In Silico Design and Analysis of Genetic Circuit-Based Whole-Cell Biosensors for Detecting Groundwater Contaminants

Samuel Fajemilua, Solomon Bada, and M. Ahsanul Islam*

Department of Chemical Engineering, Loughborough University, Loughborough,
Leicestershire LE11 3TU, UK

*Corresponding author: m.islam@lboro.ac.uk

49   **Abstract**

50   Contaminants of emerging concern (CEC) such as tetracycline, erythromycin, and salicylic
 51   acid in groundwater can seriously endanger the environment and human health due to their
 52   widespread and everlasting harmful effects. Thus, continuous monitoring of various CEC
 53   concentrations in groundwater is essential to ensure the safety, security, and biodiversity of
 54   natural habitats. CECs can be detected using whole-cell biosensors for environmental
 55   surveillance and monitoring purposes, as they provide a cheaper and more robust alternative
 56   to traditional and expensive analytical techniques. In this study, various genetic circuit
 57   designs are considered to model three biosensors using the genetic design automation (GDA)
 58   software, iBioSim. The genetic circuits were designed to detect multiple CECs, including
 59   atrazine, salicylic acid, and tetracycline simultaneously to produce quantitative fluorescent
 60   outputs. The biosensor responses and the viability of the genetic circuit designs were further
 61   analysed using ODