were selected to be identified, characterised, and applied to develop genetically encoded biosensors. As an extensive characterisation of genetically encoded biosensors is usually lacking [11], the TF-dependent orthogonality, specificity, and sensitivity of the selected biosensors were also set to be estimated.

The aim of the project was the identification and characterisation of gallic and *p*-coumaric acid-inducible gene expression systems and their application for biosensor development.

Objectives:

1. to identify gene operons responsible for gallic and *p*-coumaric acid catabolism and predict relevant inducible gene expression systems that control the expression of these operons;
2. to use bioinformatics tools and identify genetic elements, such as promoters and the binding site of transcription factors in gallic acid-inducible systems;
3. to generate plasmid constructs with gallic and *p*-coumaric acid-inducible gene expression systems;
4. to characterise gallic and *p*-coumaric acid-inducible systems;
5. to develop biosensors for detection and quantitative analysis of gallic and *p*-coumaric acid.

1. Literature review

1.1. Phenolic acids

Phenolic acids belong to the large class of phenolic compounds, known as secondary metabolites, which are synthesised by plants and fungi. In general, phenolic acids contain a carboxyl group and one or a few hydroxyl groups as essential features [1], determining the bioactive properties of these compounds [2]. The beneficial characteristics and functions of phenolic acids were investigated in agricultural, biological, chemical, and medical research studies as summarised in [16], as they possess many beneficial properties [17]. These compounds are highly essential for human health and plants [2], and it is, therefore, relevant to have a precise understanding of phenolic acids' concentrations and biological activities in cells.

1.1.1. Functions, sources, and applications of phenolic acids

Phenolic acids mold one of the largest and most essential groups of secondary metabolites and bioactive compounds in plants [2]. They are of high importance, exhibiting antioxidative, anti-inflammatory, anti-aging, and possible anticarcinogenic characteristics. The damage of oxidative stress conditions in the development of chronic and age-related diseases, though antioxidants delay and reduce the risk of them. Phenolic acids act not only as reducing agents but also as free radical scavengers and suppressors of singlet oxygen formation [16]. They ensure antimicrobial activity against some strains of bacteria such as *Staphylococcus aureus* [18] and are proven to be more efficient than Vitamin C, E, and carotenoids. The ability to reduce the damage of oxidative stress depends on the number of free hydroxyl groups in the structure of phenolic acid [19]. Phenolic acids are also associated with great potential for food preservation [17], wherefore they are widely utilised in food and agricultural industries [2]. Overall, these important properties-possessing compounds are significant not only for human health but also for the plant itself.

Despite that phenolic acids do not directly affect the growth of the plant, they are still essential for its survival, impacting physiology, ecology, and development. They are significant for reproduction, conversion of nutrients, protein synthesis, enzymatic activity, and photosynthesis. Moreover, they act as structural components, signaling agents, and pigments in plants [20]. Many phenolic acids reside in plant-derived foods, including fruits, vegetables, grains, spices, tea, coffee, oilseeds, and even alcoholic beverages, such as beer and red wine [1]. The amount and content of phenolic acids depend on growing conditions and the stage of plant maturation. However, only a few exist in the free acid form, while others are linked through ester, ether, or acetal bonds to cellulose, proteins, lignin, flavonoids, glucose, or terpenes [2]. Consequently, such arrangement complicates the analysis and detection of phenolic acids [4].

1.1.2. Structure and synthesis of phenolic acids

Phenolic acids are usually divided into two main classes – hydroxycinnamic and hydroxybenzoic acids [16]. They contain at least one aromatic ring, with at least one hydrogen supplemented by a hydroxyl group [21]. The principal hydroxybenzoic derivatives are *p*-hydroxybenzoic, protocatechuic, vanillic, syringic, gentisic, and gallic acids. Hydroxycinnamic acid derivatives are the most prevalent phenolic acids in plant tissues, such as *p*-coumaric, caffeic, ferulic, and sinapic acids [1]. Table 1.1 represents all common phenolic acids of both classes.

Table 1.1. Common phenolic acids [4], [21]

Hydroxybenzoic acids	Hydroxycinnamic acids
p-Hydroxybenzoic acid (*p*-salicylic acid)	*p*-Coumaric acid
o-Hydroxybenzoic acid (salicylic acid)	*o*-Coumaric acid
m-Hydroxybenzoic (*m*-salicylic acid)	*m*-Coumaric acid
Protocatechuic acid (3,4-Dihydroxybenzoic acid)	Ferulic acid
Vanillic	Sinapic acid
Gallic acid	Caffeic acid
5-Hydroxy-vanillic acid	Hydroxyferulic acid
Gentisic acid (2,5-dihydroxybenzoic acid)	
Syringic acid	

Phenolic acids are mainly produced by the extraction from plants, and production yield depends on the plant tissue used. They can be extracted from seeds, leaves, nuts, vegetables, or berries and fruits, the latter being the most suitable choice, as such source possesses mainly free forms of phenolic acids. For instance, gallic acid was extracted from Black tea (*Clonorchis sinensis*) leaves up to 6550 μg/g dry weight [22], whereas extraction of this phenolic acid yielded up to 1485–30603 μg/g dry weight from *Byrsonima ligustrifolia* fruits [23]. However, industrial production by extraction is mainly limited due to the bound forms and inhibitory effects of extraneously synthesised compounds [2]. Notably, the number of hydroxyl groups in a phenolic acid structure, also its concentration and polarity, can restrict the extraction effectiveness [24]. Besides, phenolic acids can also be produced by the biosynthesis from organic wastes by microbial fermentation [2], resembling the reactions occurring in plants, which provides the potential to synthesise these compounds more readily.

Phenolic acids emerge via the shikimate pathway in microorganisms and higher plants, where shikimic acid is the central metabolite (Figure 1.1). This pathway links primary and secondary metabolism, beginning from the conversion of simple carbohydrate molecules, such as glucose, into aromatic amino acids [24]. In plants, the shikimate pathway consists of seven reactions that occur in the plastids (chloroplasts). It starts with a condensation of phosphoenolpyruvic acid and D-erythrose-4-phosphate, producing 3-deoxy-D-arabino-heptulosonic acid 7-phosphate. A final and essential compound of the shikimate pathway is chorismic acid [20], which is generally converted into L-phenylalanine or L-tyrosine. Phenylalanine ammonia-lyase then converts L-phenylalanine into hydroxycinnamic acids and releases ammonia, although some plants use another enzyme - tyrosine ammonia-lyase, converting L-tyrosine into *p*-coumaric acid [24]. After deamination, hydroxycinnamic acid is converted into benzoic acid under the action of oxidase, while synthesised *p*-coumaric, salicylic, and *p*-hydroxybenzoic acids are hydroxylated and methylated during subsequent reactions, resulting in the synthesis of other hydroxycinnamic or hydroxybenzoic acids [2], [21]. Overall, the synthesis of phenolic acids in plants is specific and similar to the reactions of microbial pathways, contributing to the perception of regulatory mechanisms of metabolic pathways proceeding in microorganisms.

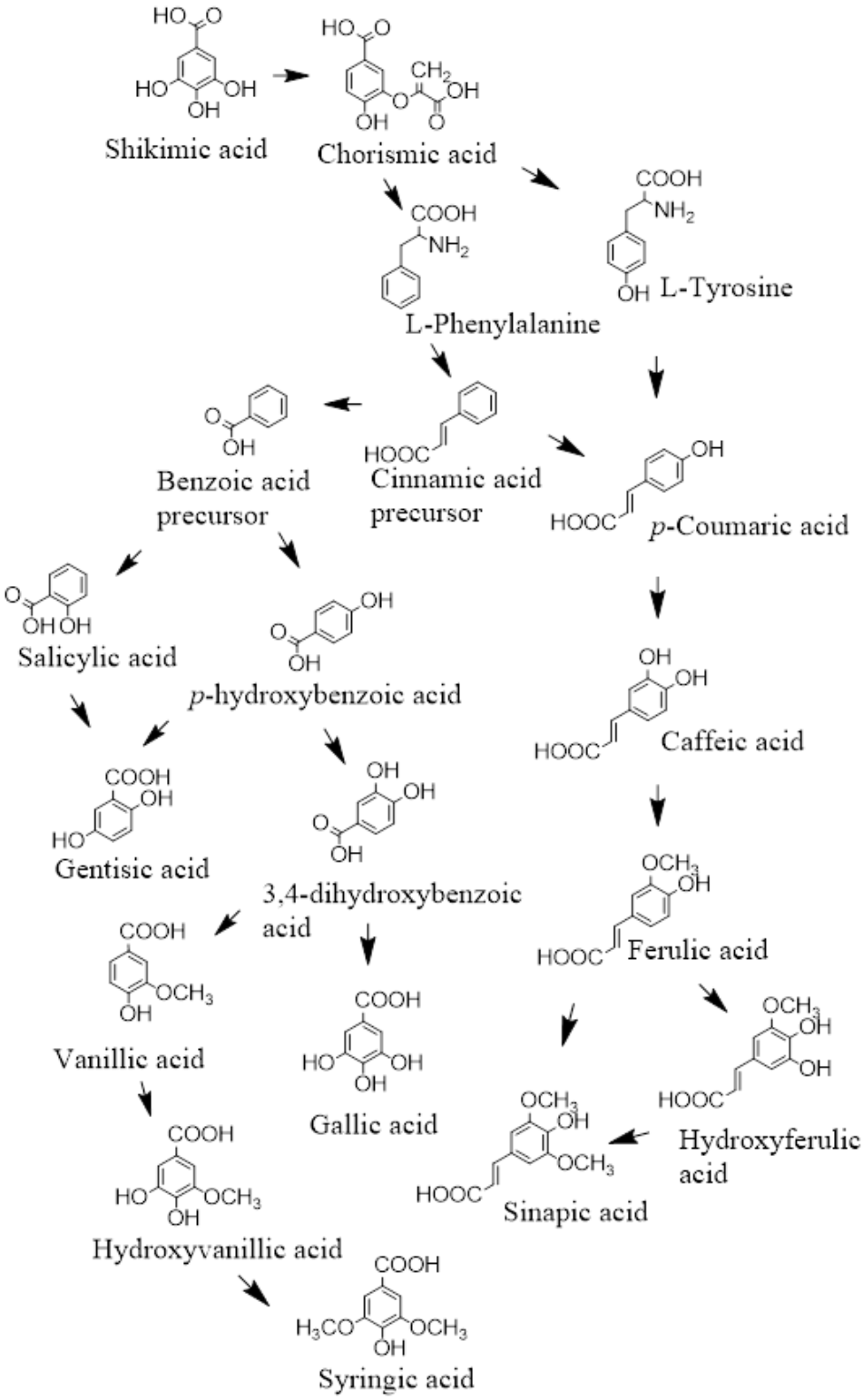


Fig. 1.1. Schematic representation of phenolic acids synthesis in plants

1.2. Use of inducible gene expression systems in biosensors

Living organisms have adapted to a constantly unstable and fluctuating environment by developing sophisticated sensing and regulatory mechanisms [25], which occur at the level of transcription, translation, or post-translational modification. Transcriptional regulation adjusts the consumption of resources and energy, suppresses the formation of growth-inhibiting reaction intermediates, and preserves the presumable virulence of the organism [26]. Plenty of distinct regulatory mechanisms

modulating gene expression were already discovered, which are used in conjunction to control the expression of a single gene or operon [10]. As a result, the investigation of transcriptional regulation by employing inducible gene expression systems can improve the perception of metabolic engineering and increase the understanding of reactions occurring at the single-cell level.

1.2.1. Inducible gene expression systems

Although inducible regulation proceeds only in specific genes, as a reaction to elevated temperature, pH fluctuations, light, metabolites, or toxic molecules, however, it is known that regulators, known as transcription factors (TFs), can sense various inducers [14]. TF, which is also called a transcriptional regulator or sequence-specific DNA (deoxyribonucleic acid)-binding factor, is a protein that monitors the rate of transcription or the DNA conversion to messenger RNA. The essential feature of TFs is that they have at least one DNA-binding domain, which connects to the DNA strand, usually next to the genes that they regulate. Equally, they contain a ligand-binding domain, which can connect with an inducing molecule. TFs play a relevant role in gene expression, turning on and off genes to maintain required gene expression in the right cell at the proper time and in a suitable amount [26]. Generally, an inducible gene expression system consists of a regulatory gene, promoter, operator, and structural (regulated) genes. During transcription, RNA polymerase (RNAP) binds to the promoter and begins the transcription of specific genes. Regulatory genes are responsible for the synthesis of TFs, which can either form a complex with inducing molecules and bind to the DNA operator site or cause the release of the TF from the DNA strand, therefore activating or repressing the gene expression [2].

TFs are commonly described with compatible promoters, which are called TF/promoter pairs. However, many bacterial strains have difficulty in expanding the amount of these pairs [27]. Generally, TFs are divided into three main groups – activators that can stimulate the transcription, repressors that manage to inhibit the transcription, and dual-function TFs that operate either as activators or repressors [25]. These gene expression-modifying factors are divided into more than 50 diverse families, including MarR, PadR, DesR, GntR, depending on the genes that they regulate [28]. TFs which can recognize aromatic compounds are divided into three families – XylS-AraC, NahR-LysR, and XylR-NtrC [26], of which the LysR family of TFs compiles the most widespread family found in bacteria [29]. Microorganisms, essentially prokaryotes possess plenty of TFs, currently are known more than 400 unique ligand-TF pairs [25], of which 230 are found in *Escherichia coli* [30]. Accordingly, inducible gene expression systems are mainly expressed in *E. coli* and *Bacillus subtilis* – two mostly employed microorganisms in synthetic biology [31]. However, transferring inducible systems from one host to another can lead to various drawbacks, such as unpredicted behaviour, usually proceeding due to distinct cellular machinery [14]. For instance, it was observed that inducible system tools that perfectly functioned in *E. coli* do not behave as expected in other bacteria [31]. Overall, it is still complicated to apply TFs for the development of a robust inducible system applicable in diverse hosts.

1.2.2. Transcription factor-based biosensors

In molecular biology, TF-based biosensors are also known as genetically encoded or whole-cell biosensors. Such biosensors are simply the inducible gene expression systems transferred into a host microorganism (Figure 1.2). An activator-type biosensor is initiated when an inducing molecule binds to a specific TF, creating a complex, which can merge with a region in a DNA strand called operator

(O). Such binding induces RNAP holoenzyme to connect with a promoter (P) and the transcription to begin, leading to the expression of the regulated gene, which is replaced with a reporter-encoding gene [32]. A reporter gene, or simply a reporter, encodes one of the visualisable proteins, such as an enzyme or a fluorescent protein. Reporters are classified into constitutively expressed (always-on) and inducible reporters, the latter being applied in biosensors since they are sensitive to inducing molecules and the action of TFs [26]. Inducible expression of the fluorescent protein is helpful to monitor the efficiency of induction, and it is significant to ensure that the inducible gene expression system does not operate constitutively [33]. Fluorescent proteins are generally used as reporters because they are stable, direct, and convenient tools [34]. Hence, when the reporter gene is induced, transcription is initiated, demonstrating the functioning of the developed biosensor and the presence of the inducing molecule [9]. Overall, TF-based biosensors demonstrate the reaction occurred in such a way that its response can be readily visualised.

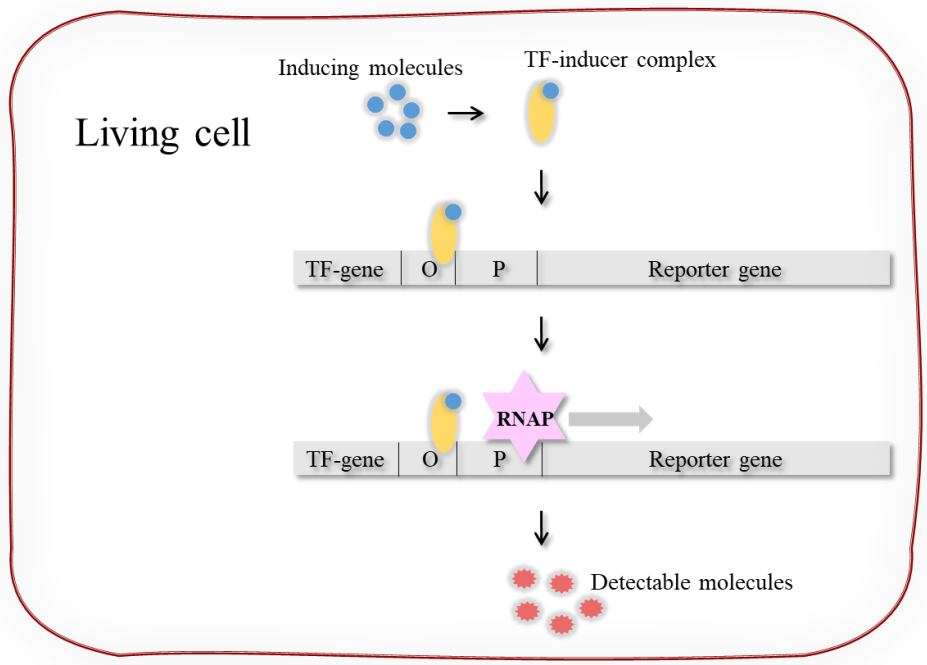


Fig. 1.2. Mechanism of an activator-type TF-based biosensor, O – operator, P – promoter, RNAP – RNA polymerase, TF – transcription factor

On the contrary to conventional biosensors, TF-based biosensors are known as more sensitive and universal, as they can detect a considerable number of various molecules. They are also capable of genetic modifications, and their operating conditions' range is broader [32]. The magnitude of TF-based biosensors can be constructed, depending on the location of the TF-gene in the DNA strand, thus developing a specific type of biosensors. An activator-type biosensor is usually developed by inserting a TF-gene and the promoter site of the regulated gene upstream of the reporter gene. A repressor-type biosensor is developed by inserting only the promoter site of the structural genes upstream of the reporter gene [33]. Although the rules of biosensor design and engineering are presently lacking [35], the development of differently arranged inducible systems, necessarily containing the inducible promoter for its corresponding effector, enables the screening of various inducible systems [10]. Overall, the large modification ability of genetically encoded biosensors provides a lot of potentials to optimise their sensitivity and use them in a wider field of applications.

1.2.3. Applications of transcription factor-based biosensors

Usually, investigation mechanisms for the detection of specific molecules are too expensive, time-consuming, requiring special techniques and experience in such a field [36]. However, the techniques of genetic engineering can overcome such disadvantages. Since mechanisms of bacterial gene regulation are too complex and multi-layered, they easily cannot be used for genetic engineering. Although the perception of appealing properties of specific and well-characterised molecules, which can be captured, engineered, and optimised for application in synthetic biology, contributes to the application of TF-based biosensors [10]. As TFs activate or repress gene expression, biosensors are used to observe this genetically encoded process. Therefore changes in intra- and extracellular fluid can be observed, as the use of biosensors enables cells to report their success by chemical production in real-time [11], [37] Overall, biosensors are successful because of their simplicity and only minimal demand for intricate equipment and skills.

Since naturally occurring microorganisms are usually inadequate for large-scale applications [38], genetically encoded biosensors, which are based on engineered microorganisms, can be applied industrially and synthetically [10], for instance, applying them in so-called high-throughput screening. This way of application was demonstrated in the dynamic regulation of the synthesis of various valuable compounds, such as fatty acids and muconic acid [39]. Such screening highly contributes to the detection of natural products, which are used in pharmaceutical, food, agricultural, environmental, and other industries by, for example, identifying producer strains [13], [34]. For instance, 186 strain variants from a 10-million-mutant library of *L*-lysine-producing *E. coli* variants were screened in the action of TF-based biosensors. Moreover, two best-performing variants were selected and applied for more effective lysine production [40]. Besides, whole-cell biosensors in high-throughput screening are often combined with fluorescence-activated cell sorting (FACS) [41], [42], [43]. Such combination demonstrated the ability to indicate yeast cells producing *p*-coumaric acid [33]. Biosensors combined with FACS also helped to improve the activities of the specific enzymes in significant biosynthetic pathways [44] and enabled the screening of suitable enzymes in the metagenomic library [45]. Altogether, biosensors applied in high-throughput screening enable the quicker selection of producer strains, help to improve the production yields or activities of enzymes.

Furthermore, genetically encoded biosensors are used to detect the release of specific low molecular weight chemical compounds. One study demonstrated that TF-based biosensors enabled the detection of phenolic acids from such recalcitrant polymers as lignin and contributed to achieving a sustainable and economically attractive valorisation process [46]. The release of aromatic monomers from lignocellulose can be detected, and this way applied in a real-time monitoring, high-throughput screening, enabling dynamic pathway control of enzymatic pathways, or benefit in selection of high-producing strains through analysis of adaptive evolution (Figure 1.3). Such a detection mechanism can improve the enzymatic approach towards lignin valorisation and, therefore, aid in a sustainable biorefinery process [43]. Overall, this application of TF-based biosensors could aid to enhance the sustainable processes, from biomass utilisation to production of aromatic compounds.

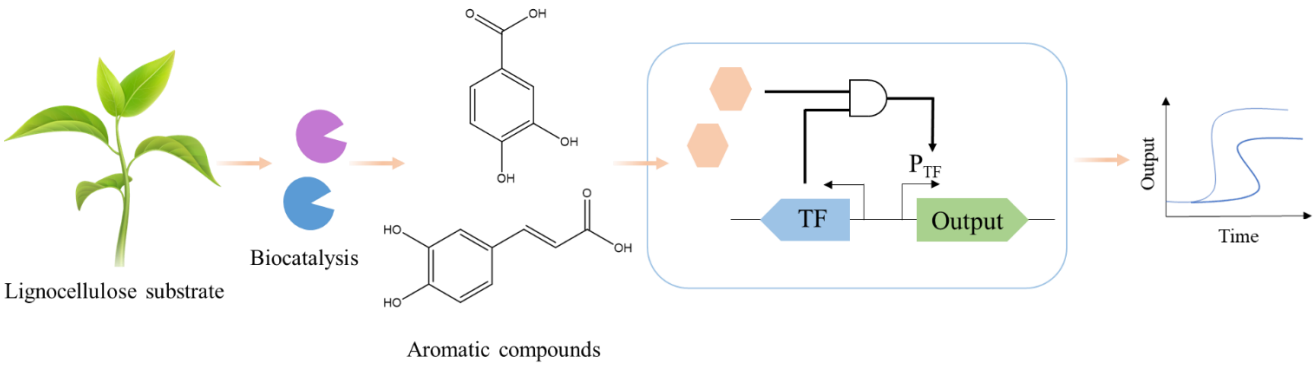


Fig. 1.3. Application of genetically encoded biosensors for lignin valorisation, figure was modified by [43]

Moreover, whole-cell biosensors are not only of scientific and industrial interest but also lead to an application in human and animal disease prevention and health care, including diagnostics, therapeutics, gene therapy [10], and understanding of the pharmacological activity of the specific drugs [32]. The biomedical field benefits from whole-cell biosensors, as they are employed for the detection of chronic pathologies, for instance, the changes in mitochondrial Ca^{2+} concentrations were detected by the genetically encoded biosensor, indicating the impaired cellular energy metabolism, apoptosis, and autophagy, even surpassing the chemical Ca^{2+} indicators [47]. Interestingly, whole-cell biosensors were indicated as suitable tools for detection of tetracycline in the intestines, demonstrating the ability to detect this antibiotic before the accumulation of tetracycline-resistance genes in bacteria colonizing the gastrointestinal tract. Such detection mechanism indicated a potential utility of the biosensor, which could optimise the therapy with tetracycline without harming the patient's organism [48]. In conclusion, demonstrated applications of whole-cell biosensors provide the expectance to develop more specific and highly applicable biosensors.

1.3. Transcription factor-based biosensors for detection of phenolic acids

In recent years, biosensors have been developed as tools to detect and observe gene expression of inducing molecules, basically due to the fast response, simplicity, and effectiveness [36]. However, as biosensors are usually constructed for intracellular assessment at the single-cell level, their usage as a screening method is limited [33]. Though many biosensors that integrate environmental signals and monitor the gene expression have been constructed, only several were identified for phenolic acids, such as *p*-coumaric, salicylic, vanillic, protocatechuic, caffeic, and ferulic acids (Table 1.2). The development of specific biosensors emerges some challenges, such as suboptimal output strength, narrow dynamic range, and poor inducer specificity [39], demonstrating the need to expand the field of TF-based biosensors for detection of various molecules, including phenolic acids.

Table 1.2. Existing biosensors used for the detection of phenolic acids

Phenolic acid	TF*	Locus Tag	Origin of TF	Host of application	Reporter	Reference
***p*-Coumaric acid**	PadR	BSU_08340	*B. subtilis*	*E. coli*, *C. glutamicum*	YFP	[33]
Caffeic acid	FerC	AAP78949	*Sphingobium* sp. SYK-6	*E. coli*	GFP	[46]

Salicylic acid	NahR	HK44_029505	*P. fluorescens* HK 44	*P. fluorescens* HK 44	Bioluminescence	[49]
***p*-Salicylic acid**	HbaR	TX73_RS03475	*Rhodopseudomonas palustris*	*E. coli*	LacZ	[50]
Vanillic acid	VanR	CCNA_02475	*Caulobacter crescentus CB15N*	*E. coli*	RFP	[51]
Protocatechuic acid	PcaU	ACIAD1702	*Acinetobacter sp. ADP1*	*E. coli*	GFP	[52]
Ferulic acid	FerC	AAP78949	*Sphingobium* sp. SYK-6	*E. coli*	GFP	[46]

*TF (transcription factor) is a specific DNA-binding protein, which modifies RNAP allowing the transcription to begin.

1.3.1. Transcription factor-based biosensors of *p*-coumaric acid

p-Coumaric acid, or 4-hydroxycinnamic acid, is the most common isomer of coumaric acids in nature. It is known as a significant precursor for the synthesis of many plant natural products, such as flavonoids, lignans, polyphenols, stilbenes, and coumarins [39], [53]. It is also being used as a substrate to produce protocatechuic acid [21], and it plays an important role in the synthesis of various phenolic acids in plants [24]. Therefore, *p*-coumaric acid demonstrates the necessity to develop genetically encoded whole-cell biosensors.

A lot of potential to develop a functional *p*-coumaric acid-inducible biosensor is observed by employing the *Bacillus* genus bacteria. There is still no indicated TF, which would react to *p*-coumaric acid in the genome of *E. coli* [41], but the repressor-encoding gene *padR* and its promoter exist in *Bacillus subtilis*. This gene encodes the PadR, which is responsive to *p*-coumaric acid, inhibiting the expression of the *padC*, encoding the phenolic acid decarboxylase. PadC molds an operon with a genetic sequence called *yveFG*, which has long been considered as a gene of unknown function [33], but recently was shown to be responsible for the activation of *padC* transcription as well [39]. However, *padR* expression in *B. subtilis* is low and seems to be constitutive [12]. However, it was indicated that *padR* expression in *E. coli* is too high, resulting in acute inhibition of its downstream promoter [54]. It is known that excessive gene expression leads to constitutive repression, while inadequate expression does not cause repression, even if the inducer is not present. The expression of repressors should be therefore highly regulated [55]. Supposedly, *B. subtilis* contains an additional regulatory element, which helps to maintain low expression of the *padR* [33].

Accordingly, Siedler et al. [33] developed a biosensor, which consisted of the *padR*, along with *yveFG/padC* gene regions and yellow fluorescent protein (YFP) encoding gene inserted instead of the native *padC* gene (Figure 1.4). In normal conditions, PadR dimerizes and binds to the PadR binding site located in the promoter site of *padC*, repressing the transcription of *yveFG* and *YFP*. In the presence of *p*-Coumaric acid, it binds to PadR and releases it from the operator site, which leads to the de-repression of *yveFG* and *YFP*. Subsequently, the transcription of *YFP* is not repressed, and there will be a fluorescence signal present. The strength of the ribosomal binding site was modified to lower the expression of the *padR* in *E. coli*, demonstrating the specific induction up to 130-fold in the presence of 2 mM *p*-coumaric acid. Additionally, this biosensor construct was transformed into a

Corynebacterium glutamicum mutant, which is unable to convert *p*-coumaric acid into *p*-coumaroyl-CoA, whereas wild-type *C. glutamicum* manages to implement this reaction. The biosensor was tested with both bacteria and demonstrated the fluorescence response in both [33]. This result suggests that the biosensor responds precisely to *p*-coumaric acid and not to *p*-coumaroyl-CoA, to which this phenolic acid is generally converted.

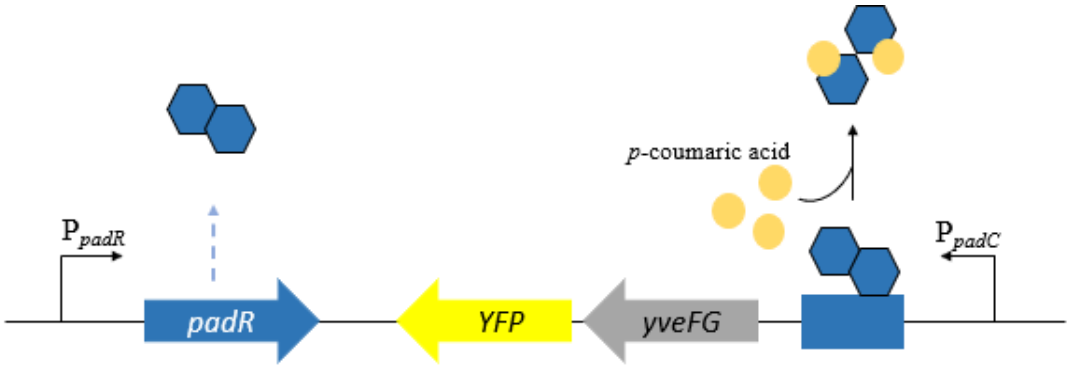


Fig. 1.4. The principal scheme of *padR* regulatory system

In conclusion, due to a wide application field of *p*-coumaric acid and a demonstrated potential to construct a robust and effective biosensor, there is a need to identify other microorganisms to use PadR as a TF, such as *Lactobacillus plantarum* or *Bacillus megaterium*. Besides, out of 400 existing PadR-like regulators, only a few were already investigated [12]. Other isomers of *p*-coumaric acid, such as *m*-coumaric and *o*-coumaric acids, should be also investigated furtherly to be applied in genetically encoded biosensors, as no positive induction effects were described in the literature yet.

1.3.2. Transcription factor-based biosensors of caffeic acid

Caffeic acid, also known as 3,4-dihydroxycinnamic acid, is naturally found in coffee, olive oil, vegetables, fruits, and wine [2], [56]. This phenolic acid is known for its high antioxidative, anti-inflammatory, anti-microbial, anti-viral, and anti-carcinogenic properties [17]. Moreover, caffeic acid inhibits the degradation of lipids, eliminates free radicals [36], and is used for the biosynthesis of lignin. However, the biosensors for detection of caffeic acid described in the literature are usually amperometric or electrochemical [2], still lacking an optimal and universal caffeic acid-inducible genetically encoded biosensor.

To create a biosensor for the detection of lignin-derived substrates, a whole-cell biosensor, sensitive to caffeic acid, was developed [46] since the detection of lignin degradation monomers, including *p*-coumaric, ferulic, and caffeic acids, enable the monitoring of its degradation. The created biosensor is based on a FerC repressor, which normally binds to the DNA sequence upstream of the *ferB* gene, encoding feruloyl-CoA hydratase in *Sphingobium* sp. SYK-6 [57]. FerC was inserted upstream of the GFP (green fluorescent protein) reporter gene. However, this biosensor was activated not only by caffeic but also by 1 mM ferulic, sinapic, umbellic, 5-hydroxyferulic, and iso-ferulic acids. Results demonstrated moderate activation with an 11.2-fold dynamic range after adding 1 mM caffeic acid and induction with 13 other compounds [46]. Although this biosensor being non-caffeic acid-specific, it could be applied for the optimisation of various chemical or enzymatic processes, including the degradation of lignin [43]. It could be also utilised in the processes using the mentioned inducing molecules as substrates for production of high-value chemicals, such as flavonoids and vanillin, and to investigate the substrate/product transport systems across biological membranes [46]. Overall,

there is still a need to develop a specific caffeic acid-inducible biosensor, which could contribute to investigate the metabolism pathways of this phenolic acid or help to maximise its production yields.

1.3.3. Transcription factor-based biosensors of salicylic acid

Salicylic acid, also known as *o*-hydroxybenzoic acid, was usually used as an inducer due to its stability under normal growth conditions and the ability of *E. coli* not to metabolise this compound [58]. Nevertheless, low concentrations of salicylic acid should be detected when no inhibition of bacterial growth is observed [59], therefore complicating the investigation of inducible systems sensitive to this phenolic acid.

The salicylic acid-inducible system was predicted based on two operons called *nah* and *sal* in *Pseudomonas fluorescens* HK44. Analysis of gene expression in *P. putida* has shown that NahR acts as a regulator [60], activating the transcription of *nah* and *sal* genes in the presence of salicylic acid. Insertion of a bioluminescent reporter in *Pseudomonas fluorescens* HK44 generated a luminescence induced by salicylic acid, which is also included in its metabolism. However, strain HK44 causes a problem in real sample biosensing applications because the selectivity of bioluminescent response is not highly specific [49]. On the other hand, one study revealed that using a NahR-based inducible system increased RFP (red fluorescent protein) output by 292-fold after the addition of salicylic acid, which was an exceptionally strong RFP expression. Moreover, this system was transformed into two different host microorganisms – *E. coli* Top10 and *P. putida* KT2440, demonstrating the activation of the system and higher than 75-fold induction in both tested microorganisms. Salicylic acid was also shown to be a specific inducer in this genetically encoded biosensor [15], demonstrating specific applicability in diverse hosts.

Considering that salicylic acid is economically advantageous, as it is relatively low in price as an inducing agent, it is not metabolised in *E. coli*, and stable at a variety of temperatures, it could be a beneficial tool in the field of synthetic biology [59]. However, there is a need to investigate the isomers of salicylic acid as inducing compounds, as most of the studies are carried out specifically on *o*-salicylic acid. The degradation pathway of *m*-salicylic acid (3-hydroxybenzoic acid) is encoded by five genes called *mhbRDHMI*, which are clustered in the order regulator-gentisate dioxygenase-furnaryl pyruvate hydrolase-3-hydroxybenzoate monooxygenase-maleylpyruvate isomerase. These genes encode the conversion of 3-hydroxybenzoic acid to fumaric acid and pyruvate in *Klebsiella pneumoniae.* Moreover, salicylic acid was proven to be a non-metabolisable inducer analogue in such a pathway. *E. coli* was used as a host of an application, and this degradation pathway allowed it to grow on *m*-salicylic acid [61]. However, the biosensors of *m*-salicylic acid are still not applied in metabolic engineering and should be further investigated.

Moreover, a genetically encoded biosensor was developed for the detection of 4-hydroxybenzoic acid (*p*-salicylic acid). Genes of *Rhodopseudomonas palustris* encode enzymes responsible for the conversion of *p*-salicylic acid to benzoyl-CoA, regulated by HbaR, which is a member of the FNR-CRP superfamily of TFs. HbaR is related to the transcriptional regulators, which are involved in the regulation of nitrate reduction. It has been shown that HbaR regulates the *hbaA* gene, which encodes *p*-salicylic acid-CoA ligase – the first enzyme of the *p*-salicylic acid degradation pathway. Moreover, HbaR activated the expression of *hbaA-lacZ* fusion in *Pseudomonas aeruginosa* cells, which were cultivated aerobically, thus HbaR not being sensitive to oxygen [50] and demonstrating a potential to be applied in genetically encoded biosensors.

1.3.4. Transcription factor-based biosensors of vanillic acid

Vanillic acid is low in toxicity, and it is an intermediate compound for the production of vanillin, which is one of the main flavour additives by volume worldwide [51]. It is being used in the food, beverage, cosmetics, and pharmaceutical industries [62]. Vanillic acid demethylase is encoded by the *vanAB* gene, and it catalyses the conversion of vanillic acid to protocatechuic acid [63]. Nevertheless, this enzyme is remarkably air-sensitive and unstable, thus a measurement of its activity is complicated [64]. Moreover, the regulator of *vanAB* expression is known to be a member of the GntR family of repressor proteins, called VanR [63]. A significant application of vanillic acid and already known regulatory genes of vanillic acid metabolism arise an opportunity to develop a vanillic acid-inducible whole-cell biosensor.

Vanillic acid-inducible biosensor, consisting of the *Caulobacter crescentus* VanR-VanO system, was designed in *E. coli.* Such construct demonstrated excellent substrate selectivity with no response to iso-vanillic acid and only minimal response to structurally similar pathway intermediates, including protocatechuic acid, protocatechuic aldehyde, and vanillin [51]. VanR exists in distinct bacteria, such as *Acinetobacter, Caulobacter, Corynebacterium, Myxococcus,* and *Pseudomonas* [65]. Various repressor variants bind to different operator *vanO* sites. In a developed biosensor, the VanR-VanO system was linked to the expression of *RFP*. The most optimal variant demonstrated high sensitivity to vanillic acid and exquisite molecular specificity of the VanR in selected dose-responses occurring from 10 μM to 100 μM, which indicated a full activation of the biosensor [51] and very high sensitivity. Considering that vanillic acid biosensor produces proportional fluorescence at sub-millimolar concentrations of substrate, possesses low background fluorescence, and does not demonstrate induction on similar molecules, it could be applied in metabolic engineering.

1.3.5. Transcription factor-based biosensors of protocatechuic acid

Protocatechuic acid is also known as 3,4-dihydroxybenzoic acid [66] and, as well as vanillic acid, it is low in toxicity, therefore attracting investigation as a chemical inducer of biosensors [51]. Moreover, direct production of protocatechuic acid could compensate petrochemicals and contribute to the production of synthetic products [52]. However, some of the protocatechuic acid-inducible genetically encoded biosensors are known to be also activated in the presence of vanillic acid [66], therefore several mistakes can emerge while investigating such biosensors.

A biosensor for the detection of protocatechuic acid was designed and developed to improve nylon synthesis and apply biosensors in the biotechnological industry since protocatechuic acid is an essential precursor of this polymer [52]. One of the most investigated protocatechuic acid-inducible systems was applied, which is based on the catabolic pathway of this phenolic acid in *Acinetobacter sp*. strain ADP1. In this pathway, PcaU, which is a member of the IclR family of TFs, binds to the promoter P_{pca} enabling RNAP to initiate the transcription of the structural *pca* genes [66]. Identified inducible gene expression system was applied by transforming it into *E. coli* and using GFP as a reporter protein. The developed biosensor was sensitive to protocatechuic acid at >5 μM, and it was investigated in combination with FACS. Interestingly, a second plasmid containing AsbF transcriptional regulator from *Bacillus cereus* was transformed into the same *E. coli* cell. After the addition of protocatechuic acid, the double transformant demonstrated not only the presence of this phenolic acid but also the functioning of dehydroshikimate dehydratase, converting dehydroshikimate to protocatechuic acid [52]. Protocatechuic acid was shown to increase cell density

more than twofold, which is observed due to ligand catabolism. It is known that the consumption of the inducer can affect the kinetics of induction [15], which must be estimated while investigating protocatechuic acid-inducible biosensors.

1.3.6. Transcription factor-based biosensors of ferulic acid

Ferulic acid, also known as 4-hydroxy-3-methoxy-cinnamic acid, is characterised by antimicrobial and antifungal properties. There is a growing interest in the application of ferulic acid to obtain valuable compounds, such as vanillin [62]. Microbial pathways of catabolism of ferulic acid are therefore widely investigated [67], contributing to the development of functional ferulic acid-inducible biosensors.

As mentioned above, the biosensor developed to achieve the lignin valorisation process though being non-specific, but it demonstrated the 26.2-fold increase in GFP fluorescence after adding 1 mM ferulic acid. Ferulic acid is converted into feruloyl-CoA, which de-represses *ferC*, therefore activating the GFP expression. Moreover, the analysis of the dose-response was performed of the three developed biosensor variants, which were constructed by varying promoter-operator sequences. The first one demonstrated the saturation at ≥100 μM of ferulic acid, along with incomplete de-repression, corresponding to ~85% of the reporter's output. The second one indicated complete de-repression at ≥40 μM of ferulic acid, and the third variant was fully de-repressed but had limited utility because of the narrow signal range. Therefore, the second biosensor was more suitable for the detection of ferulic acid, but greater efficiency was observed with the first one [46]. Such results indicate that there is a potential to develop ferulic acid-inducible whole-cell biosensors, which could be furtherly modified to obtain a more specific signal.

1.4. Justification of the project aim and objectives

It is evident from the literature review that phenolic acids, including gallic and *p*-coumaric acids, are important bioproducts. These phenolic acids are extensively used in various industries for their beneficial properties such as antioxidant, antibacterial, and anti-inflammatory. This project was aimed at the identification of gallic and *p*-coumaric acid-inducible systems to expand a range of genetic and analytical tools for controlling phenolic acid metabolism and relevant metabolite monitoring. The identified inducible systems were aimed to be applied and characterised as genetically encoded biosensors, enabling the measuring at the single-cell level, high-throughput screening, and detection of phenolic acids in real-time.

2. Materials and methods

2.1. Equipment, tools, and materials

The equipment and tools used for this research (supplier, country): laminar flow cabinet (ESCO, USA), micropulser (Bio-Rad, USA), block heater (VWR, USA), vortex mixer (Vortex V-1 plus Biosan, Latvia), shaking incubator MaxQ 6000 (Thermo Fisher Scientific, USA), incubator (Memmert, Germany), centrifuge (Eppendorf, Germany), refrigerated centrifuge (Eppendorf, Germany), heated bath (Grant Instruments, UK), DNA electrophoresis apparatus POWER PRO-300 (Cleaver Scientific, UK), DNA electrophoresis bath (Cleaver Scientific, UK), UVITEC Essential Imaging System (UVITEC, UK), UV lamp Safe Light-Box 20-blue Q9 PLUS (UVITEC, UK), scales (Sartorius Lab Instruments, Germany), Nanophotometer N60 (Implen, Germany), spectrophotometer (Thermo Fisher Scientific, USA), Tecan Infinite 200 Pro plate reader (TECAN, Austria), 96-well black plates clear bottom with Lid tissue culture treated polystyrene (Corning Incorporataed, USA), microbank vials (PRO-LAB DIAGNOSTICS, Canada), thermal cycler Mastercycler Nexus X2 (Eppendorf, Germany), autoclave (CertoCLAV Labor-Autoklav, Austria), 0.2 cm gap width electroporation cuvettes (Bio-Rad, USA).

Chemical materials used for this research (supplier, country): Phusion™ High–Fidelity DNA Polymerase (Thermo Fisher Scientific, USA), GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich, USA), GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, USA), NEBuilder HiFi DNA Assembly Master Mix (New England BioLabs, UK), Zymoclean Gel DNA Recovery Kit (Zymo Research Corporation, USA), DreamTaq Green PCR Master Mix 2X (ThermoFisher Scientific, USA), LB Agar Miller (Fisher Bioreagents, Belgium), LB Miller (Fisher Bioreagents, Belgium), SOB medium (Hanahan's Broth, Sigma), TAE Buffer 50X (Thermo Fisher Scientific, Lithuania), agarose (Sigma-Aldrich, USA), GeneRuler 1 kb DNA Ladder (Thermo Fisher Scientific, Lithuania), loading dye TriTrack 6X (Thermo Fisher Scientific, Lihuania), loading dye for agarose gel SYBR Safe (Thermo Fisher Scientific, USA), HEPES (Sigma, USA), NaOH (Eurochemicals, Lithuania), glycerol (Eurochemicals, Lithuania), DMSO (Eurochemicals, Lithuania), ethanol (MV group production, Lithuania), chloramphenicol (TCI, USA), tetracycline (Sigma-Aldrich, USA), FastDigest restriction enzymes-endonucleases (Thermo Fisher Scientific, Lithuania). All phenolic acids used in this study are listed in Table 2.1.

Table 2.1. Phenolic acids used in this study

Chemical	Supplier	Catalogue number
p-Hydroxybenzoic acid	Fluorochem	047887
o-Salicylic acid	Sigma-Aldrich	71945-250G
m-Hydroxybenzoic acid	Sigma-Aldrich	H20008-100G
Vanillic acid	Sigma-Aldrich	94770-10G
Iso-vanillic acid	Alfa Aesar	A13709
Gallic acid hydrate	Fluorochem	242843
Protocatechuic acid	Alfa Aesar	B24016
Syringic acid	Alfa Aesar	A11725
Gentisic acid	Sigma-Aldrich	149357-25G
α-Resorcylic acid	Sigma-Aldrich	D110000-100G

β-Resorcylic acid	Sigma-Aldrich	D109401-100G
γ-Resorcylic acid	Sigma-Aldrich	D109606-25G
Orsellinic acid	Acros Organics	A0411384
6-Methylsalicylic acid	Acros Organics	341500
o-Coumaric acid	Sigma-Aldrich	H22809-5G
m-Coumaric acid	Sigma-Aldrich	H23007-5G
p-Coumaric acid	Sigma-Aldrich	C9008-25G
Ferulic acid	Sigma-Aldrich	128708-25G
Sinapic acid	Alfa Aesar	A15676
Chlorogenic acid	TCI	C0181

2.2. Strains and plasmids

Bacterial strains used in this final project are represented in Table 2.2. All plasmids designed and constructed in this study are described in Tables 3.3 and 3.4 (see Results and discussion section).

Table 2.2. Bacterial strains used in this study

Bacterial strains	Characteristic	Use in this study	Supplier
Escherichia coli Top10	F- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Φ80*lacZ*ΔM15 Δ*lacX*74 *recA*1 *araD*139 Δ(*araleu*)7697 *galU galK rpsL* (StrR) *endA*1 *nupG*	For plasmid propagation and cloning	Thermo Fisher Scientific
Cupriavidus necator H16	Wild type strain	For cloning	DSM 428
Pseudomonas putida KT2440	Wild type strain	For cloning	DSM 291

2.3. Methods

2.3.1. Bioinformatics

All information provided in this final project is based on the published literature while the plasmid, the sequences of insert genes and promoters were found at the National Center for Biotechnology Information (NCBI, ncbi.nlm.nih.gov) [68] and analysed using the Basic Local Alignment Search Tool (BLAST, blast.ncbi.nlm.nih.gov/Blast.cgi). Usually, this program is used for various comparisons, such as amino-acid sequences of proteins or RNA sequences of nucleotides [69]. However, in this project, it was applied for comparing DNA or protein sequences, enabling the comparison of the DNA sequence of interest (the query) with the ones in a database, indicating the resembling sequences at a respective percentage.

The collected information revealed the specifics about the relevant TF, such as the bacteria from which it originates, locus tag, DNA sequence, protein sequence, and further information. Additionally, it helped to elucidate the DNA sequence of the plasmid of interest, thus enabling the visualisation of construction using the SnapGene program (from Insightful Science; available at snapgene.com), which also simulated the DNA manipulations readily, alerting to errors before they practically occurred. Subsequently, all variations of plasmids and their inserts were automatically recorded, which helped to observe and save the variations of the constructs.

Intergenic regions were analysed using a Sequence Logo Generator (WebLogo, weblogo.berkeley.edu/logo.cgi), which generates graphical representations of the patterns within a multiple sequence alignment. Such simulations provide a clear and precise view of sequence similarity, enabling to readily reveal essential features of the alignment, which would be difficult otherwise to visualise. The height of each logo letter is proportional to the frequency of the corresponding nucleotide of the analysed sequences [70]. Moreover, the RNAP binding sites – promoter sequences containing -10 and -35 regions and transcription start sites, were predicted using the Neural Network Promoter Prediction tool [71].

2.3.2. Cell cultivation

Bacterial cells were cultivated in Luria Bertani (LB) medium, containing the appropriate antibiotic. The powder of LB medium was weighed at 40 g/L and dissolved in sterile water, autoclaved at 120°C for 15 min, and after cooling down the medium to 60°C, the appropriate antibiotic was added. *E. coli* cells were cultivated in LB medium, containing 25 µg/mL of chloramphenicol at 37°C with aeration, at 225 revolutions per minute (rpm) for 16-20 hours. *P. putida* and *C. necator* bacteria were cultivated in LB medium, containing 10 µg/mL of tetracycline or 50 µg/mL of chloramphenicol, respectively, and incubated at 30°C with aeration at 225 rpm for 16-20 hours. Solid medium was prepared by adding 15 g/L agar.

2.3.3. Microbial genomic DNA purification

GenElute Bacterial Genomic DNA Kit was used for genomic DNA purification. After 16 hours of cultivation, the tube with a grown culture was centrifuged at 8000 relative centrifugal force (rcf) for 5 min in a refrigerated centrifuge. The supernatant was discarded, the pelleted cells were resuspended with 180 µL of Lysis Solution T and mixed by vortexing. Then, 20 µL of Rnase A was added, mixed, and incubated for 2 min at room temperature. Afterward, 20 µL of Proteinase K was added, mixed, and incubated at 55 °C for 30 min. After incubation, 200 µL of Lysis Solution C was added, the tube was mixed by vortexing for 15 s and incubated at 55 °C for 10 min. Genomic DNA Purification Column was assembled with a collection tube, 500 µL of Column Preparation Solution was added and centrifuged at 12000 rpm for 1 min. The flow-through was discarded, 200 µL of ethanol (95%) was added to the lysate and mixed thoroughly by vortexing for 10 s. The contents of the tube were transferred into the prepared column and centrifuged at 6500 rpm for 1 min. The collection tube was discarded, and the column was placed into a new collection tube. Furtherly, 500 µL of Wash Solution 1 was added to the column and centrifuged at 6500 rpm for 1 min, the flow-through was discarded, and 500 µL of Wash Solution Concentrate (diluted with ethanol) was added to the column and centrifuged at 12000 rpm for 3 min. The flow-through was discarded, and an empty column was centrifuged at 12000 rpm for an additional 1 min to completely remove the residual ethanol. Subsequently, the collection tube was discarded, and the purification column was placed into a new collection tube. Then, 200 µL of the Elution Solution was pipetted directly onto the center of the column and centrifuged at 6500 rpm for 1 min to elute the DNA. The concentration of genomic DNA was measured by using Nanophotometer N60.

2.3.4. Plasmid DNA purification

Plasmid DNA from *E. coli* was purified using Thermo Scientific GeneJET Plasmid Miniprep Kit to perform cloning. Overnight cultures of *E. coli* were centrifuged at 8000 rcf for 2 min in a refrigerated centrifuge. The supernatant was discarded, the pelleted cells were resuspended in 250 µL of the

Resuspension Solution, and the suspension was transferred to a microcentrifuge tube. Then, 250 µL of the Lysis Solution was added and mixed thoroughly by gently inverting the tube 4-6 times. Afterward, 350 µL of the Neutralisation Solution was added and mixed by inverting the tube 4-6 times, and the tube was centrifuged at 14000 rpm for 8 min. The supernatant was transferred to the Miniprep Kit Column by pipetting and avoiding the transfer of white precipitate. The column was centrifuged at 12000 rpm for 1 min, the flow-through was discarded, and 500 µL of the Wash Solution (diluted with ethanol by ratio 1:4) was added, and the column centrifuged at 12000 rpm for 30 s. The flow-through was discarded, and 500 µL of the Wash Solution was added again and centrifuged at 12000 rpm for 30 s. The flow-through was discarded, and an empty column was centrifuged at 12000 rpm additionally for 2 min to remove the residual ethanol. The column was transferred to a new collection tube, and 20 µL of the Elution Buffer, heated up to 37°C, was added directly to the center of the column membrane. The column was centrifuged at 12000 rpm for 1 min to elute the plasmid DNA. The concentration of the purified plasmid DNA was measured by using Nanophotometer N60.

2.3.5. DNA amplification

DNA amplification was performed by using a simple molecular biology technique – polymerase chain reaction (PCR), enabling to make a lot of copies of a specific DNA fragment. However, as the design of appropriate primer pairs is critical for the success of PCR because they are complementary to the DNA target region and initiate the elongation of the double-stranded DNA, they were therefore designed formerly and synthesised by Thermo Fisher Scientific, Inc., UK (Table 2.3). Primers were diluted with nuclease-free water up to 100 µM, and thereafter dilutions with water by a ratio of 1:10 were performed to achieve an operative concentration stock of primers.

Table 2.3. Primers used in this study

Primer name	Primer sequence (5‘→3‘)
EV001-PP_2515_for	GGGCCTTTCGTTTTATGACGTCTCACGCCTGCTCGGTGAT
EV001A_for	GGGCCTTTCGTTTTATGACGTCAAGGGCTCCCGGCGCTCGGCTCGTGA
EV002-PP_2515A_rev	CGTCTTCGCTACTCGCCATATGCCTGTGCAGGGCACTAATG
EV003-PP_2515B_rev	CGTCTTCGCTACTCGCCATATGTCGTTGGCTCCTGCAGTGG
EV003B-PP_2515B_rev	CAGGATGGCCTTCTGCTTAAGTTTCGTTGGCTCCTGCAGT
EV001A_for	GGGCCTTTCGTTTTATGACGTCAAGGGCTCCCGGCGCTCGGCTCGTGA
EV001B_for	CAGGATGGCCTTCTGCTTAACTTATCACGCCTGCTCGGTGAT
EV003C_rev	GGGCCTTTCGTTTTATGACGTCTTAGTTTCGTTGGCTCCTGCAGT
EV008B_rev	GGGCCTTTCGTTTTATGACGTCGGAAGGCCCTGTATCAATCCC
EV008-PP_rev	TTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACGGAAGGCCCTGTATCAATCCC
EV009-PP_for	CGTCTTCGCTACTCGCCATATGTGGTCACCTTTGTTCTTGTATTGG
EV009B-PP	CGTCTTCGCTACTCGCCATATGAAAGTCCTCTGTTACAGGTTACGG
IK003_AscI_345_rev	TCGTTTTATGGCGCGCCAGGCCGGCCAATTAGAAGGCCGCCAGAGAGG
IK004_PmeI_344_for	ACTAGTACTGTTTAAACCGCTCACAATTCCACACAA
EV025_rev	CGTCTTCGCTACTCGCCATATGTATTCCTTCTTTCATTAGAATGATGTTGTTTCAAACATGTC
EV026_for	GGGCCTTTCGTTTTATGACGTCAATAACTCTCATGATGTGCCTCC

EV027_for	GGGCCTTTCGTTTTATGACGTCCCGACCATTGTCCTTAGAGC
EV028_rev	CGTCTTCGCTACTCGCCATATGACACACTCTCCTTAGTCTTTTACTG
EV029_for	GGGCCTTTCGTTTTATGACGTCTAGCTTCAGACAAGGACTGTCT
EV030_rev	TTTGCGTTTCTACCGAGTGGTAGCTTCAGACAAGGACTGTCT
EV031_for	CCACTCGGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAATTAGGCTAAAACAGATAGAATCAGCG
EV031B_for	GGGCCTTTCGTTTTATGACGTCGGCTAAAACAGATAGAATCAGCG
EV032_rev	AGGTTAAATGTTTTGTCTGGCTGTTTCGCAAAACA
EV033_for	AGCCAGACAAAACATTTAACCTAATACTACATTAGACATAATGAGAGTATTAAAATACGCCATATTAGGGCT
EV034_rev	GGGCCTTTCGTTTTATGACGTCAGCCTGAACTGCGGTTAATC
EV034B_rev	CCACTCGGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAACAGCCTGAACTGCGGTTAATC
EV035_rev	CGTCTTCGCTACTCGCCATATGTATATCGCTCCTTTGATTTATAACACTTTGTATATGTCT
EV036_for	GGGCCTTTCGTTTTATGACGTCTTAAGCACTCATAGCATTCAACCTCT
EV037_for	GCGTTTCTACAGCACTCATAGCATTCAACCTCT
EV038_rev	CTATGAGTGCTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAATCCTCAAATTAGTTCGAACACTCA
EV038B_rev	GGGCCTTTCGTTTTATGACGTCTCCTCAAATTAGTTCGAACACTCA
EV039_for	GGAAGAAATAATGAAACATTTAACCTAATACTACATTAGACATAATGAGAGTATTAAAATATGCCATATTAGGGCTCT
EV040_rev	GGTTAAATGTTTCATTATTTCTTCCATCAGTATAGCA
EV041_for	GGGCCTTTCGTTTTATGACGTCTTAAGTCATGATCCGCTCACCT
EV041B_for	CTATGAGTGCTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAAAGTCATGATCCGCTCACCT
EV074_for	GGGCCTTTCGTTTTATGACGTCTTACTTGTCGATCGGCAATGGAG
EV075_rev	CGTCTTCGCTACTCGCCATATGTATATGTCCTTTTGATCCGTTGGC
EV076_for	GGGCCTTTCGTTTTATGACGTCTTACATATAGCCCAGGCGCAGTG

The PCR was performed to obtain the proper fragments to be ligated or assembled with NEBuilder HiFi DNA Assembly kit, generating the required constructs. The PCR tube was set up on ice, and the following were added: 0.5 µM of each primer, 0.02 U/µl of Phusion™ High–Fidelity DNA Polymerase, 3% of the final concentration of DMSO, 100 ng of genomic template DNA or 10 ng of plasmid DNA, 10 mM of dNTPs (final concentration was 200 µM of each), and 4 µl of 5x Phusion Buffer. Then, nuclease-free water was added up to 20 µL. Reaction mixtures were gently mixed and placed for PCR reaction into a thermal cycler using the recommended conditions (Table 2.4). The annealing temperature and extension time was modified depending on the size and structure of the fragment to be amplified and the primers used. The annealing temperature was set to T_m value of primers minus five, and extension time was set to depending on the size of the fragment to be amplified, for instance, 60 s of an extension was used to generate a 1000 bp long insert. The reaction products were visualised with agarose gel electrophoresis.

Table 2.4. Thermal cycling conditions for PCR

Step	Initial Denaturation	Denaturation	Annealing	Extension	Final extension	Hold
Temperature	98 °C	98 °C	63 °C	72 °C	72 °C	15 °C
Time	120 s	10 s	30 s	90 s	10 min	
Number of cycles	1	25-40			1	

2.3.6. Agarose gel electrophoresis

In this final project, agarose gel electrophoresis assisted in the purification of the digested fragments. Moreover, it assured that either PCR reaction was performed correctly, or plasmids were properly constructed. Since samples are loaded into wells of an agarose gel, which is subjected to an electric field, negatively charged nucleic acids move toward the positive electrode. Shorter DNA fragments move more rapidly, whereas longer ones remain closer to the negatively charged electrode, which results in separation based on DNA fragment size [72].

Firstly, 1% of agarose gel solution was prepared by dissolving 1 g of agarose in 100 mL of TAE buffer (diluted with distilled water at 1:50) and heated up until the agarose completely dissolves. Then, agarose gel was cooled down to 60°C and poured into a casting tray along with special combs and 1 µL of loading dye, which ensures the visualisation of DNA fragments. The gel was held for 30 min until it cooled down and solidified, thus creating a gel slab with a row of wells at the top. The tray with a ready-to-use gel was placed into a TAE buffer-filled chamber, where TAE buffer provided ions for setting up an electric field. The gel slab was positioned that the wells would be closest to the negative electrode of the chamber. When visualising the properly digested plasmids or ensuring the success of colony PCR, 5 µL of each DNA sample was loaded into a respective chamber well. Also, 5 µL of the Generuler Ladder was loaded into one of the wells creating a DNA ladder with the known size fragments, enabling to assess the size of the fragment of interest. Subsequently, the positive and negative leads were connected to the chamber and a power supply, providing 120 V of power, and creating an electric field. Negatively charged DNA samples migrated through the gel towards the positive electrode. The results were visualised under UV light.

2.3.7. Extraction of DNA fragments from agarose gel

The purification of the amplified fragments or digested backbone vectors from agarose gel was implemented using a Zymoclean Gel DNA Recovery Kit. In the case of purification, agarose gel was made with the wider combs. The total reaction volume of the fragments to be purified was loaded into the corresponding wells along with 2 µL of a loading dye. After the migration of the fragments, the results were visualised on the UV lamp, and gel slices were excised using a sterile scalpel. The gel slices were placed into the pre-weighed tubes, and the weights of the slices were recorded. Three volumes of agarose dissolving buffer (ADB) was added to each volume of agarose excised from the gel, for instance, 300 µL of ADB was used for 100 mg of agarose gel slice. The mixture was incubated at 55°C for 5-10 min and vortexed until the gel slice completely dissolved. The melted agarose solution was transferred into a Zymo-Spin™ Column in a Collection Tube and centrifuged for 60 s at 10000 rcf, and the flow-through was discarded. Afterward, 200 µL of DNA Wash Buffer (diluted with ethanol by ratio 1:4) was added to the column and centrifuged for 30 s under the same conditions. The flow-through was discarded and the wash step repeated. Then, an empty column was centrifuged

again for 2 min to completely remove residual ethanol. Subsequently, 12 µL of DNA Elution Buffer was added directly to the column matrix. The column was placed into a new tube and centrifuged for 60 s at 10000 rcf to elute the DNA. The concentration of the purified DNA was measured using Nanophotometer N60.

2.3.8. Plasmid construction

Plasmid construction included restriction digestion, agarose gel electrophoresis for fragments extraction, and ligation or HiFi reaction. The schematic view of constructs assembly using either ligation or HiFi assembly method is presented in Appendices 1 and 2, respectively.

2.3.8.1. Restriction digestion

In this study, a restriction digestion procedure was applied to prepare DNA fragments for ligation or HiFi DNA assembly reaction. The use of specific enzymes – restriction endonucleases (Reases) enabled to cut the DNA sequence at specific locations, called restriction sites, in such a manner that the digestion result is suitable to join the desired DNA fragments with suitable plasmids. Reases are naturally occurring bacterial enzymes that recognise a high range of DNA sequences, thus there is a variety of endonucleases and the unique sites they recognise. The selected Reases for the construction of plasmids were four different enzymes, which ensured the correct orientation of the insert and prevent the plasmid from ligating to itself during ligation [73].

At first, the backbone plasmid pBRC1 was digested using Reases AatII, which cuts at the palindromic GACTGT*C site, and NdeI, which cuts at the palindromic CA*TATG site (Figure 2.1). In the case of ligation, the same Reases were used to digest the amplified inserts. The usage of such a pair of Reases allowed the ligation or HiFi reaction to assemble constructs. Moreover, Reases PmeI (cutting at GTTT*AAC) and AscI (cutting at GG*CGCGCC) were also used in this study to cut and replace the antibiotic-resistance genes when required.

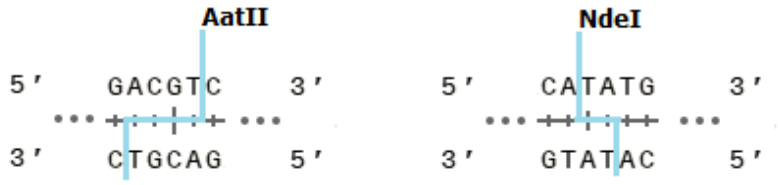


Fig. 2.1. AatII and NdeI restriction sites

Restriction mixture for digestion of the backbone plasmid was produced by mixing 1000 ng of the plasmid (pBRC1), 1/10 of the total volume of Green Buffer 10X, 1 µL of each specific Rease, and nuclease-free water up to 20 µL. Whereas the restriction mixture for the treatment of the insert contained 300 ng of the amplified insert, 1/10 of the total volume of Green Buffer 10X, 1 µL of each specific Rease, and nuclease-free water up to total volume (10 µL). Reaction mixtures were incubated at 37°C for 16-20 hours. Then, the mixtures were incubated at 80°C for 5 min to heat-inactivate the AatII and NdeI, and in the case of AscI and PmeI, the mixtures were incubated at 65°C for 20 min for heat-inactivation.

2.3.8.2. Ligation

The ligation reaction was performed using digested pBRC1 plasmid and amplified inserts, allowing to assemble the plasmids correctly. PCR product and plasmids were digested with the same Reases –

AatII and NdeI, generating sticky ends, which must stick together upon enzyme ligase T4 action, because of the formation of covalent phosphodiester linkages joining the nucleotides together.

At first, agarose gel electrophoresis was performed as described above to ensure that restriction digestion is performed successfully and enable the extraction of the digested products out of the gel. Next, fragments of the insert and plasmids were cut out of the gel and purified as described. The required ligation reaction volumes were calculated using a ligation calculator (Insilico, available at http://www.insilico.uni-duesseldorf.de), applying molar plasmid and insert ratio – 1:5, respectively. The ligation reaction mixture was prepared by adding 100 ng of digested backbone plasmid and the calculated amount of the insert, 2 µL of T4 Ligase buffer, 1 µL of T4 ligase, and nuclease-free water up to 20 µL. The ligation mixture was incubated for 1 hour at room temperature. Negative control was also performed to ensure the success of the ligation reaction by adding every same component except for the insert into another tube.

2.3.8.3. HiFi DNA assembly

Alternatively, NEBuilder HiFi DNA Assembly Reaction was implemented instead of the ligation reaction, providing that it is superior when inserting two or more fragments into a backbone plasmid. HiFi DNA Assembly method is based on the conjugation of overlapping DNA fragment ends, even those with 5'- and 3'-end mismatches, by receiving the fully assembled DNA. The principle of this method is that exonuclease cuts 5' ends, creating 3' overhangs, then DNA polymerase fills in gaps and
DNA ligase seals nicks in the assembled DNA strand. This high-fidelity assembly method provides a more effective way to connect two or more DNA fragments since there is no need to digest the PCR product, the fragment to be inserted. To insert one to three fragments into a plasmid, 5 µL of NEBuilder HiFi DNA Assembly Master Mix was added into the tube, along with 0.03-0.2 pmols of each fragment and 50 ng of digested backbone plasmid (plasmid and insert ratio 1:2, respectively), as well as nuclease-free water up to 10 µL. Total amounts of components were calculated using ligation calculator. The mixture was incubated at 50°C for 60 min and placed at -20°C until the subsequent steps to be performed.

2.3.9. Preparation of competent cells

To transfer the created constructs into bacterial cells, they must be prepared foremost. Since bacterial cells are not capable of naturally receiving plasmids from the environment, there is a need to increase the permeability of their membranes and make them competent – capable of taking up external DNA. In this final project, three diverse microorganisms were used for cloning, and all of them were prepared by applying different methodologies. The essential thing is to keep the temperature low when preparing competent cells, as they are very susceptible and fragile, and the cold helps to avoid cell death during processing.

Competent *E. coli* Top10 cells were prepared by incubating the culture on LB agar plates overnight at 37ºC. Then, they were cultivated in 5 mL of LB medium overnight in a shaking incubator at 37°C and 200 rpm. Then, 0.5 mL of the overnight culture was added to 50 mL of LB medium and incubated in a shaking incubator at 37°C and 200 rpm to OD_{600} = 0.4-0.8. The culture was cooled down on ice for 10 min and then centrifuged for 6 min at 4000 rpm. The supernatant was discarded, cells were resuspended in 15 mL of 0.1 M $MgCl_2$, and once again centrifuged at the same conditions. Afterward, the supernatant was discarded, cells were resuspended in 15 mL of 0.1 M $CaCl_2$ and left on ice for 20

min. The cells were centrifuged for 6 min at 4000 rpm, the supernatant was discarded, and cells were resuspended in 3 mL of 0.1 M $CaCl_2$ – 15% glycerol solution. The prepared cells were frozen at -80°C in aliquots of 200 µL.

Equally, electro-competent *Cupriavidus necator* H16 cells were prepared by incubating the culture on LB agar plates overnight at 30ºC, and 10 mL of SOB medium was inoculated with a heavy loop of fresh cells and vortexed thoroughly. The mixture was diluted to 10^{-3} and incubated in a shaking incubator at 30ºC and 200 rpm per night. After incubation, the OD_{600} of the suspensions was measured, 25 mL of SOB medium was inoculated into the culture to an OD_{600} of 0.055-0.075, and suspensions were incubated repeatedly for 2 hours at 30ºC and 250 rpm. Thereafter, the culture was centrifuged at 7000 rpm for 10 min at 4 ºC, the supernatant was discarded and the harvested cells were washed with a pre-chilled sterile 10 mL of 1 mM HEPES (pH 7.0, adjusted with NaOH) and centrifuged as described above. The supernatant was discarded, the wash step was repeated with 5 mL of 1 mM HEPES and centrifuged as before, then the supernatant was discarded once again, and the pellet was resuspended in a sterile 1 mL solvent, composed of 9 mL of 1 mM HEPES buffer and 1 mL of 100% glycerol to an OD_{600} of 5. Subsequently, the aliquots of 100 µL of cell suspensions were frozen at -80°C.

Moreover, electro-competent *P. putida* KT2440 cells were prepared by cultivating fresh cells in 5 mL of LB medium at 30°C and 200 rpm per night. Then, 1 mL of the overnight culture was centrifuged for 5 min at 16 000 rpm and washed with 1 mL of ice-cold 10% glycerol, centrifuged for 5 min at 4000 rpm, the supernatant was discarded, and the wash step was repeated. Finally, 1 mL of ice-cold 10% glycerol was added to the suspension, and the aliquots of 200 µL of cell suspensions were frozen at -80°C [74].

2.3.10. Transformation and screening

After construction, the proper fragments were inserted into the plasmids, and the constructs could be transferred into bacterial cells for propagation, a process called transformation. In this study, the transformation was implemented in *E. coli* cells using the heat-shock method, whereas transformation to *P. putida* and *C. necator* was implemented applying the electroporation method. After transformation, the grown colonies indicated that the transformation was performed successfully. However, there is a need to ensure that the desired genetic insert is present in bacteria since many possible scenarios could have occurred, such as an improperly assembled plasmid containing more fragments or plasmid DNA lacking an insert. Therefore, plasmid screening was accomplished by performing colony PCR and secondary restriction digestion.

2.3.10.1. Transformation

The procedure was performed as 50 µL of competent *E. coli* cells, and 4 µL of ligation or HiFi mixture was added into a sterile tube. The mixture was resuspended, mixed, and incubated on ice for 5 min. Afterward, the transformation was performed by incubation of the mixture at 42°C for 1.5 min. Then the tube was placed on the ice and incubated once again for 5 min. Subsequently, 1 mL of LB medium was added, and the tube was incubated in a shaking incubator at 37°C for 1 hour at 225 rpm.

On the contrary, competent bacterial cells (either *P. putida* or *C. necator*) and the produced plasmid were mixed in a pre-chilled electroporation cuvette, and by using a Bio-Rad Micropulser the electroporation was conducted at 2.5 kV, 200Ω and 25µF. Immediately after electroporation, 0.9 mL

of SOC medium (SOB medium supplemented with 20 mM glucose) was added to the cell suspension, the mixture was transferred into a sterile 50 mL Falcon tube and incubated at 30°C for 2 h, at 250 rpm.

Subsequently, after incubation of the transformed *E. coli*, *P. putida*, or *C. necator* cells, the mixtures were centrifuged for 2 min at 8000 rcf. The supernatant was discarded, and 200 µL of cells were injected on a *Petri* dish, which was formerly prepared with a selective LB medium, containing an appropriate antibiotic. The injection was distributed through the plate with a sterilised glass spatula. Finally, the *Petri* dish was left in an incubator at 37°C per night for *E. coli*, or 30°C for *P. putida* and *C. necator*.

2.3.10.2. Colony PCR

The colony PCR demonstrates the output of a known product size only if the desired genetic insert is present since the same primers are adjusted for amplification of the insert. This method is beneficial because of the high sensitivity of PCR – the small amount of the template DNA gives a convenient visualisation on an agarose gel, as well as DNA amplification can be recovered from a very plain preparation of cells [75]. Therefore, colony PCR is an effective measure to readily and rapidly screen colonies to distinguish true positives from false positives.

The colony PCR was performed using the DreamTaq Green PCR Master Mix. The following were added into the PCR tube: 7.5 µL of DreamTaq Green PCR Master Mix (containing DreamTaq DNA polymerase, optimised DreamTaq Green buffer, which monitors electrophoresis progress, 4 mM $MgCl_2$, and dNTPs – 0.4 mM of each), 0.45 µL of DMSO, 0.75 µL of the forward primer and 0.75 µL of reverse primer, and nuclease-free water up to 15 µL. The reaction mixture was mixed with a sample of a bacterial colony and placed into a thermal cycler using the recommended conditions (Table 2.5). Agarose gel electrophoresis was performed as described above to visualise the size of PCR products.

Table 2.5. Thermal cycling conditions for colony PCR

Step	Initial denaturation	Denaturation	Annealing	Extension	Final extension	Hold
Temperature	95 °C	95 °C	54 °C	72 °C	72 °C	15 °C
Time	5 min	30 s	30 s	2 min	10 min	
Number of cycles	1	35			1	

2.3.10.3. Secondary restriction digestion

Since the size of the insert and plasmid backbone is known, secondary restriction digestion can be effectively used to verify the suitably assembled constructs. Secondary restriction digestion was performed by using the same Reases as adjusted in the assembling of the constructs. Colonies containing potentially correct constructs were inoculated into 5 mL of LB medium containing the appropriate antibiotic and incubated overnight at 37°C. The grown bacterial culture was centrifuged, and plasmid DNA purified as described above. Then restriction digestion was performed by adding 100 ng of purified plasmid DNA, 1 µL of Green buffer 10X, a pair of Reases 0.5 µL each, and nuclease-free water up to 10 µL into a tube. The reaction mixture was incubated for 2 hours at 37 °C. After incubation, Reases were heat-inactivated, and agarose gel electrophoresis was performed as

described above – two same fragments of correct length should be observed, indicating that bacteria possess the correctly assembled constructs.

2.3.11. Replacement of antibiotic resistance gene

The identified inducible system-based constructs were transferred to a few other microorganisms usually applied in synthetic biology, contributing to the development of an effective biosensor, which is suitable for diverse hosts, ensuring great optimisation, productivity, and high yields of biosensing [76]. Besides, *P. putida* KT2440 has adapted capabilities to thrive in various environments [77], therefore constructed gallic acid-inducible systems in this study were also transferred into *P. putida* KT2440. Moreover, this microorganism naturally possesses the necessary transportation mechanisms for gallic acid to enter the cell [3], ensuring that gallic acid will be present in the tetsed environment. Since *E. coli* Top 10 and *C. necator* H16 bacteria are sensitive to chloramphenicol, whereas *P. putida* KT2440 is capable to grow with this antibiotic, the chloramphenicol-resistance gene was replaced through PmeI and AscI restriction sites with the tetracycline-resistance gene that was PCR amplified using oligonucleotide primers IK003_AscI_345 and IK004_PmeI_344 from a plasmid pME6000. The ligated constructs were transformed into *E. coli* at first, the plasmid containing the tetracycline-resistance gene was purified and finally transformed into *P. putida*. The scheme of antibiotic resistance gene replacement is represented in Appendix 3. The grown cells on additionally added tetracycline indicated that the construct contains a tetracycline-resistance gene, therefore such bacteria were not additionally screened.

2.3.12. Preservation of suitable biosensors for long-term storage

The assembled biosensor constructs, which were verified by colony PCR and secondary restriction digestion, represent that all steps, including plasmid DNA purification, restriction digestion, plasmid assembly, and transformation, were performed correctly. Such biosensors must be kept for future biosensor output testing and investigation, using either microbank stocks or glycerol.

2.3.12.1. Preservation of biosensors to MicroBank

MicroBank freezing method is a convenient, ready-to-use, and simple technique. It is based on porous glass beads, which enable microorganisms to adhere readily onto the bead surface. A single bead is easily removed from the vial and inoculated directly into the suitable medium when required. The preservation was performed by mixing bacterial cells, which contain the suitable biosensor construct, with 2 mL of LB medium containing an appropriate antibiotic and incubating at a respective temperature in a shaking incubator at 225 rpm per night. Then, the liquid from the MicroBank vial was removed using an aseptic technique, and 500 µL of the grown suspension was transferred into a prepared vial, which was inverted 4-6 times to emulsify the suspension. The bacteria must bind to the beads, therefore a vial with the liquid was maintained for 1 min. Subsequently, the liquid was removed, and the vial was frozen at -80°C for long-term storage.

2.3.12.2. Preservation of biosensors with glycerol

This preservation method is simple, though requires initial preparation of glycerol and could be not as accurate as other methods. Glycerol stabilises bacteria and prevents damage to the cell membranes, keeping them alive, and such stock, therefore, can be stored at -80°C for many years. Bacterial cells containing the suitable biosensor construct were grown in LB medium containing proper antibiotic at

37°C or 30°C, depending on a bacterium that has been used as a host microorganism, in a shaking incubator at 225 rpm overnight. Then, 850 µL of the overnight culture was mixed with the 150 µL of 20% glycerol. The prepared mixture was frozen at -80°C for long-term storage.

2.3.13. Screening of inducible systems

In this project, the Tecan Infinite 200 Pro plate reader was used for the evaluation of fluorescence and absorbance of the developed biosensors. The bacterial cells possessing plasmid constructs were grown overnight in 2 mL of LB medium with appropriate antibiotics at 30°C and 200 rpm. After incubation, the suspensions were diluted with the same medium to OD_{600}=0.05 and incubated again to OD_{600}=0.1. Then, a 96-well black plate was filled with 142.5 µL of logarithmically growing cells, along with the respective inducer-specific phenolic acid added to the respective wells, of which volume was chosen to be 1/20th of the cell culture. Equally, respective wells were filled with the same components, except replacing the phenolic acids with sterile water or DMSO, to determine the value of negative control. Moreover, some wells were filled with respective media, in which the cells were grown, and sterile water or DMSO, respectively, to evaluate the background fluorescence and absorbance of the medium and calculate the absolute normalised absorbance and fluorescence values.

Inducers were dissolved in either sterile water or DMSO at the appropriate concentrations. Gallic acid was dissolved in sterile water to a concentration of 25 mM by adding it to the corresponding well and reaching a final concentration of 1.25 mM. While testing the sensitivity of biosensors, gallic acid was dissolved at 60°C [78] up to 100 mM. The 1:2 dilutions with sterile water (heated up to 60°C) were made until the concentration of 0.19 mM was reached. *p*-Coumaric acid and all other phenolic acids, which have been tested in this study, were dissolved in DMSO to a concentration of 100 mM and added to the wells by achieving a final concentration of 5 mM. When determining the sensitivity of a selected biosensor, *p*-coumaric acid was dissolved in DMSO up to a concentration of 200 mM, and 1:2 dilutions were performed until a concentration of 0.19 mM was reached.

The RFP fluorescence was measured using 585 nm as excitation wavelength and 620 nm as emission wavelength, with 9 nm and 20 nm bandwidths, respectively. Parallelly, the absorbance was measured using a wavelength of 600 nm with a 9 nm bandwidth, settle time was 5 ms. For both fluorescence and absorbance measurements, the number of flashes was set to 20, and a gain factor was set to 120%. The measurements were taken every 10 min for 22 h.

2.3.14. Mathematical modelling

To elucidate if a specific phenolic acid induces the biosensor by comparing the reporter output between induced and uninduced samples, the obtained values of absorbance and fluorescence were recalculated, as the medium can absorb and fluoresce itself. Since the plate reader provides an output of the absolute fluorescence and absorbance value of the culture medium (RFP_{medium} and OD_{medium}), the corrected values of bacterial absorbance were assessed by subtracting the absorbance of the medium from the absolute absorbance values (OD_{raw}). Equally, to evaluate the values of bacterial fluorescence, the normalised fluorescence values were calculated by subtracting absolute fluorescence values of the medium from the absolute fluorescence (RFP_{raw}). Then, the corrected absolute fluorescence value (RFP) was divided by corrected absorbance value (OD) [25]by receiving the value of absolute normalised fluorescence (RFP/OD) (1). The values of absolute normalised fluorescence after 6 or 12 hours of induction were received by calculating the average of four meanings before and after the corresponding hour. That means that each replica gave nine values,

which were accordingly recalculated by deriving the average of three replicas and receiving only one value at each time point.

$$\frac{RFP}{OD} = \frac{RFP_{raw} - RFP_{medium}}{OD_{raw} - OD_{medium}} \quad (1)$$

Furthermore, to elucidate the kinetic characteristics of inducible systems, specific parameters, such as the Hill coefficient, K_m value, and dynamic range (μ), were determined, contributing to distinguish one inducible gene expression system from another. Absolute normalised fluorescence values were expressed as a function of inducer concentration using software GraphPad Prism 9, as well as the specific binding with Hill slope was evaluated using the Hill function (2), indicating how the value of fluorescence depends on the concentration of inducer.

$$RFP(I) = b_{max} \times \frac{I^h}{K_m{}^h + I^h} + b_{min} \quad (2)$$

Parameter K_m is characterised as the concentration of the appropriate inducer that mediates half-maximal reporter output, parameter I match the concentration of an inducer, h matches the Hill coefficient, which describes the steepness of the curve and indicates the range of inducer concentration, where the fluorescent protein synthesis change [15]. Dynamic range is the maximum level of reporter output (b_{max}) relative to basal expression levels (b_{min}) (3). Besides, statistical analysis was performed with the same program – GraphPad Prism 9, by applying the unpaired *t*-test, identifying the statistically significant values. The p values were chosen either $p<0.05$, $p<0.01$, or $p<0.001$.

$$\mu = \frac{b_{max}}{b_{min}} \quad (3)$$

Furthermore, to estimate the real difference between induced and uninduced samples while testing the specificity of chosen biosensors, the relative induction was evaluated. Essentially, it is a ratio between the absolute normalised fluorescence value of any of the inducers, which are not specific to a biosensor ($FL_{compound}$), and an inducer, for which the biosensor was designed ($FL_{primary\ inducer}$) and expressed as a percentage (4). Formerly, the values of absolute normalised fluorescence of an uninduced sample ($FL_{uninduced}$) are subtracted from both values of specific and not specific inducers' values [15]. The calculation of relative induction was also estimated in testing the sensitivity of selected biosensors, only the absolute normalised fluorescence value of diverse concentrations of specific inducers was used instead of non-specific inducers' value ($FL_{compound}$), and the absolute normalised fluorescence value of the highest concentration was used instead of a specific inducer' ($FL_{primary\ inducer}$) fluorescence value. Overall, such an assessment of the results equated the fluorescence value of a specific inducer or concentration to 100%, while the uninduced sample was equated to 0%, enabling the determination of systems' parameters.

$$\text{Relative induction (\%)} = 100 \times \left(\frac{FL_{compound} - FL_{uninduced}}{FL_{primary\ inducer} - FL_{uninduced}}\right) \quad (4)$$

3. Results and discussion

3.1. Identification of gallic acid-inducible system

Gallic acid (3,4,5-trihydroxy benzoic acid) is a prevalent phenolic acid, which is usually found in the free state or in the form of esters and ethers. It distinguishes as being the most effective free radical scavenger, compared to the other phenolic acids [79], also demonstrating cytotoxic activity against cancer cells without harming normal cells [80]. This acid and its derivatives are therefore used as antioxidants in various industries, also for drug development, and as an antimicrobial agent [81].

The commercial value of gallic acid leads to the constant development of a process to produce it at the fermenter level in high yields. For instance, genetically engineered *E. coli* strains have been reported to produce gallic acid [82], [83]. However, in the process of development and identification of producing strains, some difficulties arise, such as detection of the synthesised product. There are plenty of already developed gallic acid-inducible amperometric [84] and electrochemical [85] biosensors, which can be applied for the improved production of this phenolic acid, whereas inducible gene expression system-based biosensors are still to be investigated. Although there are several already known bacterial species, which are involved in the gallic acid degradation pathways, such as *Pseudomonas putida*, *Azotobacter vinelandii*, *Chromohalobacter salexigens*, *Serratia proteamaculans*, *Klebsiella pneumoniae*, *Citrobacter rodentium*, *Burkholderia cepacian*, *Burkholderia multivorans*, *Ralstonia solanacearum*, and *Asticcacaulis excentricus* [3], which can be potentially applied in the identification of inducible systems and development of the whole-cell biosensors.

Pseudomonas genus bacteria are frequently used to investigate inducible systems, precisely *Pseudomonas putida* KT2440, as it can readily adapt to various environments. This microorganism contains various catabolic enzymes, including oxygenases, oxidoreductases, and dehydrogenases, thus being an attractive tool to investigate metabolic responses since it possibly possesses a variety of inducible gene expression systems, containing genetic regulators [86]. *P. putida* KT2440 is known for its ability to degrade various aromatic compounds, including gallic acid as a sole carbon and energy source [87]. Gene *galR* encodes a LysR-type TF with two DNA-binding domains and was demonstrated to act as a transcriptional activator in this bacterium. Its product GalR regulates the expression of enzymes in the catabolic pathway of gallic acid (Figure 3.1, A). These enzymes- and GalR- encoding genes are located at the *P. putida* KT2440 gene cluster, including *galA*, *galB*, *galB*, *galC*, *galD*, which are responsible for the gallic acid degradation pathway, along with *galT* and *galP* genes, regulating the transportation of gallic acid into the cell (Figure 3.1, B). Additionally, this cluster is the first set of genes reported to be responsible for the use of gallic acid as a carbon source [3].

Accordingly, to analyse this putative gallic acid-inducible system, it is predicted that an intergenic region (IGR) between *galR* and *galB* genes is a significant factor in transcribing the *galR* gene, as RNAP and GalR supposedly connects to the promoter site in this IGR, thus initiating the transcription of *galR* (Figure 3.1, C). Otherwise, there is a probability that the promoter site for transcription of *galR* is located elsewhere in the monitored operon. Overall, both promoter site probabilities were investigated in this study, along with the prevalence of the operon-constituting genes in other microorganisms.

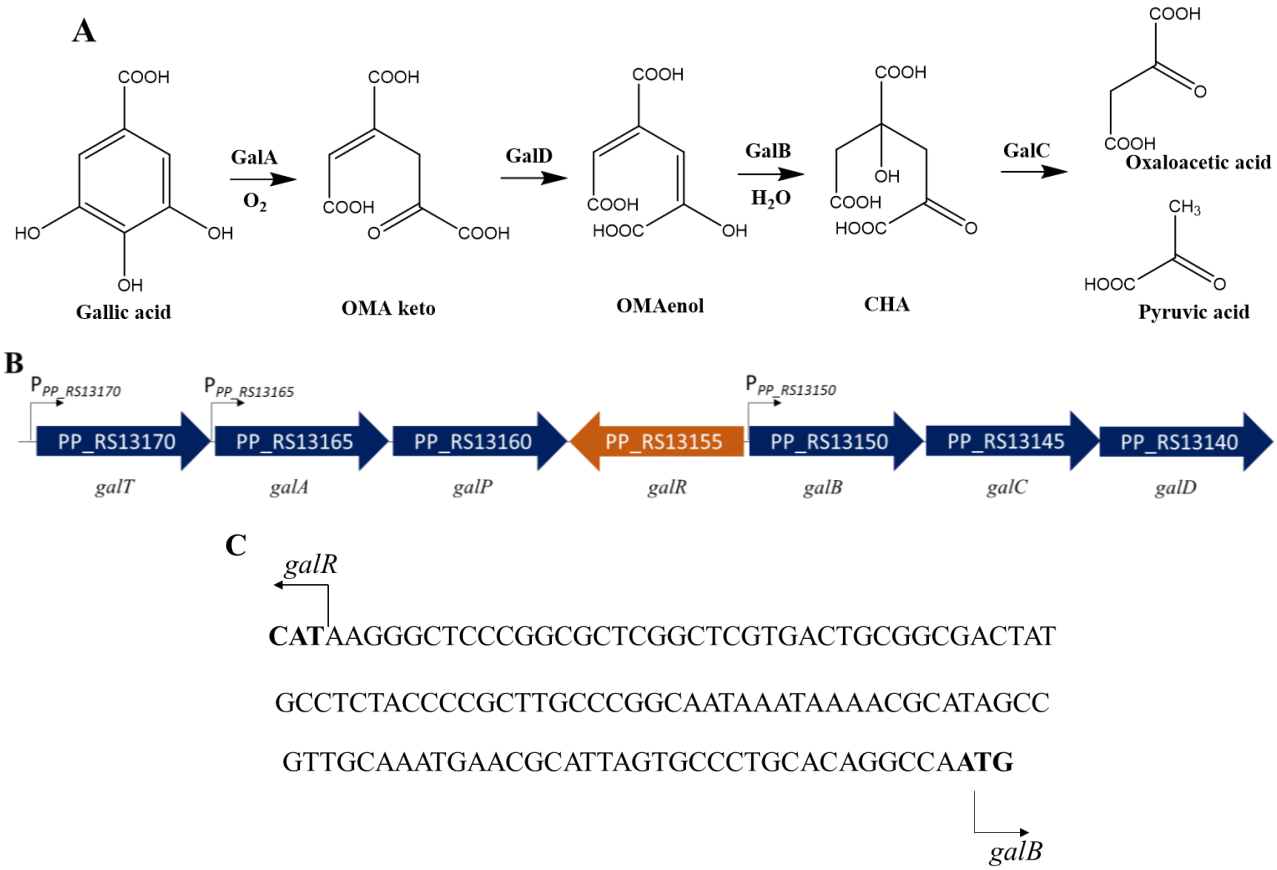


Fig. 3.1. A – gallic acid degradation pathway. GalA – gallic acid dioxygenase, GalD – OMA(4-oxalmesaconate)-keto enol tautomerase, GalB – 4-oxalmesaconate hydratase, GalC – CHA (4-carboxy-4-hydroxy-2-oxoadipate) aldolase, B – gene cluster of *P. putida* KT2440, representing the arrangement of gallic-acid inducible system-encoding genes, *galT* – MFS transporter (GalT) encoding gene, *galA* – gallate dioxygenase (GalA) encoding gene, *galP* – OprD family porin (GalP) encoding gene, *galR* – regulator GalR encoding gene, *galB* – GalB-encoding gene, *galC* – GalC-encoding gene, *galD* – GalD-encoding gene, C – the *P. putida* KT2440 *galR*/*galB* intergenic region, translational start sites are bold

3.1.1. Sequence alignment of *galR*, *galB*, and *galC* homologues

Since the gallic acid degradation pathway and its regulatory gene *galR* exists in *P. putida* KT2440, it was determined to elucidate the prevalence of this inducible system in other *Pseudomonas* species by performing BLAST analysis. The TF-encoding gene *galR* and other genes located nearby were aligned, providing information not only about the system prevalence but also about the evolutionary origin of analysed microorganisms, also assisting to investigate the properties of the putative inducible system. The genes *galR*, *galB*, and *galC* were aligned in various *Pseudomonas* species, containing a potential gallic acid-inducible system and sharing at least 90% protein sequence identity with *P. putida* KT2440 homologs (Table 3.1). The gene *galB* was aligned by two variations, as it potentially contains a longer promoter site (*galB* alternative). Moreover, supposing that the *galR* gene promoter site can be located upstream of the *galR* gene, the alignment was also performed with *glsB* (encoding glutaminase B), *galT* (encoding MFS transporter), *galA* (encoding gallate dioxygenase), and *galP* (encoding OprD family porin) genes in various *Pseudomonas* species, presuming that at least two proteins share 70% protein sequence identity (Table 3.2).

Table 3.1. List of *galR*, *galB*, and *galC* homologs from various *Pseudomonas* species sharing at least 90% protein sequence identity with *P. putida* KT2440 *galR*, *galB*, and *galC* homologs, respectively

Species	***galR* locus tag**	**Size (aa)**	**Coverage % (identity %)**	***galB* locus tag**	**Size (aa)**	**Coverage % (identity %)**	***galB* locus tag (alternative)**	**Size (aa)**	**Coverage % (identity %)**	***galC* locus tag**	**Size (aa)**	**Coverage % (identity %)**
P. putida KT2440	PP_RS_13155	398	100 (100)	PP_RS_13150	258	100 (100)	PP_RS13150	244	99 (100)	PP_RS13145	238	100 (100)
P. putida ND6	YSA_00537	397	99 (99.75)	YSA_00538	244	94 (99.18)	YSA_00538	244	99 (99.18)	YSA_00539	238	99 (99.16)
P. putida F1	Pput_3202	397	99 (99.50)	Pput_3203	244	94 (98.77)	Pput_3203	244	99 (98.77)	Pput_3204	238	99 (99.16)
P. putida DOT-T1E	T1E_3398	397	99 (99.50)	T1E_3399	258	94 (100)	T1E_3399	244	99 (100)	T1E_3400	238	99 (97.90)
P. putida H8234	L483_20165	397	99 (98.49)	L483_20170	258	94 (99.59)	L483_20170	244	99 (99.59)	L483_20175	238	99 (97.48)
P. putida W619	PputW619_2029	397	99 (96.98)	PputW619_2028	244	94 (99.59)	PputW619_2028	244	99 (99.59)	PputW619_2027	238	99 (95.80)
P. citronellolis	A9C11_06960	401	98 (77.10)	A9C11_06955	244	94 (95.90)	A9C11_06955	244	99 (95.90)	A9C11_06950	238	99 (89.50)
P. helmanticensis	EDF87_104223	411	99 (71.25)	EDF87_104222	244	94 (88.93)	EDF87_104222	244	99 (88.93)	EDF87_104221	238	98 (75.42)
P. laurylsulfatovorans	B0D71_00295	409	99 (71)	B0D71_00300	244	92 (90.38)	B0D71_00300	244	97 (90.38)	B0D71_00305	238	98 (75.42)

P. fragi	B6D87_15860	403	99 (70.68)	B6D87_15855	244	94 (87.30)	B6D87_15855	244	99 (87.30)	B6D87_15850	238	99 (73.95)
P. putida NCTC13186	NCTC13186_03432	398	100 (100)	NCTC13186_03433	259	100 (100)	NCTC13186_03433	245	100 (100)	NCTC13186_03434	239	100 (100)
Pseudomonas sp. KBS0802	FFH79_017900	398	100 (100)	FFH79_017905	259	100 (100)	FFH79_017905	245	94 (100)	FFH79_017910	239	100 (100)
Pseudomonas sp. XWY-1	PRJ_3384	398	100 (100)	PRJ_3385	259	100 (99.18)	PRJ_3385	245	100 (99.18)	PRJ_3386	239	100 (99.16)
P. putida B4	CHN49_16985	398	100 (99.50)	CHN49_16990	259	100 (99.59)	CHN49_16990	245	100 (99.59)	CHN49_16995	239	100 (98.32)
P. putida SJTE-1	A210_16225	398	100 (99.75)	A210_16230	259	100 (99.18)	A210_16230	245	100 (99.18)	A210_16235	239	100 (99.16)
P. putida KF715	KF715C_ch33490	398	100 (98.49)	KF715C_ch33500	259	100 (99.18)	KF715C_ch33500	245	100 (99.18)	KF715C_ch33510	239	100 (97.48)
P. putida JB	Q5O_15825	398	100 (100)	Q5O_15830	259	100 (99.59)	Q5O_15830	245	100 (99.59)	Q5O_15835	239	100 (99.58)
Pseudomonas sp. JY-Q	AA098_11260	398	100 (99.75)	AA098_11255	259	100 (99.59)	AA098_11255	245	100 (99.59)	AA098_11250	239	100 (97.90)
P. putida NBRC 14164	PP4_26440	401	98 (67.68)	PP4_26430	244	94 (88.93)	PP4_26430	244	99 (88.93)	PP4_26420	238	98 (73.73)
P. putida B1	CHR26_19440	398	100 (98.24)	CHR26_19435	259	100 (98.77)	CHR26_19435	245	100 (98.77)	CHR26_19430	239	100 (97.48)

Table 3.2. List of genes encoding glutaminase B, MFS transporter, gallate dioxygenase, and OprD family porin homologs from various *Pseudomonas* species, of which at least two proteins shared at least 70% protein sequence identity with *P. putida* KT2440 glutaminase B, MFS transporter, gallate dioxygenase, and OprD family porin-encoding homologs, respectively

Species	***glsB* locus tag**	**Size (aa)**	**Coverage % (identity %)**	***galT* locus tag**	**Size (aa)**	**Coverage % (identity %)**	***galA* locus tag**	**Size (aa)**	**Coverage % (identity %)**	***galP* locus tag**	**Size (aa)**	**Coverage % (identity %)**
P. putida KT2440	PP_RS13175	247	100 (100)	PP_RS13170	449	100 (100)	PP_2518	420	100 (100)	PP_2517	410	100 (100)
P. putida ND6	YSA_00525	302	100 (99.19)	YSA_00528	449	84 (96.32)	YSA_00530	420	99 (99.76)	YSA_00533	287	69 (97.21)
P. putida F1	Pput_3197	302	100 (99.60)	Pput_3199	449	84 (96.58)	Pput_3200	420	99 (99.76)	Pput_3201	415	99 (96.63)
P. putida DOT-T1E	T1E_3392	302	100 (99.60)	T1E_3394	449	84 (96.32)	T1E_3395	420	99 (99.76)	T1E_3397	415	99 (96.87)
P. putida H8234	L483_20150	248	46 (38.89)	L483_20155	449	84 (95.26)	L483_20160	420	99 (98.81)	L483_26540	418	98 (67.39)
P. putida BIRD-1	PPUBIRD1_3162	302	100 (100)	PPUBIRD1_3164	449	84 (96.84)	PPUBIRD1_3165	420	99 (99.52)	PPUBIRD1_3167	415	99 (97.83)
P. putida JB	Q5O_15805	248	82 (100)	Q5O_15810	450	93 (97.11)	Q5O_15815	421	100 (99.52)	Q5O_15820	411	100 (97.83)
P. putida NCTC13186	NCTC13186_03427	258	96 (100)	NCTC13186_03429	426	49 (42.86)	NCTC13186_03430	421	100 (100)	NCTC13186_03431	411	100 (100)
Pseudomonas sp. KBS0802	FFH79_017880	*	*	FFH79_017885	*	*	FFH79_017890	421	100 (100)	FFH79_017895	411	100 (100)
Pseudomonas sp. XWY-1	PRJ_3380	248	82 (100)	PRJ_3381	450	93 (96.58)	PRJ_3382	421	40 (38.46)	PRJ_3383	411	100 (97.35)

P. putida B4	CHN49_16965	248	82 (99.60)	CHN49_16970	450	93 (96.58)	CHN49_16975	421	100 (99.76)	CHN49_16980	411	100 (96.87)
P. putida SJTE-1	A210_16205	248	82 (99.19)	A210_16210	450	93 (96.32)	A210_16215	421	100 (99.76)	A210_16220	*	*
Pseudomonas sp. JY-Q	AA098_11280	248	82 (99.60)	AA098_11275	450	93 (96.32)	AA098_11270	421	100 (99.76)	AA098_11265	411	100 (97.11)
P. putida B1	CHR26_19470	248	82 (95.16)	CHR26_19450	450	91 (95)	CHR26_19445	421	100 (98.81)	*	*	*
P. putida KF715	KF715C_ch33420	248	82 (95.16)	KF715C_ch33470	450	91 (95.26)	KF715C_ch33480	421	100 (98.57)	*	*	*

*-gene sequence was not found on BLAST

The results indicate that the analysed *Pseudomonas* species are highly conservative and evolutionary related, implying that the putative gallic acid-inducible systems exist not only in *P. putida* KT2440 but also in other genetically related species. Notably, the identification of inducible systems in the genome of *P. putida* KT2440 was previously demonstrated for inducer 3-hydroxypropionic acid [86], where gene expression was shown to be regulated by TF encoded by the *hpdR* gene. Interestingly, the 3-hydroxypropionic acid-inducible system was indicated as outperforming the gene expression in *E. coli* and *C. necator*, compared to the inducible system identified in *C. necator* H16. Such results suggest that the gallic acid-inducible system indicated in this study will potentially be more functional and preferable compared to the inducible systems originating from other microorganisms. Therefore, subsequent investigation of conservative regions was set to be performed.

3.1.2. Sequence alignment of intergenic regions

IGR is a segment of DNA sequence between genes and historically has been called junk DNA since it does not encode any genes. However, such regions do contain functionally significant elements, such as promoters, which help RNAP to recognize the binding sites and initiate the transcription [73]. The sensory and transducer action of each biosensor can be controlled by directly designing promoter sequences [13], which is, therefore, a significant factor in the process of biosensor design. To mediate the expression of a target gene, a specific TF interacts with the DNA sequence, which comprises the inducible promoter, consisting of the TF binding site, -10 and -35 sites – RNAP recognition sequences, and a ribosome binding site (RBS) [25], all of which are responsible for the functionality of inducible gene expression system.

The alignment results of *galR*/*galB* IGRs of twenty *Pseudomonas* species are represented as a sequence similarity motif (Figure 3.2), corresponding to the nucleotide region between -1 and -105 relative to the *P. putida* KT2440 *galR* translational start site. It indicates that the part of IGR from 35 to 75 nucleotide is a potential promoter sequence, which is highly conservative in all analysed *Pseudomonas* species. Moreover, nucleotides from 61 to 66 potentially correspond to -10 site, while nucleotides from 35 to 40 match -35 site. The 76 nucleotide is a potential transcriptional start site. Moreover, a site from 12 to 35 is a potential RBS. All mentioned conserved elements indicate the RNAP binding site and imply that the expression of *galR* proceeds in other *Pseudomonas* species, suggesting that GalR possibly acts as a regulatory element in other microorganisms. However, the GalR binding site was not predicted in this IGR, as there are no visible palindromic sequences. Since GalR contains two DNA-binding domains [2], and it is an activator-type regulator, its binding site can be potentially located between RBS and promoter sites, overlapping the -35 site, as it is often observed for LysR-type activators [88]. Structures of all aligned *galR*/*galB* IGRs are represented in Appendix 4.

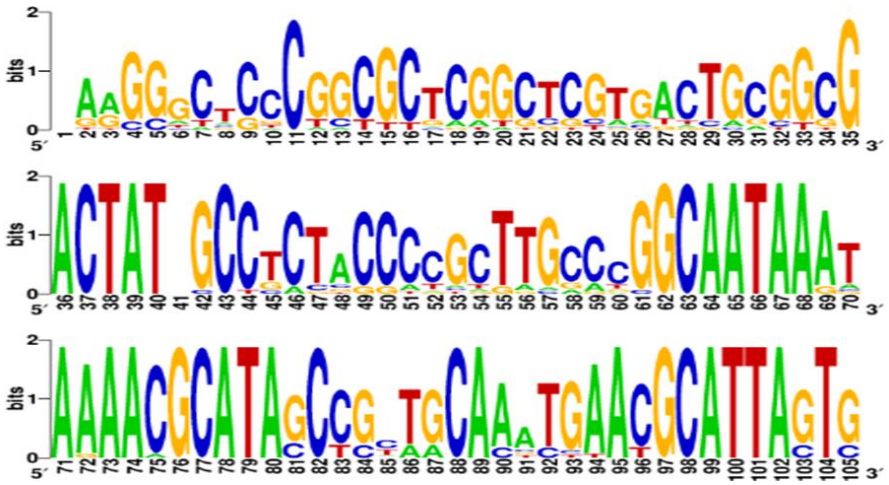


Fig. 3.2. A sequence similarity motif, which represents highly conserved nucleotides in *galR*/*galB* IGR of twenty analysed *Pseudomonas* species

Another presumable *galR* promoter site can potentially exist in the alternative IGR sequence, which is located upstream of the *galR* gene in the *glsB*/*galT* IGR. The results of such alignment are represented as a sequence similarity motif (Figure 3.3), corresponding to the nucleotide region between -1 and -180 relative to the *P. putida* KT2440 *galT* translational start site. However, analysis of fifteen *Pseudomonas* species indicated that all nucleotides from these IGRs are of high similarity level, therefore not demonstrating any specific binding sites or other relevant features for the functioning of *galR*, although re-approving a high evolutionary congeniality. Structures of all aligned *glsB*/*galT* IGRs are represented in Appendix 5.

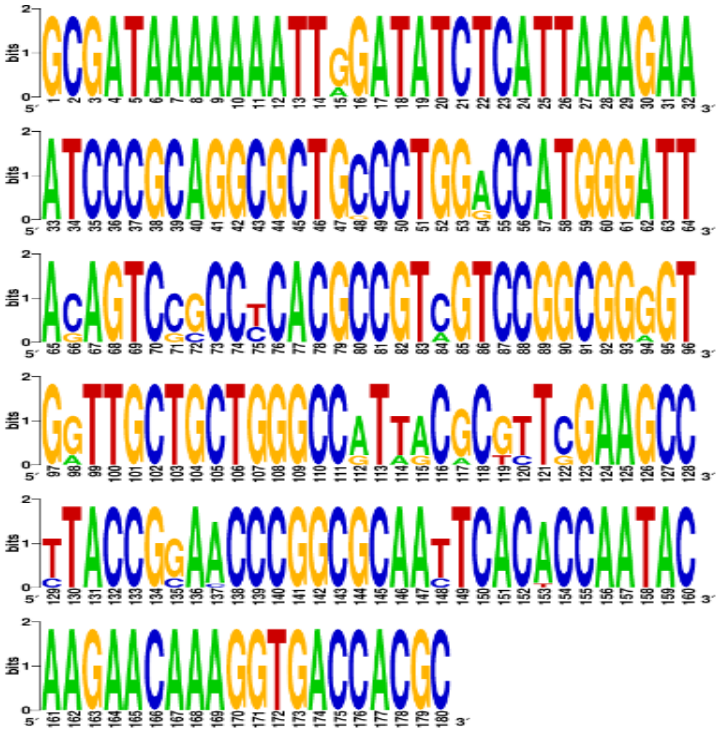


Fig. 3.3. A sequence similarity motif, which represents highly conserved nucleotides in *glsB*/*galT* IGR of fifteen analysed *Pseudomonas* species

In conclusion, such results explain how the inducible system is controlled, provide information to design genetically encoded biosensor constructs, and indicate potential microbial origins of gallic acid-inducible systems. They also reveal evolutionary characteristics, relative metabolic pathways in diverse microorganisms and aid in elucidating the gene arrangement in genomic DNA of all analysed *Pseudomonas* species.

3.1.3. Construction of gallic acid-inducible plasmid vectors

Putative gallic acid-inducible systems were predicted based on catabolic pathways of gallic acid in various microorganisms (Table 3.3), including *P. putida* KT2440, in addition to the elucidated conservative IGRs. The identified inducible systems are abbreviated depending on the origin of TF, the title of the applied TF, the direction of TF-gene according to the *RFP* direction (C-clockwise, A-anticlockwise), and the promoter used. Constructs harboring gallic acid-inducible systems were developed by amplifying specific inserts containing TF genes and promoters or promoters only, elucidating the orthogonality of the investigated regulators. Seven inducible gene expression systems were identified based on GalR, LigR, MarR, or DesR TFs. Structures of all designed and assembled biosensor constructs for the detection of gallic acid are represented in Appendix 6.

Table 3.3. Gallic acid-inducible plasmids designed and constructed in this study

Construct	Inducible system applied	Origin of TF[a]	Applied microorganism[b]
pEV001	*Pp*GalR-A/$P_{PP\_RS13150}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440/*C. necator* H16
pEV001A	*Pp*-A/$P_{PP\_RS13150}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440/*C. necator* H16
pEV002	*Pp*GalR-A/$P_{PP\_RS13150}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440/*C. necator* H16
pEV002A	*Pp*-A/$P_{PP\_RS13150}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440/*C. necator* H16
pEV005	*Pp*GalR-A/$P_{PP\_RS13170}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440/*C. necator* H16
pEV005A	*Pp*-A/$P_{PP\_RS13170}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440/*C. necator* H16
pEV013	*Pp*-C/$P_{PP\_RS13170}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440
pEV014	*Pp*GalR-C/$P_{PP\_RS13170}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440
pEV033	*Bm*GalR-A/$P_{NP\_RS00205}$	*B. multivorans* ATCC BAA-247	*E. coli* Top10/*C. necator* H16
pEV034	*Bm*-A/$P_{NP\_RS00205}$	*B. multivorans* ATCC BAA-247	*E. coli* Top10/*C. necator* H16
pEV042	*Sl*LigR-A/P_{ASE85_21430}	*Sphingobium sp*. Leaf26	*E. coli* Top10
pEV043	*Sl*-A/P_{ASE85_21430}	*Sphingobium sp*. Leaf26	*E. coli* Top10
pEV046	*Sl*-C/P_{ASE85_21505}	*Sphingobium sp*. Leaf26	*E. coli* Top10
pEV047	*Sl*MarR-C/ P_{ASE85_21505}	*Sphingobium sp*. Leaf26	*E. coli* Top10
pEV050	*Sl*DesR-A/P_{ASE85_17810}	*Sphingobium sp*. Leaf26	*E. coli* Top10/*C. necator* H16
pEV051	*Sl*-A/P_{ASE85_17810}	*Sphingobium sp*. Leaf26	*E. coli* Top10/*C. necator* H16

pEV058	*Pp*-A/$P_{PP\_RS13165}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440
pEV059	*Pp*GalR-A/$P_{PP\_RS13165}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440
pEV060	*Pp*GalR-C/$P_{PP\_RS13165}$	*P. putida* KT2440	*E. coli* Top10/*P. putida* KT2440

[a]TF (transcription factor) is a specific DNA-binding protein, which modifies RNAP allowing the transcription to begin.
[b]The applied microorganism is the bacteria, into which the specific inducible system was transferred, generating an inducible biosensor.

The constructs pEV001, pEV001A, pEV002, pEV002A were designed based on the *Pp*GalR-A/$P_{PP\_RS13150}$ inducible system, where GalR was tested for regulating the expression of *galB*. The constructs differed in the length of the promoter $P_{PP\_RS13150}$, constructs pEV002 and pEV002A possessing the longer version of it. Constructs pEV005, pEV005A, pEV013, and pEV014 were designed based on the *Pp*GalR-A/$P_{PP\_RS13170}$ inducible system, differing in the direction of TF-gene. Constructs pEV033 and pEV034 were designed based on *Bm*GalR-A/$P_{NP\_RS00205}$ inducible system from a *B. multivorans* ATCC BAA-247, where GalR potentially regulates the expression of the gallate dioxygenase-encoding gene. Constructs pEV042 and pEV043 were designed based on the *Sl*LigR-A/P_{ASE85_21430} inducible system, potentially regulating the expression of 4-carboxy-4-hydroxy-2-oxoadipate aldolase/oxaloacetate decarboxylase. Constructs pEV046 and pEV047 were developed based on inducible system *Sl*MarR-C/P_{ASE85_21505}, potentially regulating the expression of 4-hydroxybenzoate 3-monooxygenase, and constructs pEV050, pEV051 were designed based on inducible system *Sl*DesR-A/P_{ASE85_17810}, regulating the expression of carotenoid oxygenase gene. Consequently, constructs pEV058, pEV059, and pEV060 were designed based on the *Pp*GalR-C/$P_{PP\_RS13165}$ inducible system, where the regulatory effects of GalR were analysed on the expression of gallate dioxygenase. All designed and generated constructs were tested in *E. coli* primarily and afterward in selected diverse microorganisms, developing the biosensors, possibly applicable in diverse hosts.

The construction of plasmid vectors enabled the transfer of a selected inducible system into a bacterial host (Figure 3.4). At the beginning of construction, vector pBRC1 was digested with Reases AatII and NdeI and visualised with agarose gel electrophoresis (Figure 3.5, B), indicating whether it was digested suitably and enabling the purification from the gel as described in Materials and methods. Next, a target DNA was amplified by PCR from genomic DNA, using the respective pair of primers. The amplified fragments were digested with the same pair of Reases, provided that the plasmid is constructed using ligation reaction or the fragment was applied directly if the assembly was carried out by the HiFi reaction. Then, the ligation or HiFi reaction was implemented, and the transformation reaction afterward. The overnight colonies have been verified by colony PCR and secondary restriction digestion (Figure 3.5, C, D), using the same primers (for instance, EV001-PP_2515 and EV003-PP_2515B for construct pEV002) and Reases (AatII and NdeI), as were used in assembling the construct.

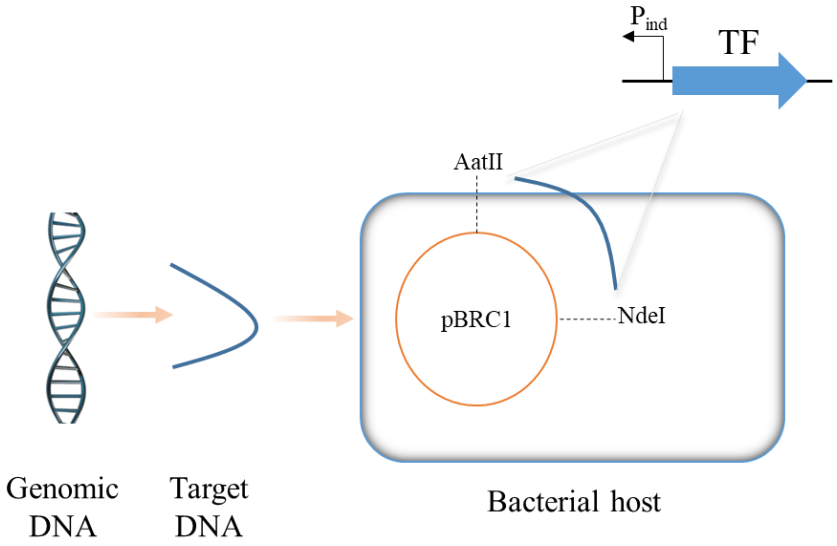


Fig. 3.4. The principal scheme of biosensor development

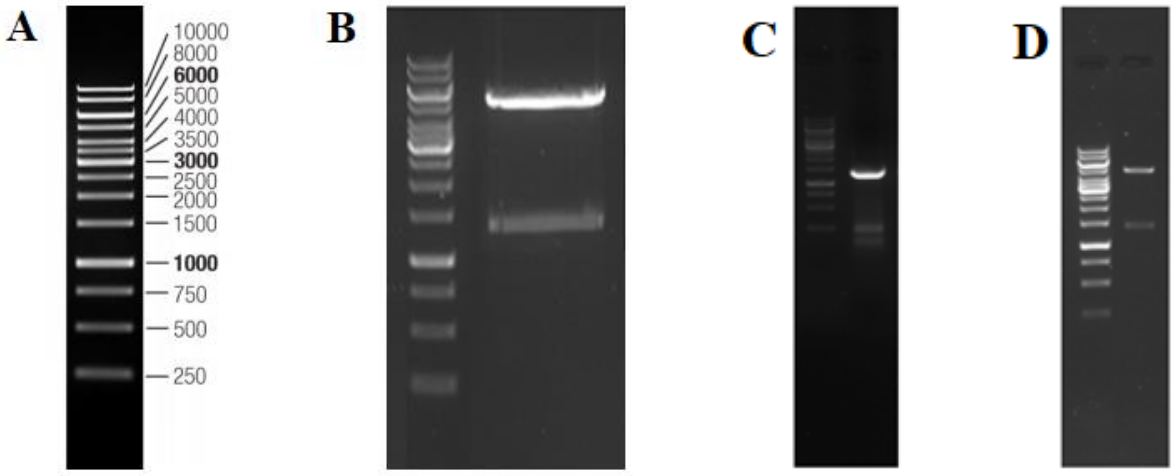


Fig. 3.5. A – the arrangement of Generuler 1 kb DNA Ladder (Thermo Fisher Scientific), which was used in this study, B – pBRC1 vector (5825 bp), which was digested with AatII and NdeI endonucleases, providing the residual unnecessary part (1263 bp) and the 4562 bp size digested fragment, which was used as a backbone plasmid in this study, C – amplified fragments of pEV002-based biosensor by colony PCR, representing 1394 bp long insert, D – digested pEV002 construct with AatII and NdeI, representing 1353 bp long insert and a residual part of the construct (4562 bp)

Generally, the functioning of gallic acid-inducible systems was monitored by inserting an RFP-encoding gene instead of structural genes, which are regulated by putative TF. For instance, the construct pEV002, based on the inducible system *Pp*GalR/$P_{PP\_RS13150}$, originating from *P. putida* KT2440, was developed by inserting the RFP-encoding gene instead of the native 4-oxalmesaconate hydratase-encoding gene *galB* (Figure 3.6). *RFP* gene was placed upstream of *galR*, encoding TF, which regulates the expression of 4-oxalmesaconate hydratase. When gallic acid is present in the cell, it binds to the GalR, creating a complex, which connects to the operator site adjacent to its regulated promoter $P_{PP\_RS13150}$. This

connection initiates the RNAP to join with the promoter site and start the transcription of *RFP*, resulting in the fluorescence signal being observed. To screen the regulator orthogonality, indicating if only GalR regulates the transcription of *galB*, an alternative construct pEV002A was developed, lacking the *galR* gene. Overall, the designed inducible systems were screened to indicate the functionality of the inducible system towards specific ligand – the process regulated by TFs.

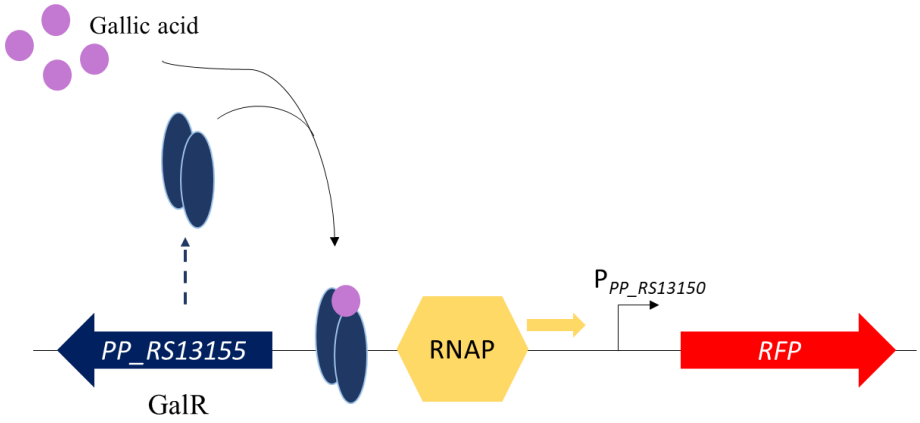


Fig. 3.6. Scheme of the functioning of inducible system *Pp*GalR/$P_{PP\_RS13150}$-based biosensor

3.1.4. Selection of the solvent for investigation of gallic acid-inducible biosensors

Study shows that the oxidation of gallic acid at neutral or alkaline pH complicates the investigation of its metabolism [3] thus it should be considered when testing gallic acid-inducible biosensors. In this study, to estimate the operating range and sensitivity of the functional biosensors, they were tested with various concentrations of gallic acid. Therefore, to elucidate an optimal and well-soluble solvent, which does not modify the results of absorbance and fluorescence, some different methodologies have been tested.

The most careful solvent for the investigation of biosensors would possibly be sterile water, however, the information in the literature indicates that gallic acid is sparingly soluble in this solvent [78], reaching only up to 20 mM concentration at room temperature. Such solvent as ethanol can be an improper choice, as it possesses an ability to readily evaporate, modifying the determined concentration of gallic acid in the well. Moreover, although DMSO being an excellent solvent, it could affect a stable cell line growth and viability [89]. The most suitable solvent should be selected carefully, therefore, gallic acid was dissolved in either DMSO, sterile water, ethanol, or ethanol at the highest concentration and performing 1:2 dilutions in sterile water. All samples were prepared up to 160 mM of gallic acid and diluted by 1:2 ratio until 0,078 mM was reached. The literature provides a lot of information on the ways to dissolve gallic acid in water by modifying the time, temperature, or even pH [78]. Therefore, in this project, a range of temperatures from 30°C to 60°C was applied to improve the gallic acid solubility in the water. Results showed that by increasing temperature to 60°C, up to 160 mmol of gallic acid can be dissolved per liter of water.

The results of bacterial absorbance while investigating the output of biosensor transferred into *P. putida*, based on pEV002 construct carrying *Pp*GalR/$P_{PP\_RS13150}$ inducible system, with gallic acid dissolved in DMSO, represent the bacterial growth-inhibiting effects (Figure 3.7), thus preventing the adequate

evaluation. However, it should be kept in mind that gallic acid enhances the growth of *P. putida*, which why the absorbance values could also be varied. Moreover, the absorption results with gallic acid dissolved in ethanol or ethanol with dilutions in water are highly distorted, possessing high error bars, especially while testing higher concentrations of gallic acid. Finally, the usage of sterile water as a solvent for gallic acid while testing biosensors demonstrates the least effects on absorbance values, suggesting that such a way of dissolution is the most suitable for stable bacterial growth.

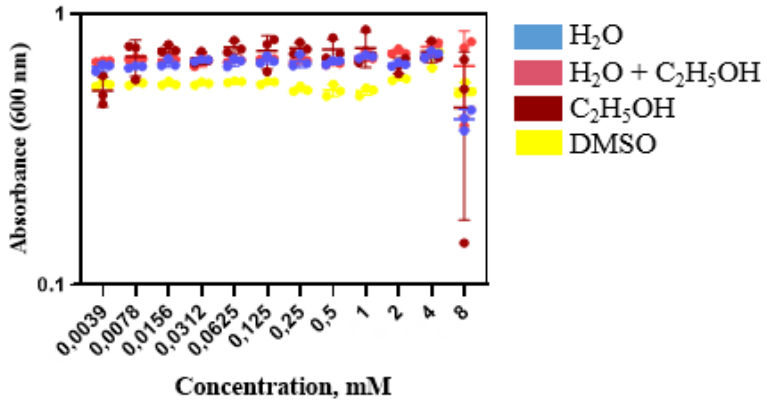


Fig. 3.7. The estimation of bacterial absorbance with various concentrations of gallic acid, dissolved in either sterile water, DMSO, ethanol, or ethanol at the highest concentration, and performing dilutions in sterile water. Results are indicated 6 hours after the inducer was added, error bars represent standard deviations of three biological replicates

Meanwhile, the analysis of fluorescence with the same biosensor indicates that the highest errors emerge while using ethanol, or ethanol and water as a diluent (Figure 3.8). The usage of DMSO as a solvent provides the fluorescence values at comparatively lower levels, which is indicated respectively with the reduced cell growth. On the contrary, sterile water as a solvent demonstrated the least effects on fluorescence, providing the lowest error bars and the highest fluorescence values overall.

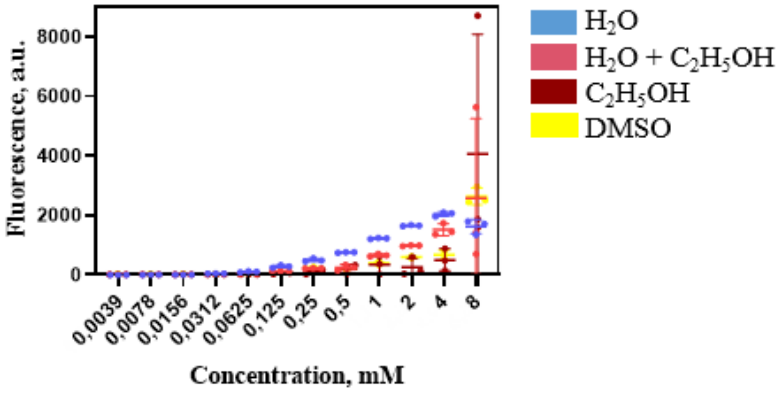


Fig. 3.8. The estimation of solvent impact on the output of the bacterial fluorescence, dissolving various concentrations of gallic acid in either sterile water, DMSO, ethanol, or ethanol at the highest concentration, and performing dilutions in sterile water. The results are indicated 6 hours after the inducer was added, error bars represent standard deviations of three biological replicates

In conclusion, such results indicate that the most suitable solvent for gallic acid is sterile water for the investigation of inducible biosensors, as it affected the results of fluorescence the least, maintaining the

fluorescence levels at the highest range and absorbance levels at a stable level. Therefore, water was selected as a solvent for further investigation of gallic acid-inducible biosensors.

3.1.5. Screening of gallic acid-inducible gene expression systems

Screening of gallic acid-inducible gene expression systems was performed by measuring absorbance and fluorescence levels of bacteria transformed with constructs harboring variously designed inducible systems. The absorbance level was measured to monitor the growth of bacteria and to assess the fluorescence by mathematical modelling. The parameters were obtained by using a plate reader, measuring the fluorescence and absorbance in real-time.

The results indicate that neither of the developed biosensors has functioned in *E. coli* and *C. necator*, although some induction effects with gallic acid were observed in *P. putida* (Figure 3.9). After 6 hours of adding 1.25 mM gallic acid, the pEV002-based biosensor demonstrated the highest dynamic range equal to 1171-fold compared to an uninduced sample. A biosensor based on pEV002A, which did not possess a TF-encoding gene, indicated induction of only up to 6.1-fold, compared to an uninduced sample. The pEV005-based biosensor demonstrated 57.5-fold induction after the addition of inducer, and a pEV005A-based biosensor indicated 112-fold induction, compared to an uninduced sample. The biosensor based on the pEV014 construct demonstrated 42-fold induction, and the pEV013-based biosensor showed 156.6-fold induction. The biosensor based on pEV058 construct inidcated 4-fold induction, pEV059-based biosensor demonstrated 6-fold induction, and pEV060-based biosensor indicated 8-fold induction, compared to uninduced samples, respectively. Subsequently, pEV001, pEV001A, pEV050, and pEV051-based biosensors did not demonstrate any induction effects with gallic acid. The screening of pEV033, pEV034, pEV042, pEV043, pEV046, and pEV047-based biosensors was not performed in *P. putida*, since the results of these biosensors using *E. coli* as a host did not represent any possible induction effects, only chaotically scattered values. All obtained statistically significant differences are based on screening inducible systems originating from *P. putida*.

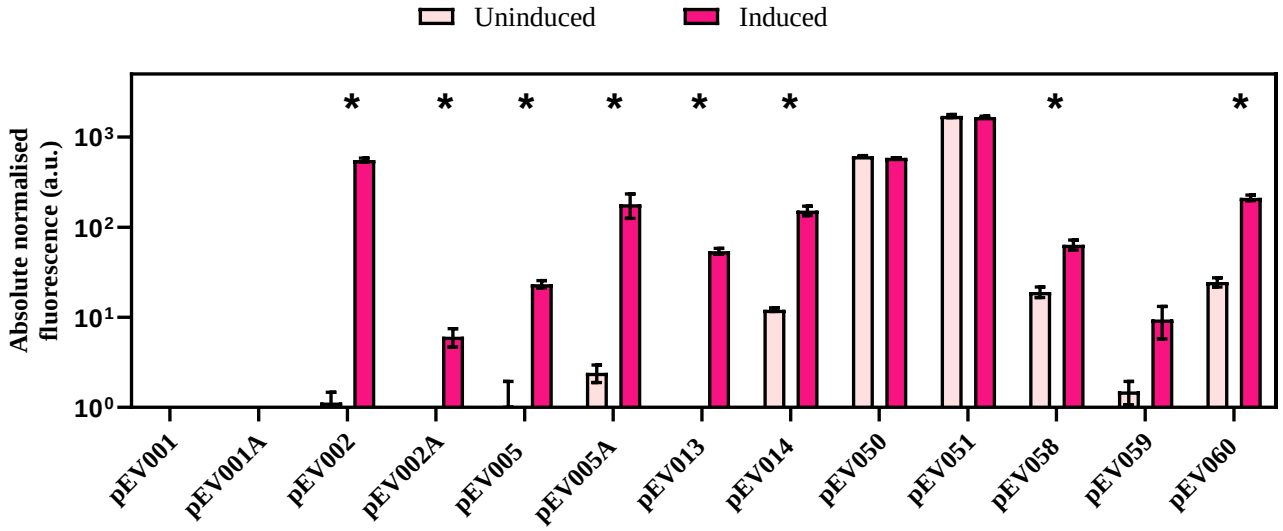


Fig. 3.9. The summary of the screened gallic acid-inducible biosensors using *P. putida* KT2440 as a host of application. The results represent the values 6 hours after 1.25 mM gallic acid was added, error bars represent

standard deviations of three biological replicates, asterisks indicate statistically significant values compared to an uninduced sample for $p<0.01$, unpaired two-tailed *t*-test

Since the host of application possesses the tested genes itself, it also naturally contains all transmission mechanisms of gallic acid into the cell [3], explaining the functional biosensors indicated in this study. However, such results can be confusing because of the same origin and the host microorganism in all biosensors. In any case, it is known that the incorporation of extraneous biosynthetic pathways into a host microorganism usually causes an imbalanced metabolic flux, leading to the toxicity caused by the accumulation of intermediates, which results in the inhibition of cell growth and decreased process performance. The lack of transport systems [90] in diverse metabolically unrelated hosts can also reduce the efficiency of the inducible system performance, wherefore functional biosensors did not demonstrate any induction in *E. coli* and *C. necator*. Moreover, the results of pEV005A and pEV013-based biosensors are controversial since the higher dynamic ranges were observed than in the biosensors containing TF-genes, based on pEV005 and pEV014, respectively. The results of pEV058, pEV059, and pEV060-based biosensors indicate that *galR* regulates the transcription of gallate dioxygenase but only at low levels. An induction level of a pEV002A-based biosensor is very low compared to the TF-encoding gene containing inducible system-based biosensor, indicating that precisely *galR* gene regulates the expression of 4-oxalmesaconate hydratase. In conclusion, the highest induction effects using gallic acid as an inducer is observed with a pEV002-based biosensor, which was therefore selected for further characterisation.

3.1.6. The specificity of gallic acid-inducible biosensor

Genetically encoded biosensors described in the literature are frequently lacking extensive characterisation, indicating only all-or-nothing induction effects, this way providing a potential to ascertain the overexpression of toxic genes but not demonstrating the information of the dosage of the inducers [11] and their specificity towards the inducer of interest. In this project, the specificity of the selected biosensor was tested to ensure accuracy, verify applicability, and perceive whether it detects only the phenolic acid of interest, as there might emerge some misunderstandings.

The testing was carried out using twenty phenolic acids with a concentration of 1.25 mM of each. It is visible that bacterial cells containing plasmid construct pEV002 grew logarithmically with all phenolic acids, indicating that a 1.25 mM concentration of tested phenolic acids does not harm bacteria and that the measurements can be adequately estimated (Figure 3.10, A). The results demonstrate that *P. putida* harboring *Pp*GalR/$P_{PP\_RS13150}$ inducible system-based construct pEV002 is highly specific to gallic acid, displaying 1171-fold induction 6 hours after the addition of the inducer and 1846-fold induction 12 hours after addition of the inducer (Figure 3.10, C, D). Moreover, neither of the other tested phenolic acids demonstrated a statistically significant difference compared to an uninduced sample.

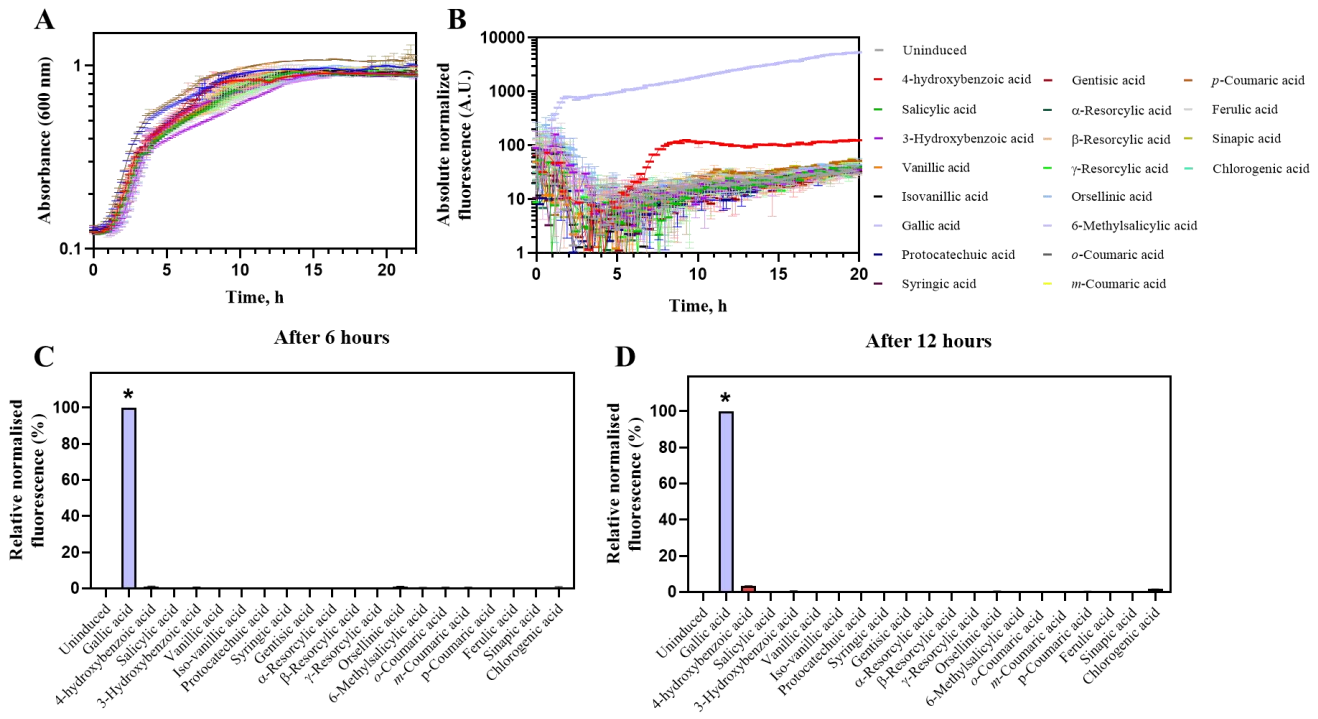


Fig. 3.10. The specificity of *P. putida* KT2440 harboring pEV002 construct. Twenty different phenolic acids of concentration 1.25 mM were tested. A – Absorbance of the cells at 600 nm, B – absolute normalised fluorescence of biosensor, using 585 nm as excitation wavelength and 620 nm as emission wavelength, C – relative normalised fluorescence of the biosensor after 6 hours, D – relative normalised fluorescence of the biosensor after 12 hours, error bars correspond to standard deviations of three biological replicates, asterisks represent statistically significant values for $p<0.001$, unpaired two-tailed *t*-test

Nevertheless, although pEV002-based biosensor appears to be highly specific, some unclear phenomenon is observed. The increase in fluorescence is observed from the 8th hour with 4-hydroxybenzoic acid (Figure 3.10, B), indicating the potential induction. This occurrence is observed possibly due to the structural similarity of this phenolic acid to gallic acid, although no statistically significant difference was observed after 6 and 12 hours of ligand addition (Figure 3.10, C, D). On the other hand, such induction can be observed due to the conversion of the tested inducer to another compound activating the inducible system. Overall, such results demonstrate a very high specificity towards relevant inducer, indicating that the developed gallic acid-inducible biosensor can be applied for orthogonal identification of this phenolic acid in the medium, even though structurally similar compounds are present.

3.1.7. The sensitivity of gallic acid-inducible biosensor

Biosensor based on pEV002 construct, harboring *Pp*GalR/$P_{PP\_RS13150}$ inducible system, was furtherly characterised elucidating its sensitivity and indicating an operating range, achieved by testing it with various concentrations of gallic acid. As well as the intrinsic strength of promoter and sensor-DNA equilibrium can vary the kinetics of induction [11], the concentration of the inducing molecule relates to the protein expression, and a degree of the cell growth affects expression kinetics. Therefore, gallic acid-inducible biosensors were tested with the maximum concentration of 2.5 mM of gallic acid, as the higher

concentrations caused the values to scatter. Moreover, since *P. putida* uses gallic acid as a sole carbon and energy source [3], higher concentrations increased absorbance values, resulting in modification of fluorescence intensity.

The results demonstrate that the pEV002-based biosensor is induced by gallic acid in a concentration-dependent manner (Figure 3.11, B). The operating range of the inducible biosensor was estimated approximately from 0.078 mM to 2.5 mM (Figure 3.11, C) 6 hours and 0.039 to 2.5 mM (Figure 3.11, D) 12 hours after various concentrations of gallic acid was added. Besides, the K_m value, corresponding to the concentration of gallic acid, mediating half-maximal reporter output, was estimated to be equal to 1.362 after 6 hours and 0.689 after 12 hours of the gallic acid addition. The Hill coefficient is calculated to be equal to 1.357 and 1.388, 6 and 12 hours after the gallic acid was added, respectively.

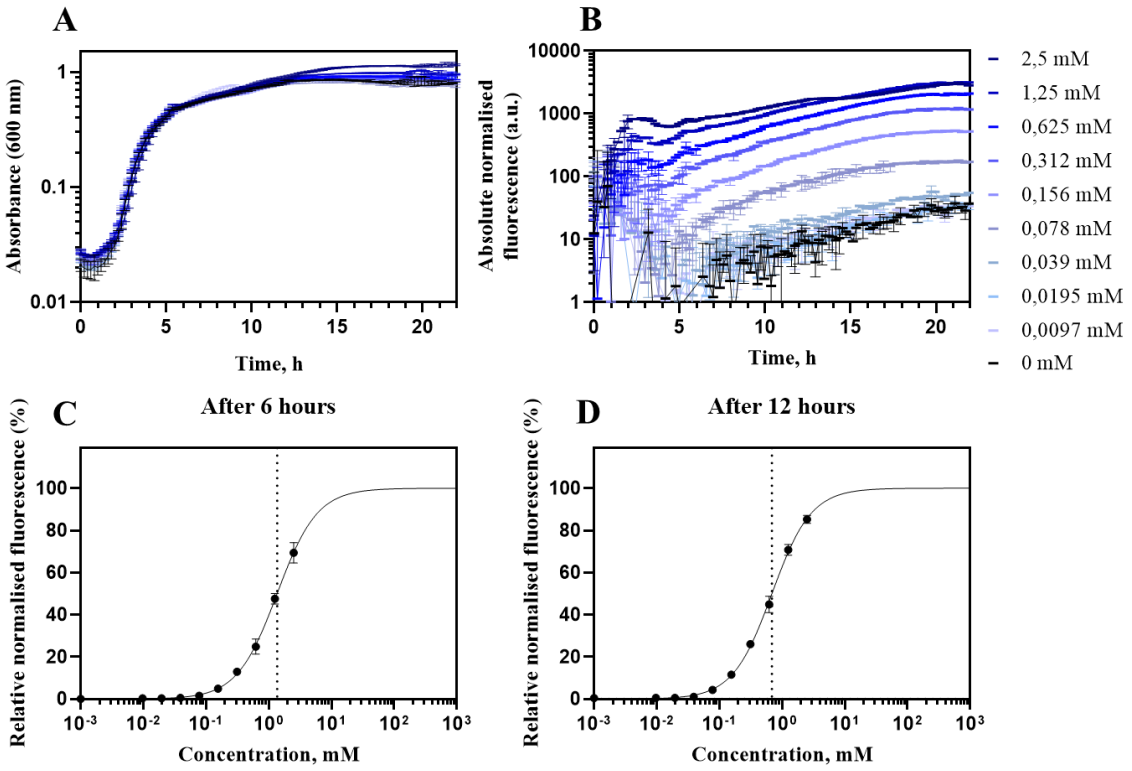


Fig. 3.11. The sensitivity to gallic acid of *P. putida* KT2440 harboring the pEV002 construct. A – absorbance of the bacterial cells at 600 nm, B – absolute normalised fluorescence of the biosensor, using 585 nm as excitation wavelength and 620 nm as emission wavelength, C - relative normalised fluorescence of the biosensor, measurements were taken 6 hours after the inducer was added, D – relative normalised fluorescence 12 hours after the addition of the inducer, K_m is indicated by a dotted line, error bars correspond to standard deviations of three biological replicates

Such results indicate that a pEV002-based biosensor can be used for the quantitative evaluation of gallic acid. The K_m values demonstrate that the system was almost saturated, and Hill coefficient values are comparatively low, which indicates that gene expression is tuneable over a quite wide range of ligand concentrations. However, the developed inducible system can be furtherly modified and improved, especially by applying it in diverse host microorganisms.

3.2. Identification of *p*-coumaric acid-inducible system

p-Coumaric acid, also known as 4-hydroxycinnamic acid, is the most abundant isoform in nature out of the three isomers. In addition to the benefits and already applied biosensors that were described above, this phenolic acid is still lacking a universal and operative TF-based biosensor [41]. Metabolic pathways of *p*-coumaric acid catabolism are thoroughly analysed in the *Bacillus* genus, where a structural gene *padC*, encoding phenolic acid decarboxylase, is induced in the presence of *p*-coumaric acid. PadC catalyses the conversion of *p*-coumaric acid to *p*-vinylphenol (Figure 3.12, A) since some phenolic acids, including *p*-coumaric acid, are toxic for gram-positive bacteria [12].

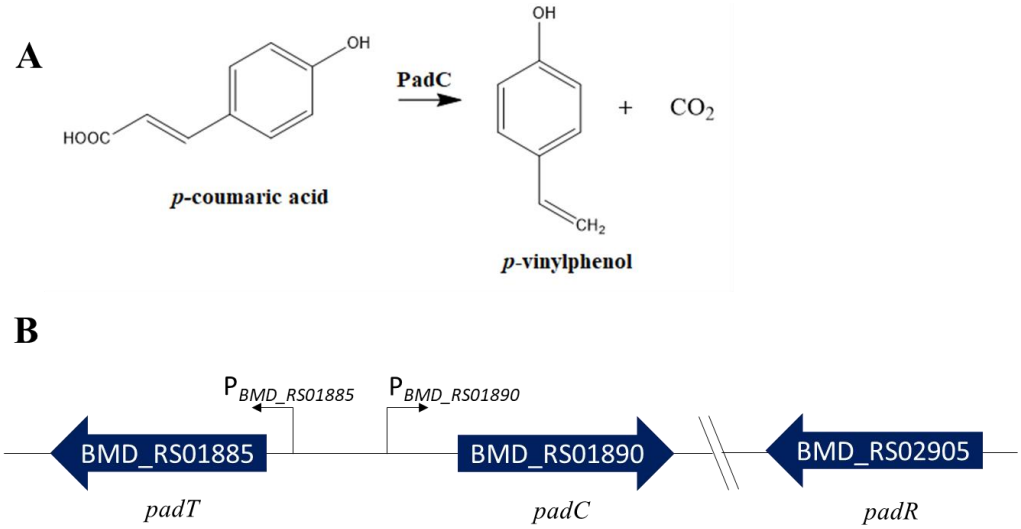


Fig. 3.12. A – the conversion of *p*-coumaric acid to *p*-vinylphenol, catalysed by phenolic acid decarbocylase (PadC), B – identified operon in *Bacillus megaterium* DSM 319, where product of the *padR* regulates the expression of *padC*, encoding phenolic acid decarboxylase, *padT* – MFS transporter-encoding gene

As already mentioned above, the biosensor based on *padR-padC* operon was developed by Siedler et al. [33], based on *Bacillus subtilis* genomic DNA. Therefore in this project, similar inducible systems were predicted in other *Bacillus* species, including *Bacillus pumilus* ATCC 7061, *Bacillus subtilis subsp. subtilis* str. 168, and *Bacillus megaterium* DSM 319 (Figure 3.12, B). Additionally, a *p*-coumaric acid-inducible system was predicted in the genome of *Sphingobium* sp. Leaf26, potentially regulating *p*-hydroxycinnamoyl CoA hydratase/lyase-encoding gene, which product catalyses hydration and two-carbon cleavage of *p*-hydroxycinnamic acids. The predicted inducible systems were applied for designing and developing potential *p*-coumaric acid-inducible whole-cell biosensors.

3.2.1. Construcion of *p*-coumaric acid-inducible plasmid vectors

All constructs for detection of *p*-coumaric acid were designed and developed in the same way as it was done with gallic-acid inducible plasmid vectors. The usage of Reases and specific primers for PCR generated target DNA, which was assembled into a pBRC1 vector. *p*-Coumaric acid-inducible constructs were designed based on 4 identified inducible systems by replacing relevant structural genes with *RFP*-gene and inserting TF-gene either clockwise or anticlockwise. 14 *p*-coumaric acid-inducible plasmid vectors were generated (Table 3.4), structures of which are represented in Appendix 7. The constructs

pIK001, pIK002 were designed by using *Bp*PadR/$P_{BPUM\_RS03685}$ inducible system from *Bacillus pumilus* to indicate whether the *padR* regulates the expression of a *padC*. The constructs pIK003, pIK004, pIK005, pIK006, and pIK007 were designed by modifying RBS strength using the inducible system *Bs*PadR/P_{BSU_34400} from *Bacillus subtilis* for testing if the *padR* regulates the expression of a *padC*, additionally inserting *yveFG* encoding genes between *RFP* and *padR*. The constructs pIK008, pIK009, pIK010, pIK011, and pIK012 were designed by predicting inducible system *Bm*PadR/$P_{BMD\_RS01890,\ BMD\_RS01885}$ in *Bacillus megaterium*, verifying if the *padR* regulates the expression of a *padC* and testing two variants of the promoter regions. Additionally, the constructs pEV044 and pEV045 were designed based on the inducible system *Sl*MarR/P_{ASE85_21485} from *Sphingobium* sp. Leaf26, indicating if MarR regulates the expression of the gene encoding *p*-hydroxycinnamoyl CoA hydratase/lyase.

Table 3.4. *p*-Coumaric acid-inducible plasmids designed and constructed in this study

Construct	Inducible system applied	Origin of TF[a]	Applied microorganism[b]
pIK001	*Bp*-A/$P_{BPUM\_RS03685}$	*B. pumilus* ATCC 7061	*E. coli* Top10
pIK002	*Bp*PadR-A/$P_{BPUM\_RS03685}$	*B. pumilus* ATCC 7061	*E. coli* Top10/*C. necator* H16
pIK003	*Bs*-A/P_{BSU_34400}	*B. subtilis subsp. subtilis* str. 168	*E. coli* Top10
pIK004	*Bs*PadR-A/P_{BSU_34400}	*B. subtilis subsp. subtilis* str. 168	*E. coli* Top10
pIK005	*Bs*PadR-C/P_{BSU_34400}	*B. subtilis subsp. subtilis* str. 168	*E. coli* Top10
pIK006	*Bs*PadR-A/P_{BSU_34400}	*B. subtilis subsp. subtilis* str. 168	*E. coli* Top10
pIK007	*Bs*PadR-C/P_{BSU_34400}	*B. subtilis subsp. subtilis* str. 168	*E. coli* Top10/*C. necator* H16
pIK008	*Bm*-C/$P_{BMD\_RS01890,\ BMD\_RS01885}$	*B. megaterium* DSM 319	*E. coli* Top10
pIK009	*Bm*PadR-A/$P_{BMD\_RS01890,\ BMD\_RS01885}$	*B. megaterium* DSM 319	*E. coli* Top10/*C. necator* H16
pIK010	*Bm*PadR-C/$P_{BMD\_RS01890,\ BMD\_RS01885}$	*B. megaterium* DSM 319	*E. coli* Top10/*C. necator* H16
pIK011	*Bm*PadR-A/$P_{BMD\_RS01890,\ BMD\_RS01885}$	*B. megaterium* DSM 319	*E. coli* Top10/*C. necator* H16
pIK012	*Bm*PadR-C/$P_{BMD\_RS01890,\ BMD\_RS01885}$	*B. megaterium* DSM 319	*E. coli* Top10/*C. necator* H16
pEV044	*Sl*MarR-A/P_{ASE85_21485}	*Sphingobium sp*. Leaf26	*E. coli* Top10
pEV045	*Sl*-A/P_{ASE85_21485}	*Sphingobium sp*. Leaf26	*E. coli* Top10

[a]TF (transcription factor) is a specific DNA-binding protein, which modifies RNAP allowing the transcription to begin.

[b]The applied microorganism is the bacteria, into which the specific inducible system was transferred, generating an inducible biosensor.

After assembly of the constructs, they were transformed into *E. coli*, and the screening was performed, including colony PCR and secondary restriction digestion, using the same primers and Reases (AatII and NdeI) as were used in the construction (Figure 3.13, A, B). Afterward, the functioning plasmid vectors were transformed into *C. necator*, indicating if the inducible system functions in a wider range of host microorganisms.

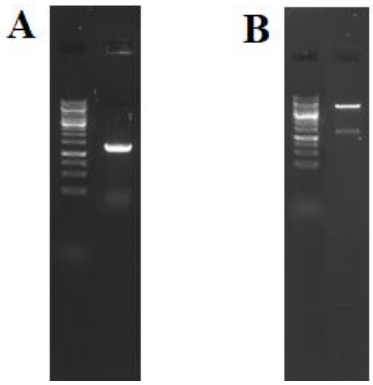


Fig. 3.13. Results of agarose gel electrophoresis of screened fragments of the construct pIK010. A - amplified fragment by colony PCR, representing 1292 bp long insert, B – digested construct with AatII and NdeI, representing 1257 bp long insert and a residual part of the plasmid (4556 bp)

3.2.2. Screening of *p*-coumaric acid-inducible gene expression systems

The screening results of *p*-coumaric acid-inducible systems using *E. coli* as a host indicate that 6 hours after 5 mM of *p*-coumaric acid was added, biosensors based on constructs pIK002, pIK007, pIK009, pIK010, pIK011, and pIK012, were induced statistically significantly (Figure 3.14). The highest dynamic range was observed screening pIK010-based biosensor, reaching up to 6-fold induction compared to an uninduced sample, therefore this biosensor was selected for further characterisation. Moreover, only the pIK010-based biosensor demonstrated statistically significant ($p<0.05$) 5-fold induction in *C. necator*, indicating wider applicability of the inducible system in diverse hosts. Besides, it can be emphasised that the pIK008-based biosensor, based on *Bm*-C/$P_{BMD\_RS01890,\ BMD\_RS01885}$ system, which corresponds to the promoter-only variant of the pIK010-based biosensor, did not demonstrate any induction, suggesting that *padR* acts orthogonally in this biosensor. The fluorescence values of the pIK008-based construct reached more than 20000 A.U. after 6 hours of the ligand addition, whereas the fluorescence values of the pIK010-based biosensor indicated the values only up to 4500 A.U. Such variation implies that the functioning biosensor can be furtherly developed to reach the values of the TF-lacking biosensor.

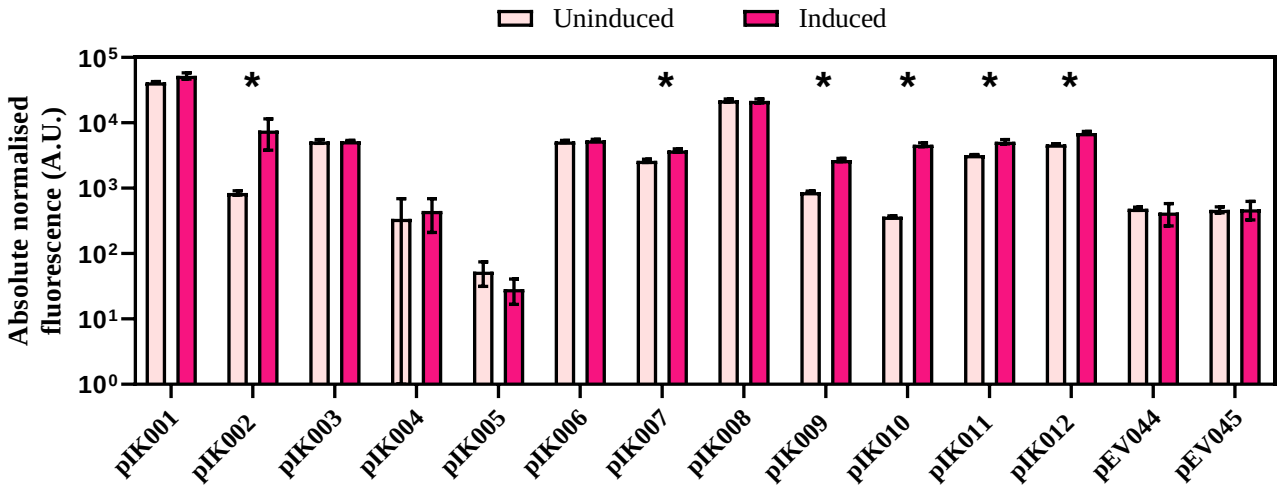


Fig. 3.14. The results of *p*-coumaric acid-inducible biosensors, tested using *E. coli* Top10 as a host of application, 6 hours after 5 mM *p*-coumaric acid was added, error bars represent standard deviations of three

biological replicates, asterisks indicate statistically significant difference compared to an uninduced sample ($p<0.05$), unpaired two-tailed *t*-test

Overall, the results indicate that the majority of the developed *p*-coumaric acid-inducible biosensors are functioning in *E. coli*. This suggests that these inducible systems can be used to orthogonally regulate the gene expression of biosynthetic pathways or other genetic circuits in the wider host range, suggesting wide applicability. Although demonstrating only moderate activation, these biosensors should be subsequently optimised to reach higher induction levels. The optimisation can be achieved by tuning the concentration of TF, modifying the plasmid copy number, changing the binding affinity of the inducer-TF complex or TF-operator, modifying the number and position of operators, changing the reporter itself, or the strength of promoter-operator and RBS [91]. The latter type of optimisation was demonstrated by Siedler et al., as the constructed *p*-coumaric acid-inducible biosensor from *B. subtilis subsp. subtilis* 168, tested in *E. coli* Top10, demonstrated 130-fold induction with 2 mM *p*-coumaric acid, reducing the expression of the *padR* in *E. coli* cells and enabling the system to be induced [33], while the non-modified version of biosensor was not induced. Although the results in this project did not display the higher induction with the modified RBS versions, other non-modified constructs can be optimised and furtherly developed by applying diverse modification methods.

3.2.3. The specificity of *p*-coumaric acid-inducible biosensor

The specificity of a *p*-coumaric acid-inducible biosensor was assessed to indicate whether the biosensor will detect only *p*-coumaric acid and not respond to structurally similar compounds. *E. coli* cells harboring the pIK010 construct gave the highest dynamic range from all developed *p*-coumaric acid-inducible biosensors, and it was selected to be furtherly characterised.

The results indicate that the developed biosensor responds not only to *p*-coumaric but also to ferulic acid (Figure 3.15, B). Such a detection possibly occurs due to the similar structure of these phenolic acids since they are both hydroxycinnamic acids, only ferulic acid containing an additional methoxy group. Moreover, bacteria demonstrated low fluorescence values while testing gallic acid as an inducer, even lower than the uninduced values, possibly because of its oxidation in neutral pH, this way darkening the tested wells. Such variation caused high absorbance values (Figure 3.15, A), which modified the expression kinetics [11], along with altering the fluorescence values. However, the dynamic range 6 hours after the addition of 5 mM *p*-coumaric acid is equal to 6-fold, while it is equal to 4-fold using 5 mM ferulic acid as an inducer (Figure 3.15, C). The dynamic range 12 hours after the addition of 5 mM *p*-coumaric acid is 6-fold, and for 5 mM ferulic acid, it is equal to 5-fold (Figure 3.15, D). The statistically significant difference between induced and uninduced samples 6 hours after the addition of salicylic and β-resorcylic acid was also observed, although it disappeared in the later hours.

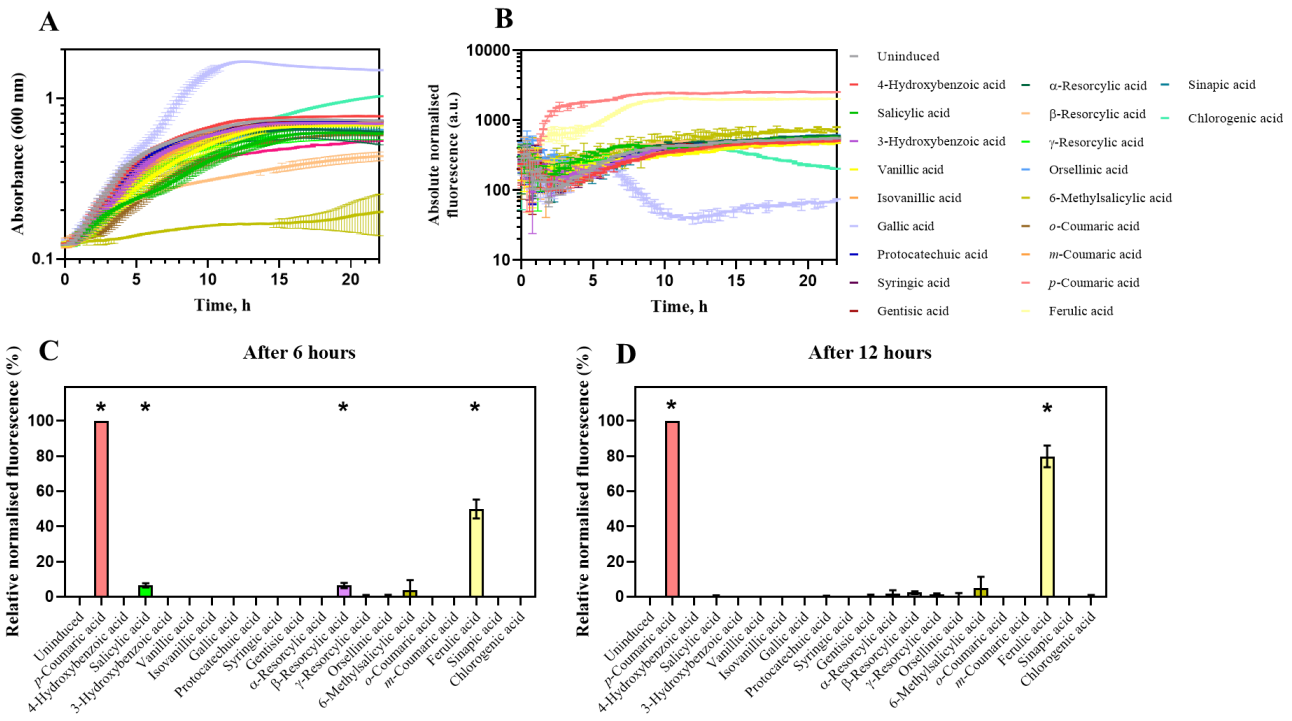


Fig. 3.15. The specificity of *E. coli* Top10 harboring the pIK010 biosensor construct. Twenty different phenolic acids of concentration 5 mM were tested. A – absorbance of the cells at 600 nm, B – absolute normalised fluorescence of the biosensor, using 585 nm as excitation wavelength and 620 nm as emission wavelength, C – relative normalised fluorescence of the biosensor after 6 hours, D – relative normalised fluorescence of the biosensor after 12 hours, error bars correspond standard deviations of three biological replicates and asterisks represent statistically significant values for $p<0.01$

The bacterial growth during the testing indicates various bacterial growth-promoting or -inhibiting effects. It is observed that 5 mM 6-methylsalicylic acid highly reduced the growth of bacteria, which is observed possibly due to the toxicity of this phenolic acid. It was previously reported that *p*-coumaric acid also contains growth-inhibiting properties [92]. Nevertheless, the biosensor containing the pIK010 construct did not demonstrate such inhibitory effects. Interestingly, 5 mM *p*-coumaric acid inhibited the growth of other tested biosensors, such as *E. coli* containing pIK001 and pIK002 constructs (data not shown). Moreover, the screening of gallic acid-inducible biosensors with 1,25 mM *p*-coumaric acid also did not demonstrate growth inhibition. Such results indicate the *p*-coumaric acid concentration- and construct-dependent inhibition of bacterial growth. The information in the literature suggests that after the addition of *p*-coumaric acid to *Shigella dysenteriae*, the antibacterial activity of hydrophobic antibiotics improves significantly, *p*-coumaric acid increases the permeability of the outer membrane, and the K^+ efflux from the cell increases extremely [92]. Overall, such growth-inhibition mechanisms could also occur in *E. coli* Top10, explaining the results of this study, although these effects should be furtherly investigated.

In conclusion, even though the developed pIK010-based biosensor tested in *E. coli* is not exceptionally specific to *p*-coumaric acid, but it indicates the applicability in diverse hosts. Moreover, *p*-coumaric acid does not cause the growth inhibition of bacteria in this biosensor, and therefore it can be tested with

higher concentrations of this phenolic acid. Besides, such an induction manner can be used to indicate if either *p*-coumaric or ferulic acid is present in the environment or if the producing strains synthesise any of these compounds.

3.2.4. The sensitivity of *p*-coumaric acid-inducible biosensor

The pIK010-based biosensor was also evaluated for its response to the range of different concentrations of *p*-coumaric acid. The results demonstrate the induction effects in a concentration-dependent manner, in the range from 0.0097 to 5 mM of *p*-coumaric acid (Figure 3.16, B). The K_m values of the tested biosensor are equal to 2.545 after 6 hours of the addition of the inducer (Figure 3.16, C) and 2.685 after 12 hours (Figure 3.16, D). Moreover, the Hill coefficient was obtained equal to 1.564 and 1.772 after 6 and 12 hours of the addition of *p*-coumaric acid, respectively.

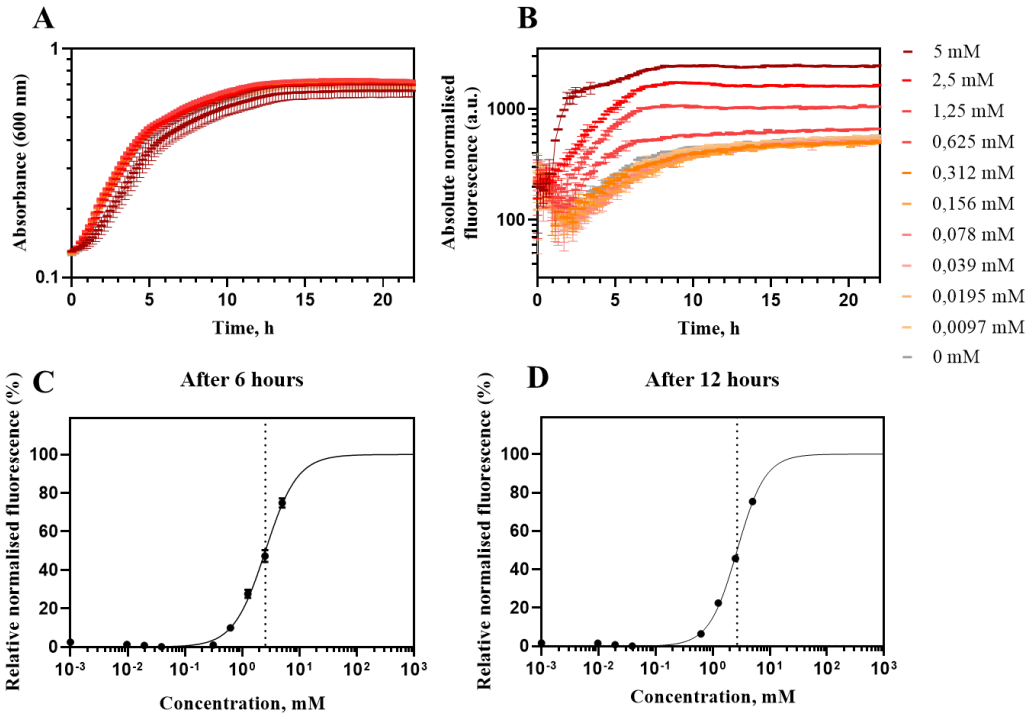


Fig. 3.16. The sensitivity to *p*-coumaric acid of *E. coli* Top10 harboring the pIK010 construct. A – absorbance of the bacterial cells at 600 nm, B – absolute normalised fluorescence of the biosensor, using 585 nm as excitation wavelength and 620 nm as emission wavelength, C - relative normalised fluorescence of biosensor in response to different concentrations of *p*-coumaric acid, measurements were taken 6 hours after the inducer was added, D – relative normalised fluorescence 12 hours after the addition of the inducer, K_m is indicated by a dotted line, error bars correspond standard deviations of three biological replicates

The results indicate that a pIK010-based biosensor can be used for a quantitative evaluation of *p*-coumaric acid. The K_m values demonstrate that the system was fully saturated, and values of the Hill coefficient show that the gene expression is tuneable over a quite wide range of inducer concentrations. Indeed, the system should be furtherly improved either to increase the response range or to improve its dynamic range overall. In conclusion, the *E. coli* pIK010-based biosensor demonstrates the concentration-dependent response, implying the possibility to apply it in either metabolic engineering or other analysis areas.

4. Recommendations part

The developed biosensors can be furtherly developed and applied to produce higher yields of specific phenolic acids indicating the best performing producer strains, to detect phenolic acids in a natural environment, to improve the performance of enzymes in significant metabolic pathways, or to test the quality of phenolic-acids rich products. To implement this, the developed functional biosensor constructs could be produced industrially. According to the information published in the literature and the investigation performed in this project, the technological scheme was designed (Figure 4.1), providing the method for generating phenolic acid-inducible plasmids, based on an example of gallic acid-inducible plasmid pEV002, containing a tetracycline-resistance gene.

The overall process to produce the plasmid DNA (pDNA) at high amounts can be divided into two sections – upstream and downstream processes, the latter consisting of four reaction steps, such as fermentation, primary recovery, intermediate purification, and final purification [93]. All devices used in a technological scheme are listed in Table 4.1. The production of pDNA starts with the upstream process, which is the initial preparation of the cell culture containing the biosensor construct. Then, the downstream process is initiated by preparing a medium in a broth storage tank (S-1), which is based on dissolving solid media in pure water and then autoclaving (A-1). The antibiotic fraction is prepared by dissolving tetracycline in pure water (S-2), which is later filter-sterilised (F-2) and added to the prepared medium up to 10 µg/mL. The biosensor construct-containing *E. coli* Top10 cells are collected from a *Petri* dish and cultivated at aerobic conditions in a medium (FL-1), which was formerly prepared and supplemented with filter-sterilised tetracycline. The cultivated cells are supplied to a bioreactor (B-1), constantly supplemented with sterile air and prepared medium, containing tetracycline at aerobic conditions at 37°C, cultivating for 16-20 hours [94]. The emitted air is constantly filter-sterilised and discarded into the environment.

After cultivation, the concentration of cells reached is approximately 7 g/L of dry cell weight. The broth is then stored in a tank (T-1) while the bioreactor is washed to prepare it for the subsequent incubation. The cells are furtherly centrifuged (C-1) at 14300 g. The supernatant is discarded, and the cell pellet is resuspended with a resuspension solution in a blending tank (S-3), this way concentrating the cells 10-15 times compared to the primary fermentation broth [93]. Then, lysis solution is added, containing 200 mM NaOH and 1% (weight per volume) sodium dodecyl sulfate as a detergent, causing cell membranes solubilisation and protein denaturation. Effective mixing is a key in this step, guaranteeing an equal distribution of pH, which protects pDNA from irreversible denaturation. pDNA and other components, including proteins, RNA, and genomic DNA, are released from the cell, therefore a neutralisation solution containing 3 M potassium acetate (pH 5.5) is added to the mixture, which causes the contaminants to precipitate.

The precipitate is then filtered (N-1), maintaining the lysate at 4°C to avoid degradation. The clarified lysate is then precipitated (P-1) with isopropyl alcohol by obtaining pDNA, which is filtered repeatedly (N-2). To remove salt ions, isopropyl alcohol is added, and the waste is discarded. The precipitate is then transferred to a solubilisation tank (SL-1) and redissolved in a solubilisation solution made of 10 mM Tris-HCl buffer (pH 8.0). Protein, endotoxin, and RNA impurities are precipitated in this step by adding

the dissolved solid ammonium sulfate up to 2.5 M, which increases the pDNA purity by 6.5-fold [93]. Then, the final filtering step is performed (F-3) by removing the precipitate. The filtered mixture is transferred to a tank (S-4) to prepare it for the final purification step.

The final purification step is performed in a hydrophobic-interaction chromatography column (CH-1). This chromatography technique is based on a separation of molecules according to their hydrophobicity, therefore all impurities, including the remaining proteins, single-stranded nucleic acids, and genomic DNA, are adsorbed by the matrix. In pDNA molecules, hydrophobic bases are hidden in the helix and only minimally interact with the matrix. Therefore, pDNA flows through the column and can be furtherly purified. The phenyl-Sepharose HIC gel matrix should be used, applying the negative mode. Tris-HCl buffer supplemented with ammonium sulfate is used as a low ionic strength buffer for isocratic elution in a step mode. The column is consequently washed with 1 M NaOH. Then, to remove the ammonium sulfate, the almost purified fraction of pDNA is diafiltrated (D-1, D-2) and consequently sterilised using microfiltration (ST-1) with a 0.2 µM pore sizes [93], to certainly assure that pDNA is sterile and lacks additional microorganisms or other contaminants. After final purification, the pDNA product obtained is about 1 mg/mL [94], which is stored (T-1) and subsequently packaged and filled into vials.

Table 4.1. The devices used in the technological scheme

Device abbreviation	Device
A-1	autoclave
A-2	air compressor
B-1	bioreactor
C-1	centrifuge
D-1; D-2	diafiltration equipment
FL-1	250 mL flask
F-1	air filter
F-2	antibiotic filter
F-3	product filter
CH-1	chromatography column
N-1, N-2	Nutche filter
P-1	precipitator
S-1	broth storage tank
S-2	antibiotic dissolution tank
SL-1	tank for solubilisation
S-3	blending tank
S-4	Intermediate storage tank
ST-1	microfiltration filter
T-1	storage tank

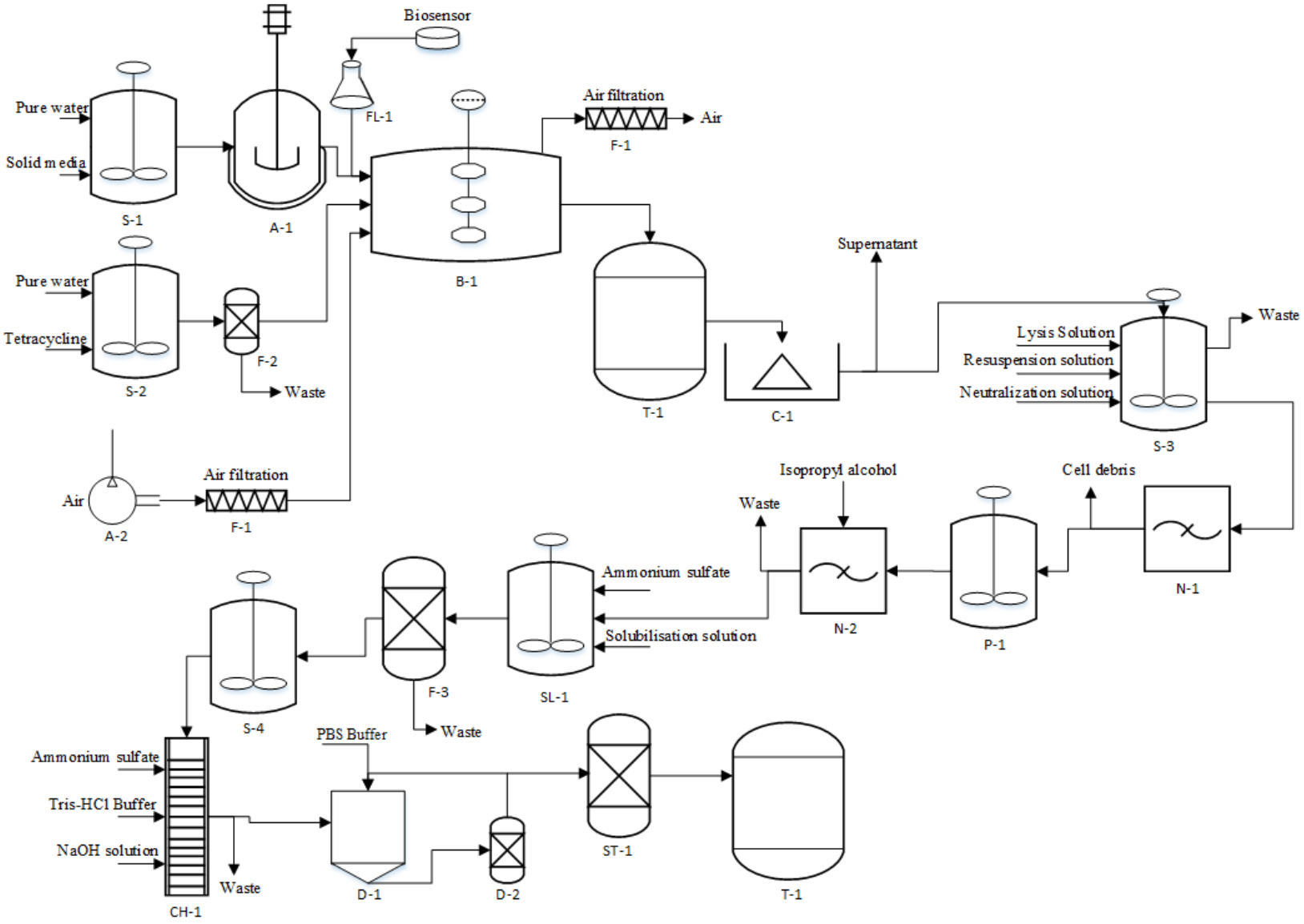


Fig. 4.1. The principal technological scheme for plasmid industrial production

Conclusions

1. Gene operons responsible for gallic and *p*-coumaric acid catabolism were identified. Subsequently, using information relevant to functional genetic elements such as promoter sequences, operators, and genes encoding transcriptional regulators, seven gallic acid and four *p*-coumaric acid-inducible gene expression systems were predicted.

2. Application of bioinformatics tools enabled identification of the potential -35 and -10 promoter sequence motifs in the intergenic region of gallic acid-inducible gene expression system in *Pseudomonas putida* KT2440. However, no sufficiently conservative sequence region corresponding to the transcription factor GalR binding site was found.

3. 19 and 14 plasmid constructs with gallic acid and *p*-coumaric acid-inducible gene expression systems were assembled, respectively.

4. Amongst potential systems identified in this study, the *Pp*GalR/$P_{PP\_RS13150}$, *Pp*$P_{PP\_RS13150}$, *Pp*GalR/$P_{PP\_RS13170}$, *Pp*$P_{PP\_RS13170}$, *Pp*-A/$P_{PP\_RS13165}$, *Pp*GalR-C/$P_{PP\_RS13165}$ inducible systems were found suitable for the detection of gallic acid, and the *Bp*PadR/$P_{BPUM\_RS03685}$, *Bs*PadR-C/P_{BSU_34400}, *Bm*PadR/$P_{BMD\_RS02900,\ BMD\_RS01890}$ inducible systems were able to respond to the *p*-coumaric acid.

5. Using *P. putida* KT2440 and *E. coli* Top10 strains as hosts, the biosensors based on the *Pp*GalR/$P_{PP\_RS13150}$ and *Bm*PadR-C/$P_{BMD\_RS02900,BMD\_RS01890}$ systems were developed, respectively. Application of screen with 20 different phenolic acids revealed that the former was specific to gallic acid in a concentration-dependent manner, whereas the latter was responding not only to *p*-coumaric acid in a concentration-dependent manner but also can be applied for ferulic acid detection.

List of references

1. BIALECKA-FLORJANCZYK, E., A. FABISZEWSKA and B. ZIENIUK. Phenolic acids derivatives - biotechnological methods of synthesis and bioactivity. Current Pharmaceutical Biotechnology. 2018, 19(14), 1098-1113. ISSN 1873-4316.
2. VALANCIENE, E., I. JONUSKIENE, M. SYRPAS, E. AUGUSTINIENE, P. MATULIS, A. SIMONAVICIUS and N. MALYS. Advances and prospects of phenolic acids production, biorefinery and analysis. *Biomolecules*. 2020, 10(6), 874. ISSN 2218-273X.
3. NOGALES, J., A. CANALES, J. JIMENEZ-BARBERO, B. SERRA, J. M. PINGARRON, J. L. GARCIA and EDUARDO DIAZ. Unravelling the gallic acid degradation pathway in bacteria: the gal cluster from *Pseudomonas putida*. *Molecular Microbiology*. 2011, 79(2), 359-374. ISSN 1365-2958.
4. STALIKAS, C. D. Extraction, separation and detection methods for phenolic acids and flavonoids. *Separation Science*. 2007, 30(18), 3268-3295. ISSN 1615-9314.
5. DIETRICH, J. A., A. E. MCKEE and J. D. KEASLING. High-throughput metabolic engineering: advances in small-molecule screening and selection. *Annual Review of Biochemistry*. 2010, 79, 563-590. ISSN 1545-4509.
6. FERNANDES, F. H. A. and H. R. N. SALGADO. Gallic acid: review of the methods of determination and quantification. *Critical Reviews in Analytical Chemistry*. 2016, 46(3), 257-265. ISSN 1547-6510.
7. GARDNER, T. S., C. R. CANTOR and J. J. COLLINS. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*. 2000, 403(6767), 339-342. ISSN 1476-4687.
8. BASU, S., R. MEHREJA, S. THIBERGE, M. CHEN and R. WEISS. Spatiotemporal control of gene expression with pulse-generating networks. *Proceedings of the National Academy of Sciences*. 2004, 101(17), 6355-6360. ISSN 1091-6490.
9. PURNICK, P. E. M. and R. WEISS. The second wave of synthetic biology: from modules to systems. *Nature*. 2009, 10, 410-422. ISSN 1473-4687.
10. BERVOETS, I. and D. CHARLIER. Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology. *FEMS Microbiology Reviews.* 2019, 43(3), 304-339. ISSN 1574-6976.
11. ROGERS, J. K., C. D. GUZMAN, N. D. TAYLOR, S. RAMAN, K. ANDERSON and G. M. CHURCH. Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Research*. 2015, 43(15), 7648-7660. ISSN 1362-4962.
12. TRAN, N. P., J. GURY, V. DARTOIS and T. K. C. NGUYEN. Phenolic acid-mediated regulation of the padC gene, encoding the phenolic acid decarboxylase of *Bacillus subtilis*. *Journal of Bacteriology*. 2008, 190(9), 3213-3224. ISSN 1098-5530.
13. KHALIL, A. S. and J. J. COLLINS. Synthetic biology: applications come of age. *Nature Reviews Genetics*. 2010, 11, 367-379. ISSN 1471-0064.
14. KIM, N. M., R. W. SINNOT and N. R. SANDOVAL. Transcription factor-based biosensors and inducible systems in non-model bacteria: current progress and future directions. *Current Opinion in Biotechnology*. 2019, 64, 39-46. ISSN 0958-1669.

15. HANKO, E. K. R., A. C. PAIVA, M. JONCZYK, M. ABBOTT, N. P. MINTON and N. MALYS. A genome-wide approach for identification and characterisation of metabolite-inducible systems. *Nature Communications*. 2020, 11, 1213. ISSN 2041-1723.
16. GHASEMZADEH, A. and N. GHASEMZADEH. Flavonoids and phenolic acids: role and biochemical activity in plants and human. *Medicinal Plants Research*. 2011, 5(31), 6697-6703. ISSN 1996-0875.
17. BALASUNDRAM, N., K. SUNDRAM and S. SAMMAN. Phenolic compounds in plants and agri-industrial by-products: antioxidant activity, occurence, and potential uses. *Food Chemistry*. 2006, 99(1), 191-203. ISSN 0308-8146.
18. MIKLASINSKA-MAJDANIK, M., M. KĘPA, R. D. WOJTYCZKA, D. IDZIK and T. J. WĄSIK. Phenolic compounds diminish antibiotic resistance of *Staphylococcus aureus* clinical strains. *Environmental Research and Public Health*. 2018, 15(10), 2321. ISSN 1660-4601.
19. DAI, J. and R. J. MUMPER. Plant phenolics: extraction, analysis and their antioxidant and anticancer properties. *Molecules*. 2010, 15(10), 7313-7352. ISSN 1420-3049.
20. SANCHEZ, N. F. S., R. SALAS-CORONADO, B. HERNANDEZ-CARLOS and C. VILLANUEVA. *Shikimic acid pathway in biosynthesis of phenolic compounds*. Oaxaca: IntechOpen, 2018. ISBN 9781789840339.
21. HELENO, S. A., A. MARTINS, M. J. R. P. QUEIROZ and I. C. F. R. FERREIRA. Bioactivity of phenolic acids: metabolites versus parent compounds: a review. *Food Chemistry*. 2015, 173, 501-513. ISSN 0308-8146.
22. BAE, I. K., H. M. HAM, M. H. JEONG, D. H. KIM and H. J. KIM. Simultaneous determination of 15 phenolic compounds and caffeine in teas and mate using RP-HPLC/UV detection: method development and optimization of extraction process. *Food Chemistry*. 2015, 172, 469-475. ISSN 0308-8146.
23. SAMPAIO, C. R. P., F. HAMERSKI and R. HOFFMANN RIBANI. Antioxidant phytochemicals of *Byrsonima ligustrifolia* throughout fruit developmental stages. *Functional Foods*. 2015, 18, 400-410. ISSN 1756-4646.
24. KUMAR, N. and N. GOEL. Phenolic acids: natural versatile molecules with promising therapeutic applications. *Biotechnology Reports*. 2019, 24. ISSN 2215017X.
25. HANKO, E. K. R., N. P. MINTON and N. MALYS. Design, cloning and characterization of transcription factor-based inducible gene expression systems. *Methods in Enzymology*. 2019, 621, 153-169. ISSN 0076-6879.
26. FERNANDEZ-LOPEZ, R., R. RUIZ, F. D. L. CRUZ and G. MONCALIAN. Transcription factor-based biosensors enlightened by the analyte. *Frontiers in Microbiology*. 2015, 6, 648. ISSN 1664-302X.
27. FAGAN, R. P. and N. F. FAIRWEATHER. *Clostridium difficile* has two parallel and essential Sec secretion systems. *Biological Chemistry*. 2011, 286(31), 27483-27493. ISSN 1431-6730.
28. PEREZ-RUEDA, E., R. HERNANDEZ-GUERRERO, M. A. MARTINEZ-NUNEZ, D. ARMENTA-MEDINA, I. SANCHEZ and J. A. IBARRA. Abundance, diversity and domain architecture variability in prokaryotic DNA-binding transcription factors. *Plos One*. 2018, 13(4). ISSN 1932-6203.

29. MADDOCKS, S. E. and P. C. F. OYSTON. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology Society*. 2008, 154, 3609-3623. ISSN 1465-2080.
30. BINDER, S., G. SCHENDZIELORZ, N. STABLER, K. KRUMBACH, K. HOFFMANN, M. BOTT and L. EGGELING. A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biology*. 2012, 13, R40. ISSN 1465-6906.
31. CHEN, P. T., J-F. SHAW, Y-P. CHAO, T-H. D. HO and S-M. YU. Construction of chromosomally located T7 expression system for production of heterologous secreted proteins in *Bacillus subtilis*. *Agricultural and Food Chemistry*. 2010, 58(9), 5392-5399. ISSN 1520-5118.
32. GUI, Q., T. LAWSON, S. SHAN, L. YAN and Y. LIU. The application of whole cell-based biosensors for use in environmental analysis and in medical diagnostics. *Sensors*. 2017, 17(7), 1623. ISSN 1424-8220.
33. SIEDLER, S., N. K. KHATRI, A. ZSOHAR, I. KJAERBOLLING, M. VOGT, P. HAMMAR, C. F. NIELSEN, J. MARIENHAGEN, M. O. A. SOMMER and H. N. JOENSSON. Development of bacterial biosensor for rapid screening of yeast *p*-coumaric acid production. *ACS Synthetic Biology*. 2017, 6(10), 1860-1869. ISSN 2161-5063.
34. SERGANOVA, I. and R. G. BLASBERG. Molecular imaging with reporter genes: has its promise been delivered? *Nuclear Medicine*. 2019, 60(12), 1665-1681. ISSN 0161-5505.
35. BEREPIKI, A., R. KENT, L. F. M. MACHADO and N. DIXON. Development of high-performance whole cell biosensors aided by statistical modeling. *ACS Synthetic Biology*. 2020, 9(3), 576-589. ISSN 2161-5063.
36. MATEJCZYK, M., R. SWISLOCKA, M. KALINOWSKA, G. WIDERSKP, W. LEWANDOWSK, A. JABLONSKA-TRYPUO and S. J. ROSOCHACKI. *In vitro* evaluation of biological activity of cinnamic, caffeic, ferulic and chlorogenic acids with use of *Escherichia coli* K-12 recA::GFP biosensor strain. *Acta Poloniae Pharmaceutica – Drug Research*. 2017, 74(3), 801-808. ISSN 0001-6837.
37. CLACKSON, T. Regulated gene expression systems. *Gene Therapy*. 2000, 7, 120-125. ISSN 1476-5462.
38. CHENG, F., X-L. TANG and T. KARDASHLIEV. Transcription factor-based biosensors in high-throughput screening: advances and applications. *Biotechnology Journal*. 2018, 13(7). ISSN 1860-7314.
39. JIANG, T., C. LI and Y. YAN. Optimization of a *p*-coumaric acid biosensor system for versatile dynamic performance. *ACS Synthetic Biology*. 2021, 10(1), 132-144. ISSN 2161-5063.
40. WANG, Y., Q. LI, P. ZHENG, Y. GUO, L. WANG, T. ZHANG, J. SUN and Y. MA. Evolving the L-lysine high-producing strain of *Escherichia coli* using a newly developed high-throughput screening method. *Industrial Microbiology and Biotechnology*. 2016, 43(9), 1227-1235. ISSN 1367-5435.
41. SIEDLER, S., S. G. STAHLHUT, S. MALLA, J. MAURY and A. R. NEVES. Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into *Escherichia coli*. *Metabolic Engineering*. 2013, 21, 2-8. ISSN 1096-7176.
42. YANG, G. and S. G. WITHERS. Ultrahigh-throughput FACS-based screening for directed enzyme evolution. *ChemBioChem*. 2009, 10(17), 2704-2715. ISSN 1439-4227.

43. ALVAREZ-GONZALEZ, G. and N. DIXON. Genetically encoded biosensors for lignocellulose valorization. *Biotechnology for Biofuels*. 2019, 12, 246. ISSN 1754-6834.
44. XIONG, D., S. LU, J. WU, C. LIANG, W. WANG, W. WANG, J-M. JIN and S-Y. TANG. Improving key enzyme activity in phenylpropanoid pathway with a designed biosensor. *Metabolic Engineering*. 2017, 40, 115-123. ISSN 1096-7176.
45. UCHIYAMA, T. and K. MIYAZAKI. Metagenomic screening for aromatic compound-responsive transcriptional regulators. *Plos One*. 2013, 8(9). ISSN 1932-6203.
46. MACHADO, L. F. M. and N. DIXON. Development and substrate specificity screening of an in vivo biosensor for the detection of biomass derived aromatic chemical building blocks. *Chemical Communications*. 2016, 52(76), 11402-11405. ISSN 1359-7345.
47. GOKERKUCUK, E. B., M. TRAMIER and G. BERTOLIN. Imaging mitochondrial functions: from fluorescent dyes to genetically-encoded sensors. Genes. 2020, 11(2), 125. ISSN 2073-4425.
48. BAHL, M. L., L. H. HANSEN, T. R. LICHT and S. J. SORENSEN. *In vivo* detection and quantification of tetracycline by use of a whole-cell biosensor in the rat intestine. *Antimicrobial agents and chemotherapy*. 2004, 48(4), 1112-1117. ISSN 1098-6596.
49. TROGL, J., G. KUNCOVA, L. KUBICOVA, P. PARIK, J. HALOVA, K. DEMNEROVA, S. RIPP and G. S. SAYLER. Response of the bioluminescent bioreporter *Pseudomonas fluorescens* HK44 to analogs of naphtalene and salicylic acid. *Folia Microbiologica*. 2007, 52(1), 3-14. ISSN 0015-5632.
50. EGLAND, P. G. and C. S. HARWOOD. HbaR, a 4-hydroxybenzoate sensor and FNR-CRP superfamily member, regulates anaerobic 4-hydroxybenzoate degradation by *Rhodopseudomonas palustris*. *Journal of Bacteriology*. 2000, 182(1), 100-106. ISSN 0021-9193.
51. KUNJAPUR, A. M. and K. L. J. PRATHER. Development of a vanillate biosensor for the vanillin biosynthesis pathway in *E. coli*. *ACS Synthetic Biology*. 2019, 8, 1958-1967. ISSN 2161-5063.
52. JHA, R. K., T. L. KERN, D. T. FOX, C. E. M. STRAUSS. Engineering and *Acinetobacter* regulon for biosensing and high-throughput enzyme screening in *E. coli* via flow cytometry. *Nucleic Acids Research*. 2014, 42(12), 8150-8160. ISSN 0305-1048.
53. MARIENHAGEN, J. and M. BOTT. Metabolic engineering of microorganisms for the synthesis of plant natural products. *Journal of Biotechnology*. 2013, 163(2), 166-78. ISSN 0168-1656.
54. LIU, Y., Y. LIU, M. WANG. Design, optimization and application of small molecule biosensor in metabolic engineering. *Frontiers in Microbiology*. 2017, 17, 8. ISSN 1664-302X.
55. RUGBJERG, P., H. J. GENEE, K. JENSEN, K. SARUP-LYTZEN and M. O. A. SOMMER. Molecular buffers permit sensitivity tuning and inversion of riboswitch signals. *ACS Synthetic Biology*. 2016, 5, 632-638. ISSN 2161-5063.
56. ARNOUS, A., D. P. MAKRIS and P. KEFALAS. Effect of principal polyphenolic components in relation to antioxidant characteristics of aged red wines. *Journal of Agricultural and Food Chemistry*. 2001, 49(12), 5736-5742. ISSN 0021-8561.
57. KASAI, D., N. KAMIMURA, K. TANI, S. UMEDA, T. ABE, M. FUKUDA and E. MASAI. Characterization of FerC, a MarR-type transcriptional regulator, involved in transcriptional regulation of the ferulate catabolic operon in *Sphingobium Sp*. Strain SYK-6. *FEMS Microbiology Letters.* 2012, 332(1), 68-75. ISSN 0378-1097.

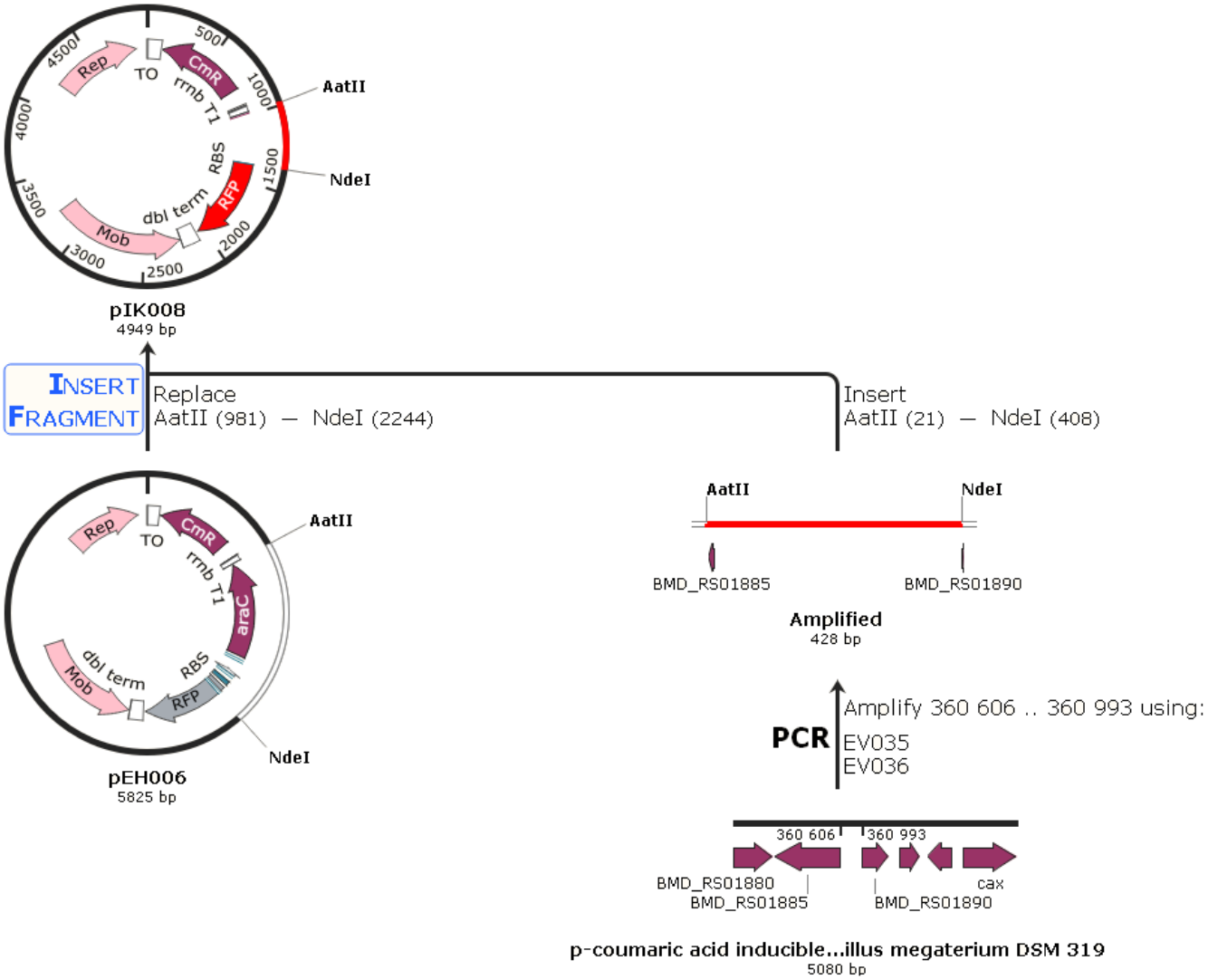
58. AHMADI, M. K., S. FAWAZ, C. H. JONES, G. ZHANG and B. A. PFEIFER. Total biosynthesis and diverse applications of the nonribosomal peptide-polyketide siderophore yersiniabactin. *Applied and Environmental Microbiology*. 2015, 81(16), 5290-5298. ISSN 0099-2240.
59. MIYAZAKI, K. Molecular engineering of the salicylate-inducible transcription factor Sal7AR for orthogonal and high gene expression in *Escherichia coli*. *Plos One*. 2018. ISSN 1932-6203.
60. SCHELL, M. A. and P. E. WENDER. Identification of the nahR gene product and nucleotide sequences required for its activation of the sal operon. *Journal of Bacteriology*. 1986, 166(1), 9-14. ISSN 0021-9193.
61. ROBSON, N. D., S. PAROTT and R. A. COOPER. In vitro formation of a catabolic plasmid carrying *Klebsiella pneumoniae* DNA that allows growth of *Escherichia coli* K-12 on 3-hydroxybenzoate. *Microbiology*. 1996, 142, 2115-2120. ISSN 1350-0872.
62. PRIEFERT, H., J. RABENHORST and A. STEINBUCHEL. Biotechnological production of vanillin. *Applied Microbiology and Biotechnology*. 2001, 56, 296-314. ISSN 0175-7598.
63. MORAWSKI, B., A. SEGURA and L. N. ORNSTON. Repression of *Acinetobacter* vanillate demethylase synthesis by VanR, a member of the GntR family of transcriptional regulators. *FEMS Microbiology Letters. 2000, 187, 65-58. ISSN 0378-1097.*
64. RIBBONS, D. W. Stoichiometry of demethylase activity in *Pseudomonas aeruginosa*. *FEBS Letters*. 1970, 8(2), 101-104. ISSN 0014-5793.
65. PRIEFERT, H., J. RABENHORST and A. STEINBUCHEL. Molecular characterization of genes of *Pseudomonas* Sp. Strain HR199 involved in bioconversion of vanillin to protocatechuate. *Journal of Bacteriology*. 1997, 179(8), 2595-2607. ISSN 0021-9193.
66. TRAUTWEIN, G. and U. GERISCHER. Effects exerted by transcriptional regulator PcaU from Acinetobacter sp. Strain ADP1. *Journal of Bacteriology*. 2001, 183(3), 873-881. ISSN 0021-9193.
67. CALISTI, C., A. G. FICCA, P. BARGHINI and M. RUZZI. Regulation of ferulic catabolic genes in *Pseudomonas fluorescens* BF13: involvement of a MarR family regulator. *Applied Microbiology and Biotechnology*. 2008, 80, 475-483. ISSN 0175-7598.
68. BENSON, D. A., M. CAVANAUGH, K. CLARK, I. KARSCH-MIZRACHI, J. OSTELL, K. D. PRUITT and E. W. SAYERS. Genbank. *Nucleic Acids Research*. 2017, 46, 41-47. ISSN 0305-1048.
69. JOHNSON, M., I. ZARETSKAYA, Y. RAYTSELIS, Y. MEREZHUK, S. McGINNIS and T. L. MADDEN. NCBI BLAST: a better web interface. *Nucleic Acids Research*. 2008, 36, 5-9. ISSN 0305-1048.
70. CROOKS, G. E., G. HON, J-M. CHANDONIA and S. E. BRENNER. WebLogo: a sequence logo generator. *Genome Research*. 2004, 14, 1188-1190. ISSN 1088-9051.
71. REESE, M. G. Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. *Journal of Computational Chemistry*. 2001, 26(1), 51-56. ISSN 0192-8651.
72. LEE, P. Y., J. COSTUMBRADO, C-Y. HSU and Y. H. KIM. Agarose gel electrophoresis for the separation of DNA fragments. *Journal of Visualized Experiments*. 2012, 62, 3923. ISSN 1940-087X.

73. SNYDER, L., J. E. PETERS, T. M. HENKIN and W. CHAMPNESS. *Molecular genetics of bacteria*. 4th edition. Herndon, 2007. ISBN 978-1555816278.
74. SAMBROOK, J. F. and D. W. RUSSELL. *Molecular cloning*: a laboratory manual. 3rd edition. New York, 2001. ISBN 0-87969-577-3.
75. BERGKESSEL, M. and C. GUTHRIE. Colony PCR. *Methods in Enzymology*. 2013, 529, 299-309. ISSN 0076-6879.
76. LIU, D., T. EVANS and F. ZHANG. Applications and advances of metabolite biosensors for metabolic engineering. *Metabolic Engineering*. 2015, 31, 35-43. ISSN 1096-7176.
77. MARTINS DOS SANTOS, V. A. P., S. HEIM, M. STRATZ and K. N. TIMMIS. Insights into the genomic basis of niche specificity of *Pseudomonas putida* KT2440. *Environmental Microbiology*. 2004, 6(12), 1264-1286. ISSN 1462-2920.
78. MOTA, F. L., A. J. QUEIMADA, S. P. PINHO and E. A. MACEDO. Aqueous solubility of some natural phenolic compounds. *Industrial & Engineering Chemistry Research*. 2008, 47, 15, 5182-5189. ISSN 0888-5885.
79. BACH, F., A. A. F. ZIELINSKI, C. V. HELM, G. M. MACIEL, A. C. PEDRO, A. P. STAFUSSA, S. AVILA and C. W. I. HAMINIUK. Bio compounds of edible mushrooms: *In vitro* antioxidant and antimicrobial activities. *LWT-Food Science and Technology*. 2019, 107, 214-220. ISSN 0023-6438.
80. BAJPAI, B. and S. PATIL. A new approach to microbial production of gallic acid. *Brazilian Journal of Microbiology*. 2008, 39, 708-711. ISSN 1678-4405.
81. OW, Y-Y. and I. STUPANS. Gallic acid and gallic acid derivatives: effects on drug metabolizing enzymes. *Current Drug Metabolism*. 2003, 4, 241-248. ISSN 1389-2002.
82. KAMBOURAKIS, S., K. M. DRATHS and J. W. FROST. Synthesis of gallic acid and pyrogallol from glucose: replacing natural product isolation with microbial catalysis. *Journal of the American Chemical Society*. 2000, 122, 9042-9043. ISSN 0002-7863.
83. YI, J., K. M. DRATHS, K. LI and J. W. FROST. Altered glucose transport and shikimate pathway product yields in *E. coli*. *Biotechnology Progress*. 2008, 19, 1450-1459. ISSN 1520-6033.
84. GIL, D. M. A., P. L. V. FALE, M. L. M. SERRALHEIRO and M. J. F. REBELO. Herbal infusions bioelectrochemical polyphenolic index: green tea – the gallic acid interference. *Food Chemistry*. 2011, 129, 1537-1543. ISSN 0308-8146.
85. KURBANOGLU, S., C. ERKMEN and B. USLU. Frontiers in electrochemical enzyme based biosensors for food and drug analysis. *Trends in Analytical Chemistry*. 2020, 124, 115809. ISSN 0165-9936.
86. HANKO, E. K. R., N. P. MINTON and N. MALYS. Characterisation of a 3-hydroxypropionic acid-inducible system from *Pseudomonas putida* for orthogonal gene expression control in *Escherichia coli* and *Cupriavidus necator*. *Scientific Reports*. 2017, 7, 1724. ISSN 2045-2322.
87. NOGALES, J., A. CANALES, J. JIMENEZ-BARBERO, J. L. GARCIA and E. DIAZ. Molecular characterization of the gallate dioxygenase from *Pseudomonas putida* KT2440: the prototype of a new subgroup of extradiol dioxygenases. *Journal of Biological Chemistry*. 2005, 280, 35382-35390. ISSN 0021-9258.

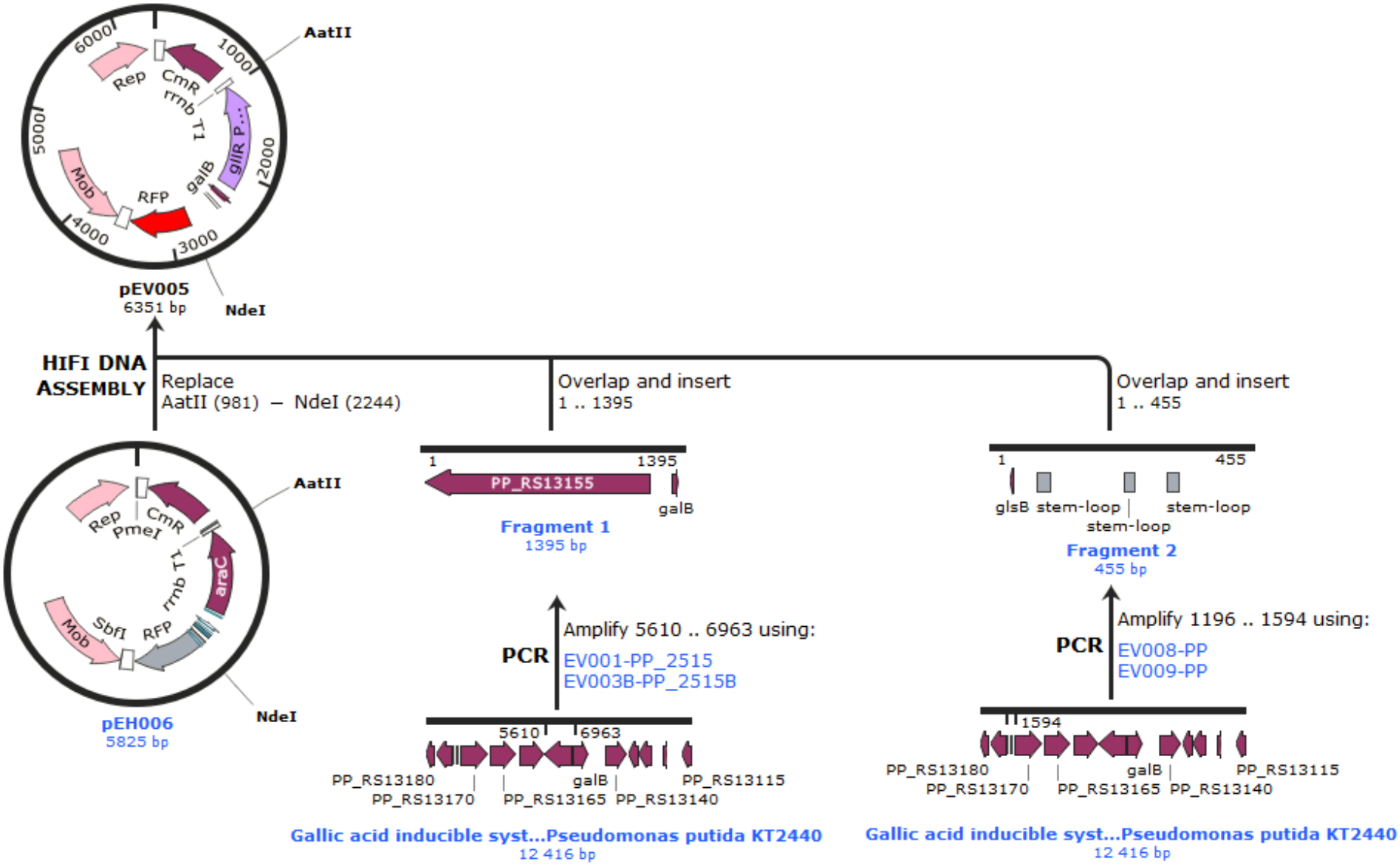
88. PORRUA, O., M. GARCIA-JARAMILLO, E. SANTERO and F. GOVANTES. The LysR-type regulator AtzR binding site: DNA sequences involved in activation, repression and cyanuric acid-dependent repositioning. *Molecular Microbiology*. 2007, 66, 410-427. ISSN 1365-2958.
89. CUSHNIE, T. P. T., B. CUSHNIE, J. ECHEVERRA, W. FOWSANTEAR, S. THAMMAWAT, J. L. A. DODGSON, S. LAW and S. M. CLOW. Bioprospecting for antibacterial drugs: a multidisciplinary perspective on natural product source material, bioassay selection and avoidable pitfalls. *Pharmaceutical Research*. 2020, 37, 125. ISSN 0724-8741.
90. KELL, D. B., N. SWAINSTON, P. PIR and S. G. OLIVER. Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis. *Trends in Biotechnology*. 2015, 33, 237-246. ISSN 0167-7799.
91. ANG, J., E. HARRIS, B. J. HUSSEY, R. KIL and D. R. McMILLEN. Tuning response curves for synthetic biology. *ACS Synthetic Biology*. 2013, 2, 547-567. ISSN 2161-5063.
92. LOU, Z., H. WANG, S. RAO, J. SUN, C. MA, J. LI. *p*-Coumaric acid kills bacteria through dual damage mechanisms. *Food Control*. 2012, 25, 550-554. ISSN 0956-7135.
93. PRAZERES, D. M. F. and G. N. M. FERREIRA. Design of flowsheets for the recovery and purification of plasmids for gene therapy and DNA vaccination. *Chemical Engineering and Processing*. 2004, 43, 615-630. ISSN 0255-2701.
94. HEINZLE, E., A. P. BIWER and C. L. COONEY. *Development of sustainable bioprocesses*: Modeling and assessment. Chichester: John Wiley & Sons, 2006. ISBN 978-0-470-01559-9.

Appendices

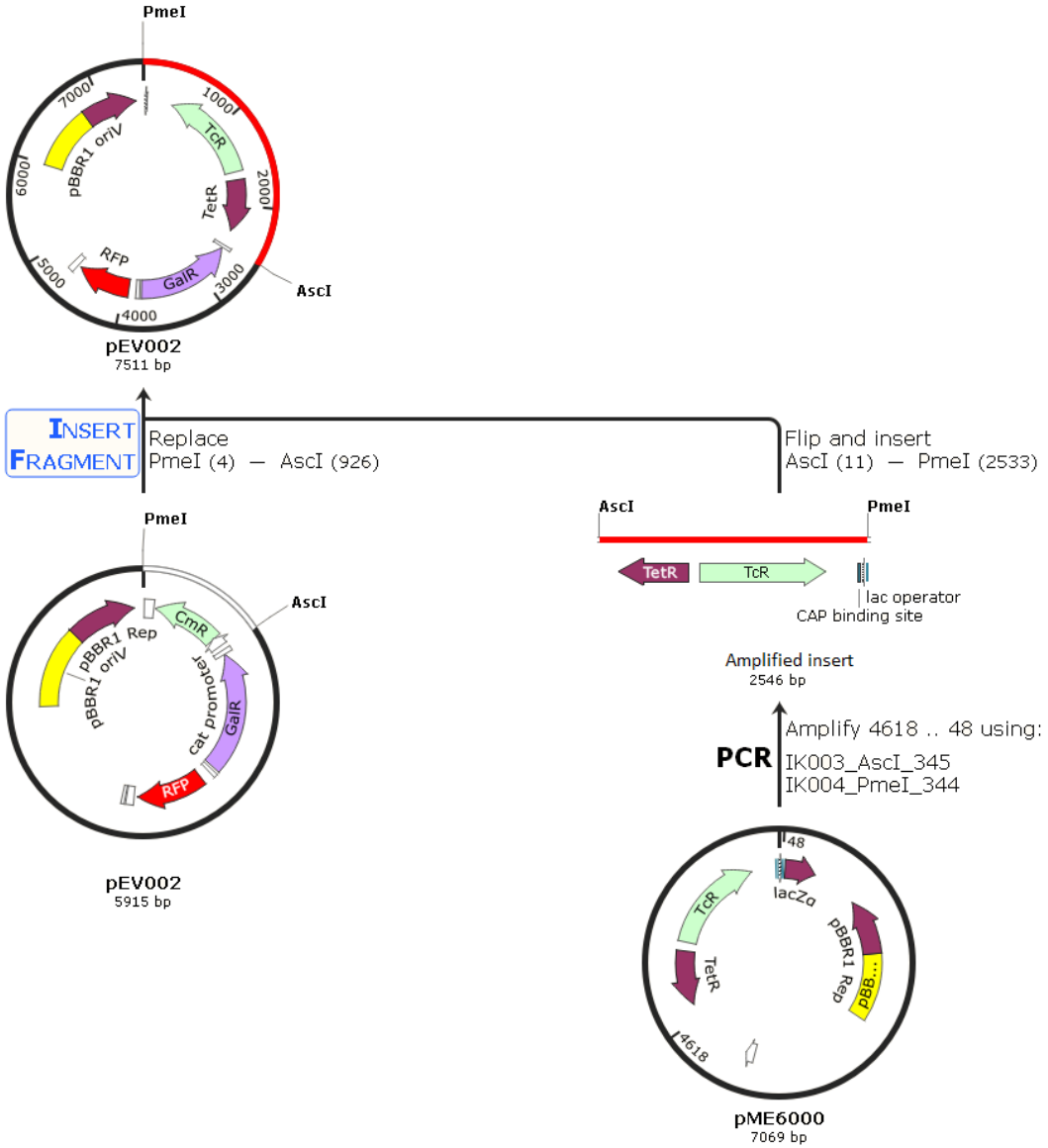
Appendix 1. Schematic view of constructs assembly using ligation reaction



Appendix 2. Schematic view of constructs assembly using HiFi assembly method



Appendix 3. Schematic view of antibiotic resistance gene replacement



Appendix 4. *galR*/*galB* intergenic regions

>*P. putida* KT2440_IGR

-aagggctcccggcgctcggctcgtgactgcggcgactat-
gcctctaccccgcttgcccggcaataaataaaacgcatagccgttgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-
ccccccactgcaggagccaacgaaac

>*P. putida* DOT-T1E_IGR

-aagggctcccggcgctcggctcgtgactgcggcgactat-
gcctctaccccgcttgcccggcaataaataaaacgcatagccgctgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-
ccccccactgcaggagccaacgaacc

>*P. putida* H8234_IGR

--------cccggcgctcggctcgtgactgcggcgactat-
gcctctaccccgcttgcccggcaataaataaaacgcatagccgttgcaaatgaacgcattagtgccctgcccaggccaatgacatcctgcgccaa-
ccccccactgcaggagccaac-----

>*P. putida* W619_IGR

aaggcttccccggcgcccggctcgtgactggggcgactat-
gcccctacccttcttgccaggcaataaataaaacgcatagccgctgcaagtgaacgcattagtgcctcgcccatgccgatgacatcctgcgtaaa-
accccactgcaggagccaacgcaca

>*P. putida*_ND6_IGR

-aagggctcccggcgctcggctcgtgactgcggcgactat-
gcctctaccccgcttgcccggcaataaataaaacgcatagccgctgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-
cccccactgcaggagccaacgaacc

>*P. putida*_F1_IGR

-aagggctcccggcgctcggctcgtgactgcggcgactat-
gcctctaccccgcttgcccggcaataaataaaacgcatagccgctgcaaatgaacgcattagtgccctgcacaagccaatgacatcctgcgccca-
cccccactgcaggagccaacgaacc

>*P. citronellolis*_IGR

-atcccctctcggcgcgcagccctactccgacgggactat-
gcccagccctccggacaagcgcaataagcagaacgcatacctccaacaactgaacgcattagtccgccacggccctcggtgacatcctcgacaca
-ttcccaaccggagcccacccac---

>*P. helmanticensis*_IGR

cgggcgaggactctgcgg-atcggcggttcaggcgactat-
gccgccggcatgtttgagtggcaataaaaaaaacgcatacctcgaacactcgaacgcattactccgtcggcggggtgcgatgcgatgctctttcgca
acgcaagaggagaatccgca----

>*P. laurylsulfativorans*_IGR

gggggacgggctctgcaaaggggcatttgagctggactat-
gccgctgccctgtttggatggcaataaaaaaaaaagcatacctcaaacacttcaatgcattactcctggggcgggtcgcggtgcgatgctcgttcgaaa
cgcaggaggagtatccgca----

>*P. fragi*_IGR

-
gggggcaggctctttgattgggcaatacagtgggactatgcctgcctggcagatgagctggcaataaataaaacgcatacctgtagcaagcttatgcattactccggggcggcaggtgggtgcgatcctttggccaactgttgaaggagcccgcac-----

>***P. putida*_NCTC13186_IGR**

-aagggctcccggcgctcggctcgtgactgcggcgactat-gcctctaccccgcttgcccggcaataaataaaacgcatagccgttgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-ccccccactgcaggagccaacgaaac

>***Pseudomonas_sp*_KBS0802 _IGR**

-aagggctcccggcgctcggctcgtgactgcggcgactat-gcctctaccccgcttgcccggcaataaataaaacgcatagccgttgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-ccccccactgcaggagccaacgaaac

>***Pseudomonas_sp*._XWY-1_IGR**

-aagggctcccggcgctcggctcgtgactgcggcgactat-gcctctaccccgcttgcccggcaataaataaaacgcatagccgctgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-ccccccactgcaggagccagcgaa--

>***P. putida*_B4_IGR**

-aagggctcccggcgctcggctcgtgactgcggcgactat-gcctctaccccgcttgcccggcaataaataaaacgcatagccgctgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-ccccccactgcaggagccaacgaa--

>***P. putida*_SJTE-1_IGR**

-aagggctcccggcgctcggctcgtgactgcggcgactat-gcctctaccccgcttgcccggcaataaataaaacgcatagccgctgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-ccccccactgcaggagccaacgaa--

>***P. putida*_B1_IGR [rev]**

--------cccggcgctcggctcgtgactgcggcgactat-gcctctaccccgcttgcccggcaataaataaaacgcatagccgttgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccaa-ccccccactgcaggagccaac-----

>***P. putida*_KF715_IGR**

--------cccggcgctcggctcgtgactgcggcgactat-gcctctaccccgcttgcccggcaataaataaaacgcatagccgttgcaaatgaatgcattagtgccctgcacaggccaatgacatcctgcgccaa-ccccccactgcaggagccaac-----

>***P. putida*_JB_IGR**

aagggttccccggcgctcggctcgtcactgcggcgactat-gcctctaccccgcttgcccggcaataaataaaacgcatagccgctgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-ccccccactgcaggagccaacgaa--

>***Pseudomonas_sp*._JY-Q_IGR**

-aagggctcccggcgctcggctcgtgactgcggcgactat-gcctctaccccgcttgcccggcaataaataaaacgcatagccgctgcaaatgaacgcattagtgccctgcacaggccaatgacatcctgcgccca-ccccccactgcaggagccaacgaa--

>***P. putida*_NBRC_14164_IGR**

-tgcgacagtcaactttc-tgcttgaggtgcgttgactat-
gccgatccgagacgaggagggcaataaggaaaacgcataccccatacacccgaacgcattactgcttcagtcagaccgaatgcgaagctgtgcct
atcgaaacagg-agcactcgca----

Appendix 5. *glsB*/*galT* intergenic regions

>*P. putida*_KT2440_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgcctcacgccgtcgtccggcggggtg
gttgctgctgggccattacgcgttcgaagccttaccggaacccggcgcaattcacaccaatacaagaacaaaggtgaccacgc

>*P. putida*_ND6_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgccccacgccgtcgtccggcggagtg
gttgctgctgggccattacacgttcgaagccttaccggaacccggcgcaattcacaccaatacaagaacaaaggtgaccacgc

>*P. putida*_F1_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgccccacgccgtcgtccggcggggt
ggttgctgctgggccattacgcgttcgaagccttaccggaacccggcgcaattcacaccaatacaagaacaaaggtgaccacgc

>*P. putida*_DOT-T1E_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgccccacgccgtcgtccggcggggt
ggttgctgctgggccattacgcgttcgaagccttaccggaacccggcgcaattcactccaatacaagaacaaaggtgaccacgc

>*P. putida*_BIRD-1_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgcctcacgccgtcgtccggcggggtg
gttgctgctgggccattacgcgttcgaagccttaccggaccccggcgcaactcacaccaatacaagaacaaaggtgaccacgc

>*P. putida*_JB_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgcctcacgccgtcgtccggcggggtg
gttgctgctgggccattacgcgttcgaagccttaccggaacccggcgcaactcacaccaatacaagaacaaaggtgaccacgc

>*P. putida*_NCTC13186_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgcctcacgccgtcgtccggcggggtg
gttgctgctgggccattacgcgttcgaagccttaccggaacccggcgcaattcacaccaatacaagaacaaaggtgaccacgc

>*Pseudomonas*_sp_KBS0802_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgcctcacgccgtcgtccggcggggtg
gttgctgctgggccattacgcgttcgaagccttaccggaacccggcgcaattcacaccaatacaagaacaaaggtgaccacgc

>*Pseudomonas*_sp._XWY-1_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgcctcacgccgtcgtccggcggggtg
gttgctgctgggccattacgcgttcgaagccttaccggaacccggcgcaactcacaccaatacaagaacaaaggtgaccacgc

>*P. putida*_B4_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgccccacgccgtcgtccggcggggt
ggttgctgctgggccattacgcgttcgaagccttaccggaacccggcgcaattcacaccaatacaagaacaaaggtgaccacgc

>*P. putida*_SJTE-1_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgccccacgccgtcgtccggcggagtg
gttgctgctgggccattacacgttcgaagccttaccggaacccggcgcaattcacaccaatacaagaacaaaggtgaccacgc

>*Pseudomonas sp*_JY-Q_glsB_IGR

gcgataaaaaaattggatatctcattaaagaaatcccgcaggcgctgccctggaccatgggattacagtccgccccacgccgtcgtccggcggggt
ggttgctgctgggccattacgcgttcgaagccttaccggaacccggcgcaattcacaccaatacaagaacaaaggtgaccacgc

>***P. putida*_B1_glsB_IGR**

gcgataaaaaaattagatatctcattaaagaaatcccgcaggcgctgc-
ctgggccatgggattagagtcgccctcacgccgtagtccggcggggtgattgctgctgggccgtagcgctctggaagccctaccgcaacccggcg
caattcacaccaatacaagaacaaaggtgaccacgc

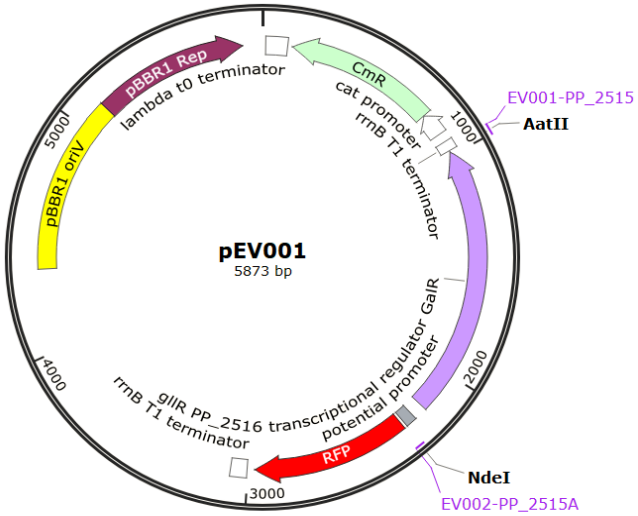
>***P. putida*_H8234_glsB_IGR**

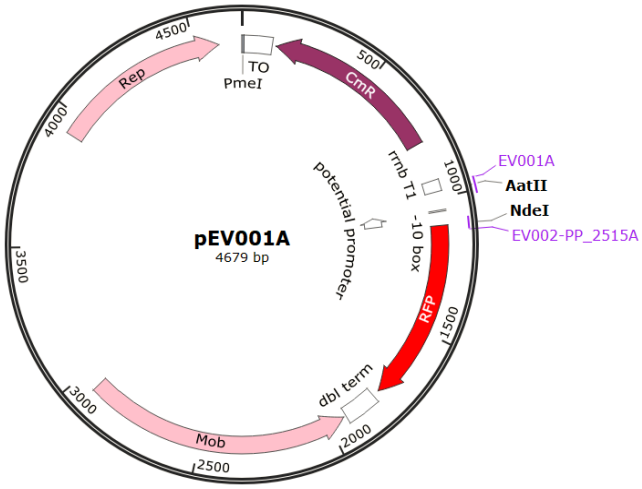
gcgataaaaaaattagatatctcattaaagaaatcccgcaggcgctgc-
ctgggccatgggattagagtcgccctcacgccgtagtccggcggggtgattgctgctgggccgtagcgctctggaagccctaccgcaacccggcg
caattcacaccaatacaagaacaaaggtgaccacgc

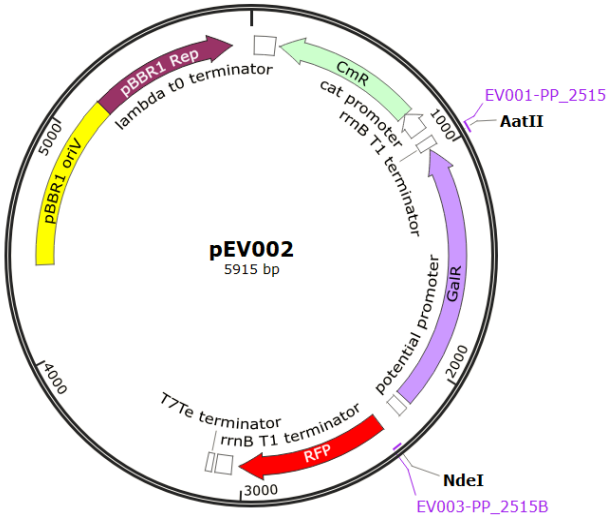
>***P. putida*_KF715_glsB_IGR**

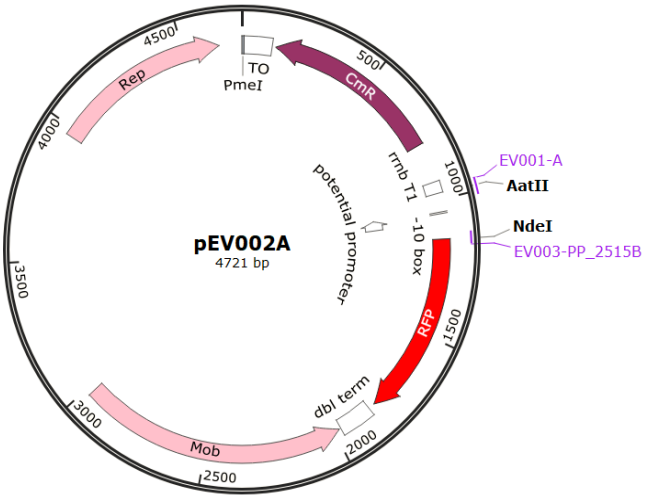
gcgataaaaaaattagatatctcattaaagaaatcccgcaggcgctgg-
ctgggccatgggattagagtcgccctcacgccgtagtccggcggggtgattgctgctgggccgtagcgctctggaagccctaccgcaacccggcg
caattcacaccaatacaagaacaaaggtgaccacgc

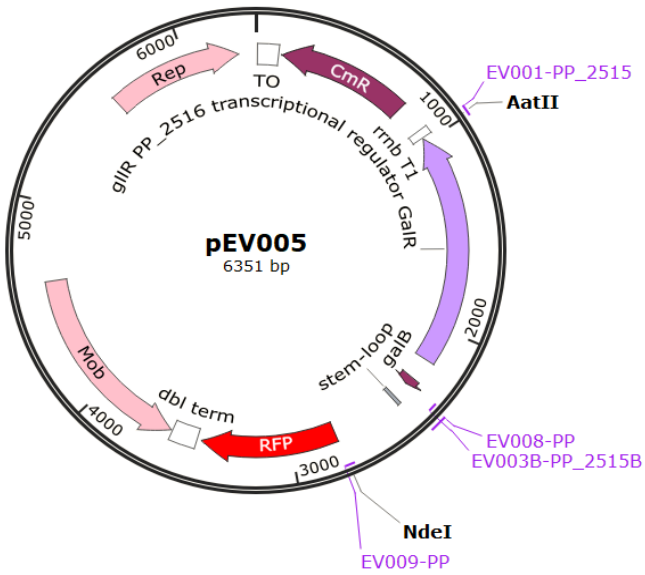
Appendix 6. Gallic acid-inducible plasmids designed and constructed in this study

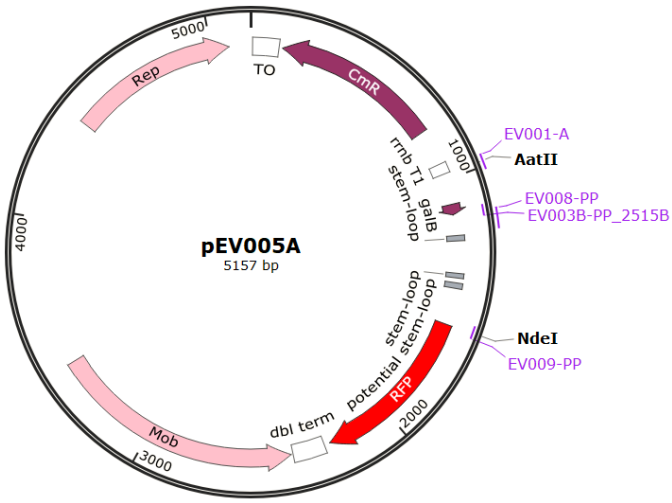


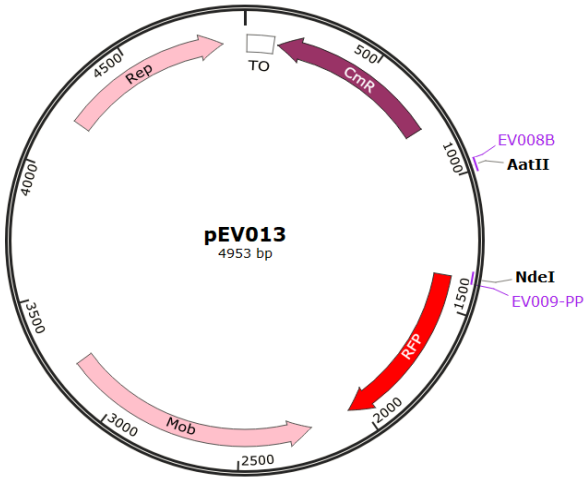


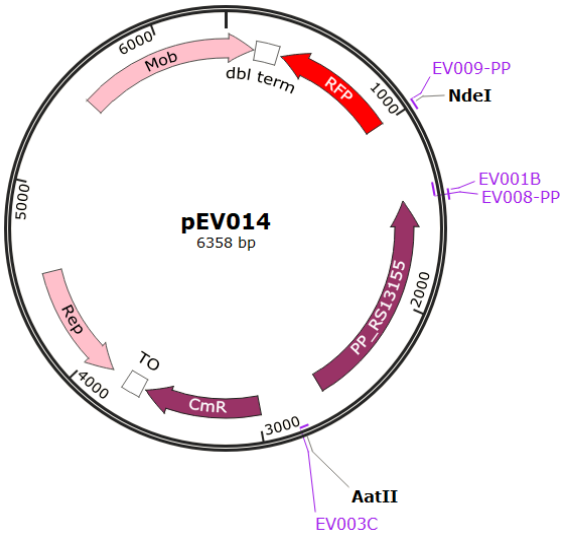


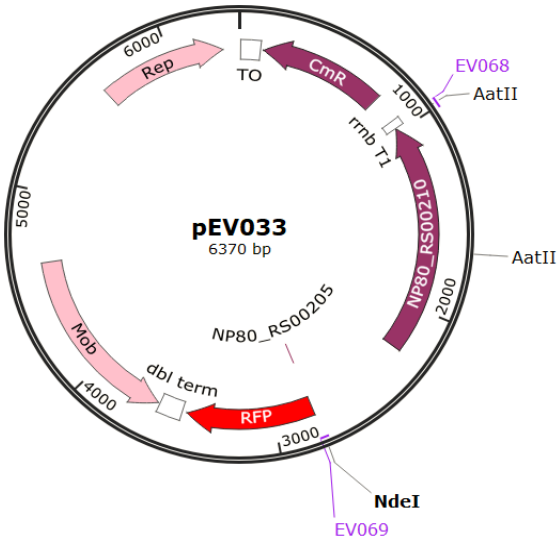


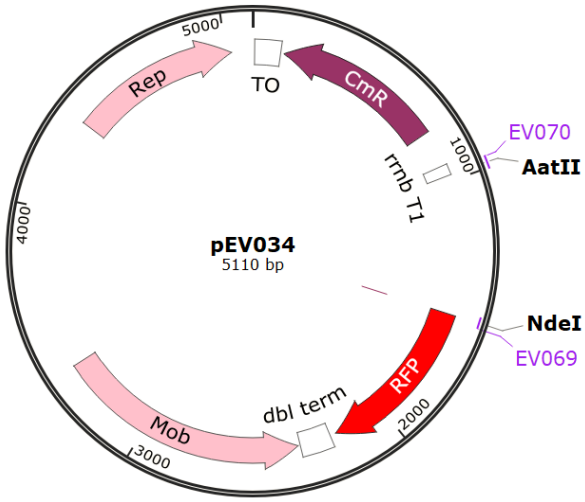


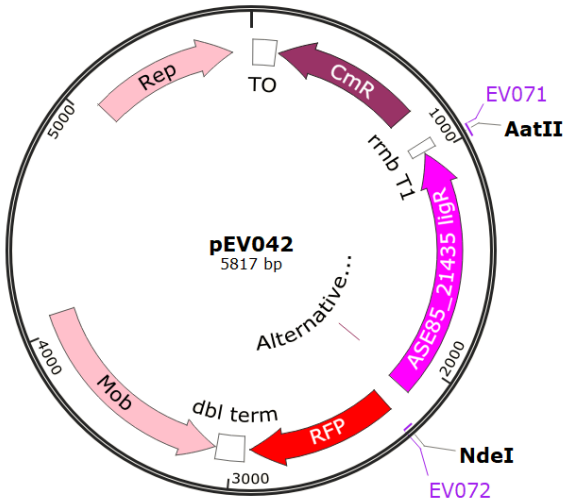


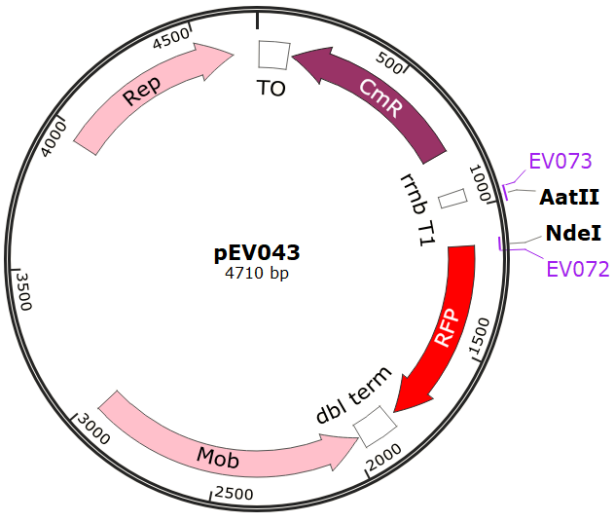

pEV013
4953 bp
TO
CmR
Rep
Mob
RFP
EV008B
AatII
NdeI
EV009-PP

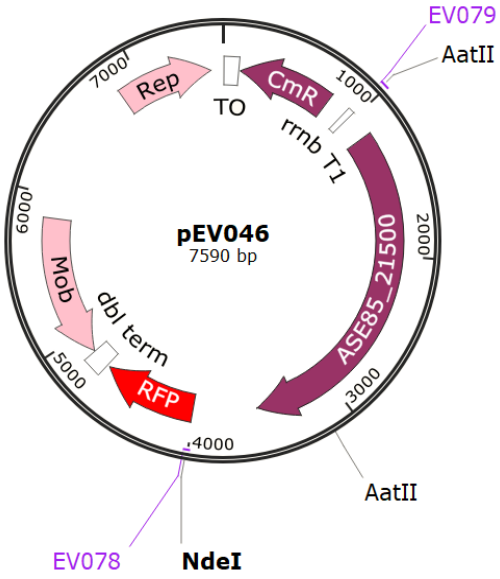

pEV014
6358 bp
TO
CmR
Rep
Mob
dbl term
RFP
PP_RS13155
EV009-PP
NdeI
EV001B
EV008-PP
AatII
EV003C

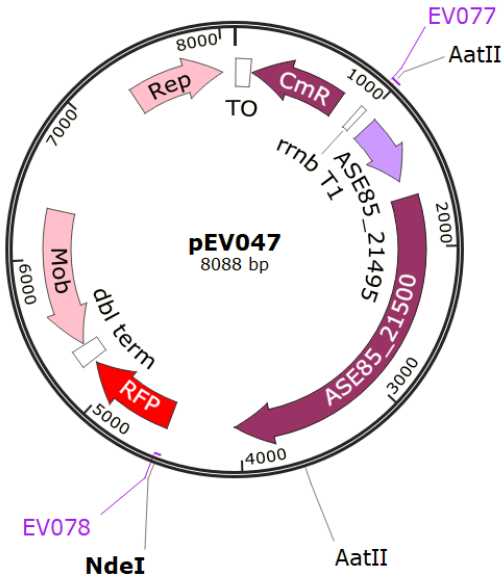

pEV033
6370 bp
TO
CmR
Rep
rrnb T1
NP80_RS00210
AatII
EV068
NP80_RS00205
Mob
dbl term
RFP
NdeI
EV069

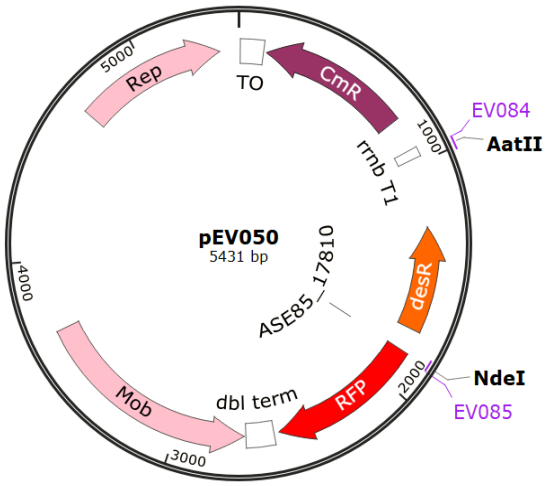

pEV034
5110 bp
TO
CmR
Rep
rrnb T1
Mob
dbl term
RFP
EV070
AatII
NdeI
EV069

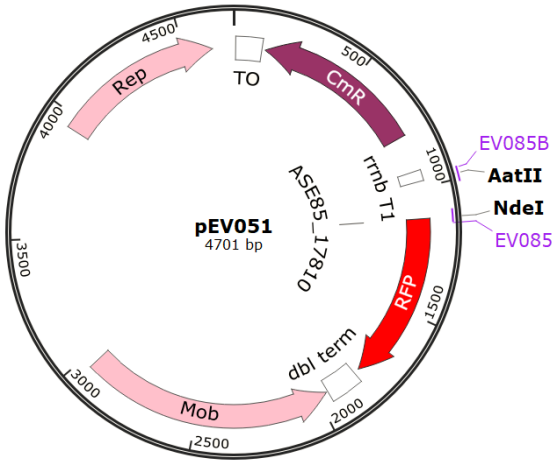

pEV042
5817 bp
TO
CmR
Rep
rrnb T1
ASE85_21435 ligR
Alternative...
Mob
dbl term
RFP
EV071
AatII
NdeI
EV072

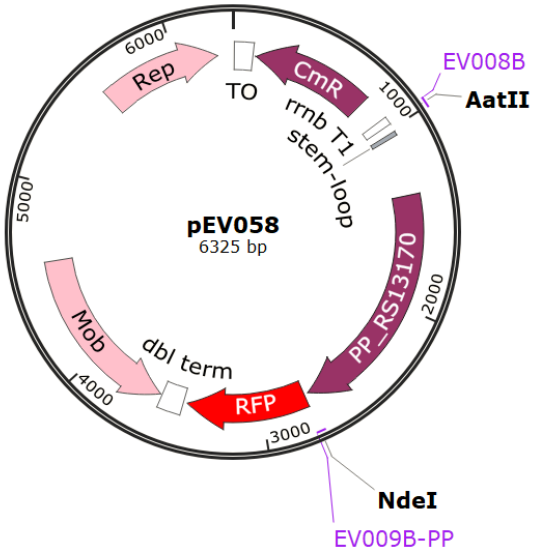

pEV043
4710 bp
TO
CmR
Rep
rrnb T1
Mob
dbl term
RFP
EV073
AatII
NdeI
EV072

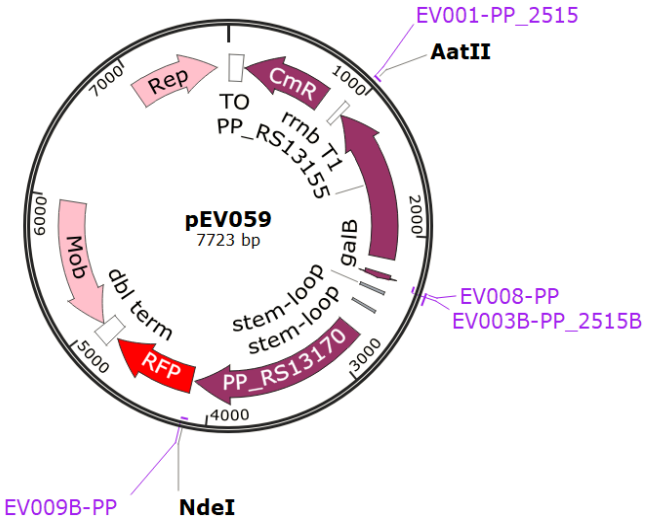

EV079
AatII
pEV046
7590 bp
Rep
TO
CmR
rrnb T1
ASE85_21500
Mob
dbl term
RFP
AatII
EV078
NdeI

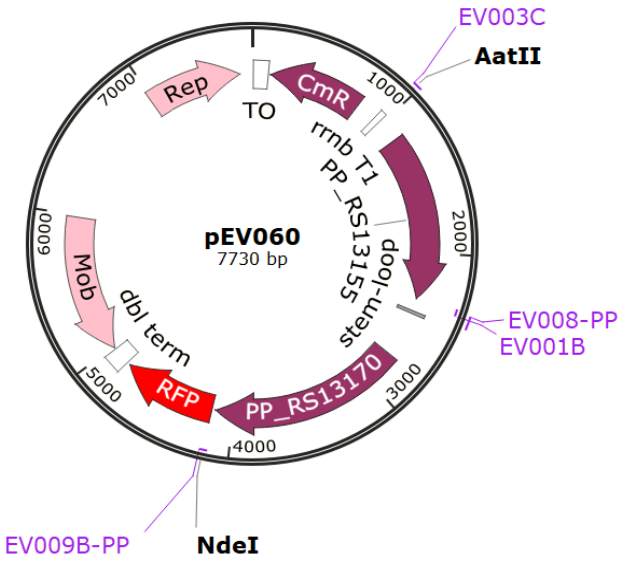

EV077
AatII
pEV047
8088 bp
Rep
TO
CmR
rrnb T1
ASE85_21495
ASE85_21500
Mob
dbl term
RFP
EV078
NdeI
AatII


pEV050
5431 bp
Rep
TO
CmR
rrnb T1
EV084
AatII
ASE85_17810
desR
NdeI
EV085
Mob
dbl term
RFP

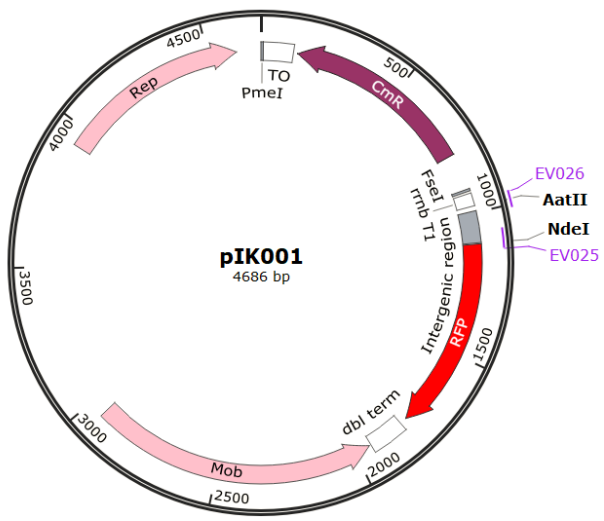

pEV051
4701 bp
Rep
TO
CmR
rrnb T1
EV085B
AatII
NdeI
EV085
ASE85_17810
RFP
dbl term
Mob

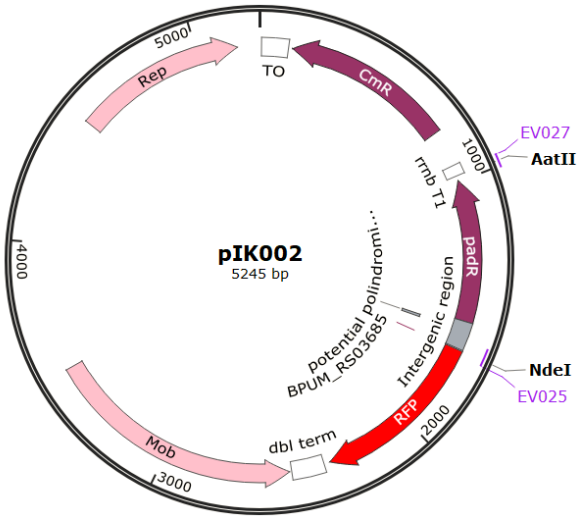

pEV058
6325 bp
Rep
TO
CmR
rrnb T1
stem-loop
EV008B
AatII
PP_RS13170
Mob
dbl term
RFP
NdeI
EV009B-PP

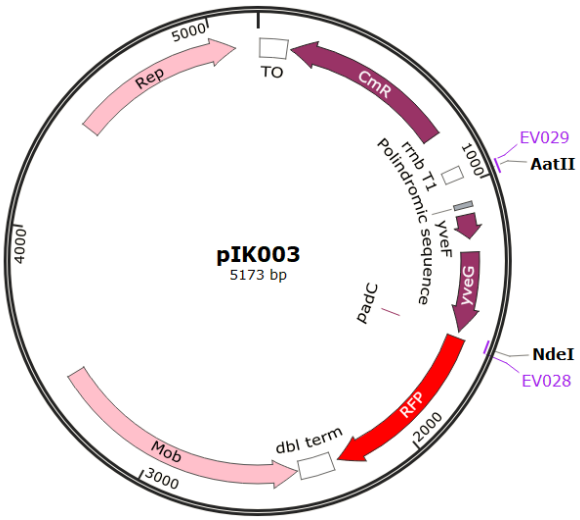

EV001-PP_2515
AatII
pEV059
7723 bp
Rep
TO
CmR
rrnb T1
PP_RS13155
Gfag
stem-loop
stem-loop
EV008-PP
EV003B-PP_2515B
Mob
dbl term
RFP
PP_RS13170
EV009B-PP
NdeI

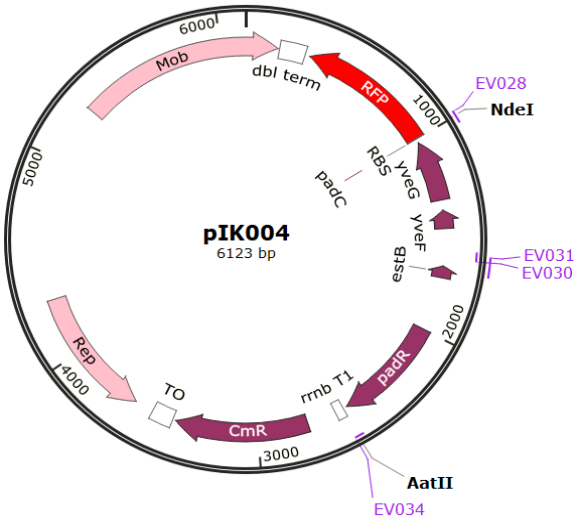

EV003C
AatII
Rep
CmR
TO
rrnb T1
PP_RS13155
stem-loop
pEV060
7730 bp
Mob
dbl term
RFP
PP_RS13170
EV008-PP
EV001B
EV009B-PP
NdeI
1000
2000
3000
4000
5000
6000
7000

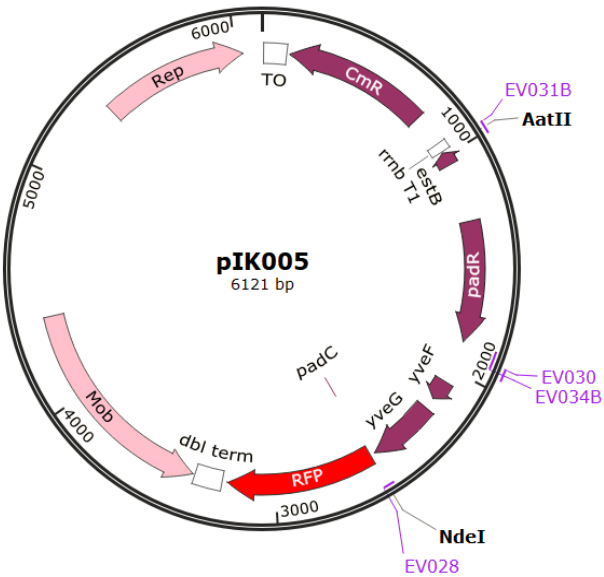
Appendix 7. *p*-Coumaric acid-inducible plasmids designed and constructed in this study

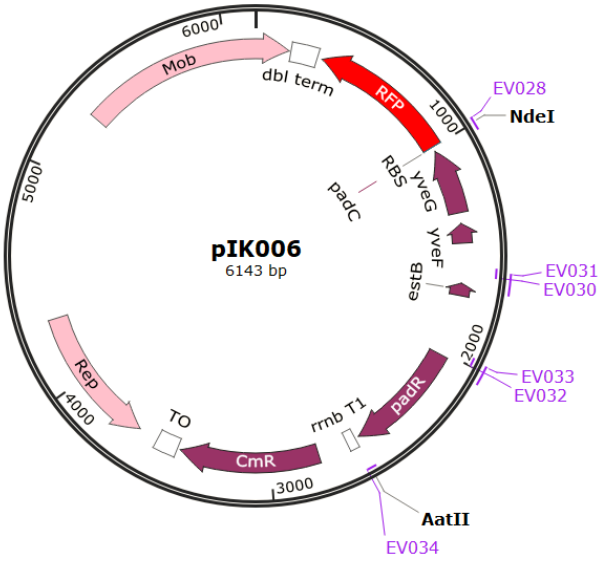


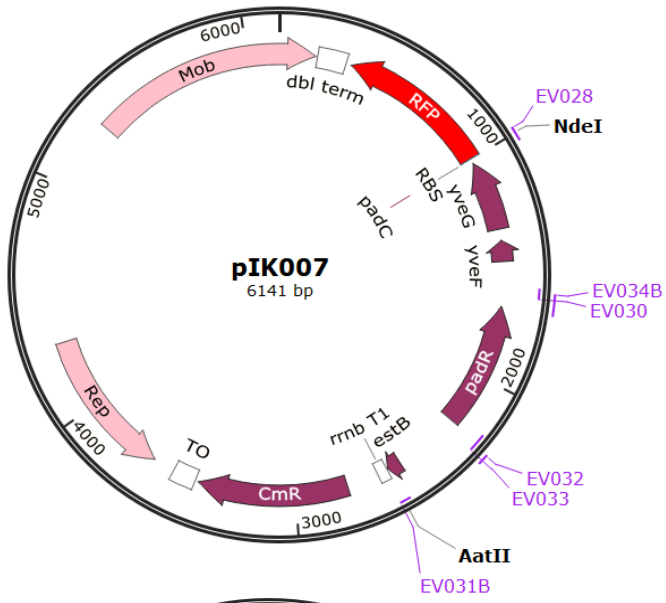


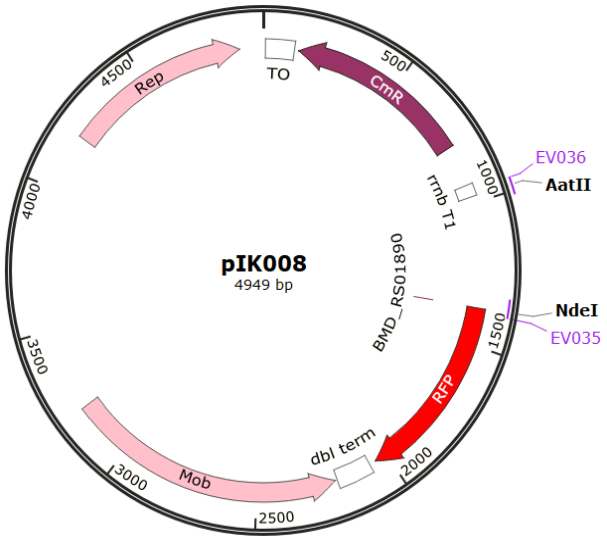


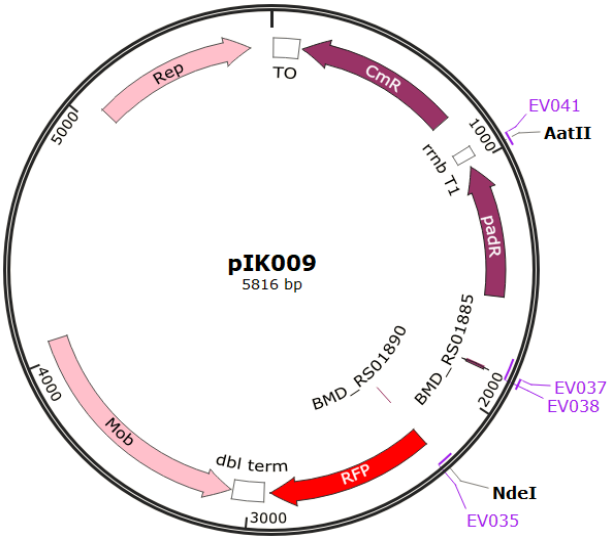


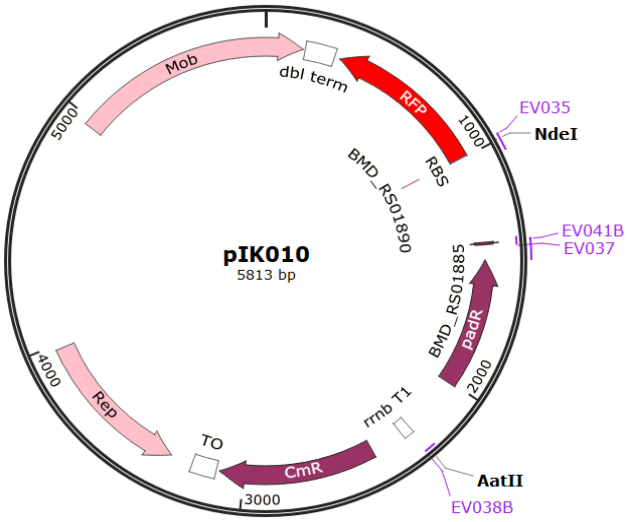


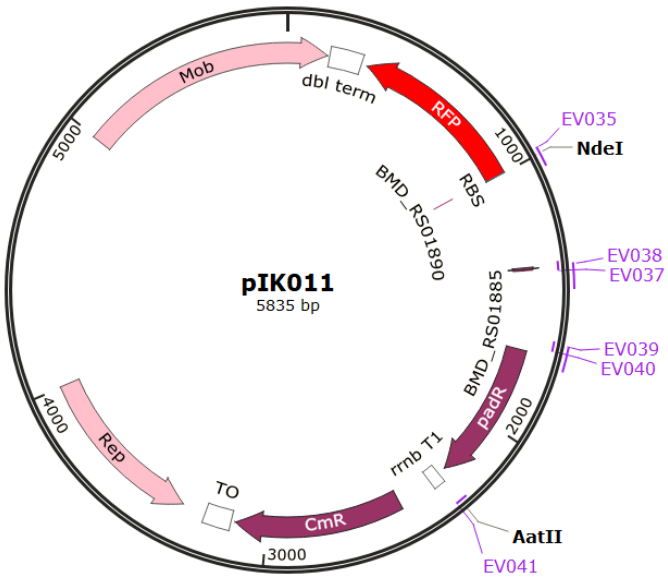


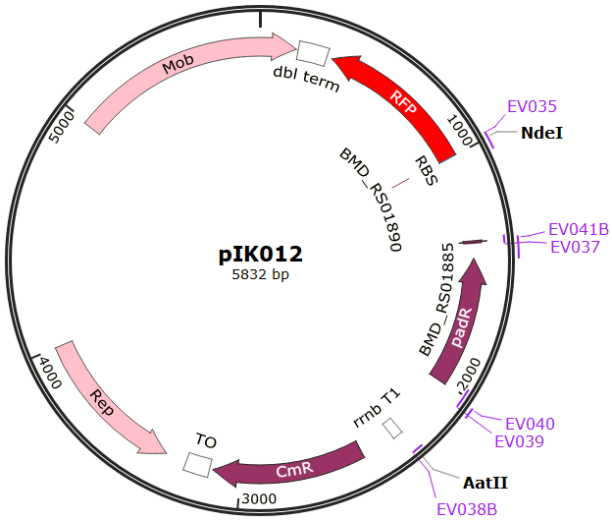

pIK007
6141 bp
Mob
dbl term
RFP
RBS
yveG
yveF
padC
padR
rrnb T1
estB
TO
CmR
Rep
EV028
NdeI
EV034B
EV030
EV032
EV033
AatII
EV031B

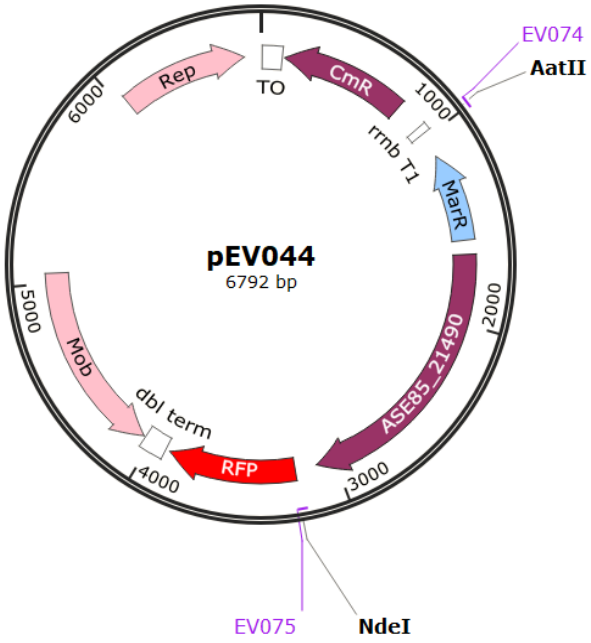

pIK008
4949 bp
TO
CmR
rrnb T1
BMD_RS01890
RFP
dbl term
Mob
Rep
EV036
AatII
NdeI
EV035

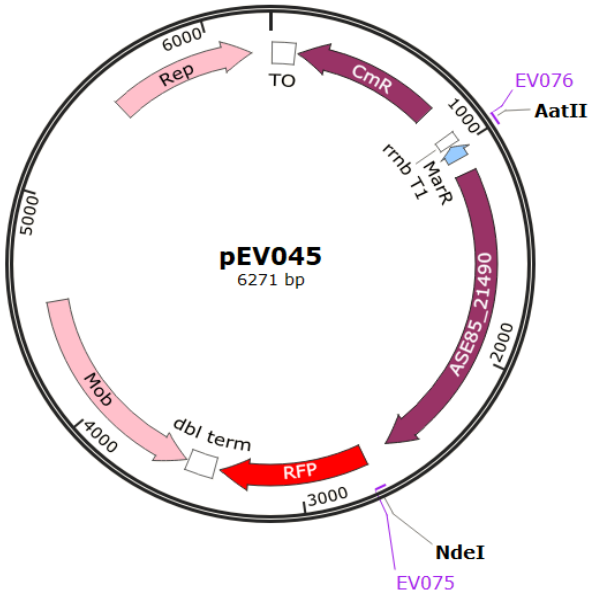

pIK009
5816 bp
Rep
TO
CmR
rrnb T1
padR
BMD_RS01890
BMD_RS01885
Mob
dbl term
RFP
EV041
AatII
EV037
EV038
NdeI
EV035


pIK010
5813 bp
Mob
dbl term
RFP
BMD_RS01890
RBS
BMD_RS01885
padR
Rep
TO
CmR
rrnb T1
EV035
NdeI
EV041B
EV037
AatII
EV038B


pIK011
5835 bp
Mob
dbl term
RFP
BMD_RS01890
RBS
BMD_RS01885
padR
rrnb T1
Rep
TO
CmR
EV035
NdeI
EV038
EV037
EV039
EV040
AatII
EV041


pIK012
5832 bp
Mob
dbl term
RFP
BMD_RS01890
RBS
BMD_RS01885
padR
Rep
TO
CmR
rrnb T1
EV035
NdeI
EV041B
EV037
EV040
EV039
AatII
EV038B


pEV044
6792 bp
Rep
TO
CmR
rrnb T1
MarR
ASE85_21490
dbl term
RFP
Mob
1000
2000
3000
4000
5000
6000
EV074
AatII
EV075
NdeI


pEV045
6271 bp
Rep
TO
CmR
rrnb T1
MarR
ASE85_21490
dbl term
RFP
Mob
1000
2000
3000
4000
5000
6000
EV076
AatII
NdeI
EV075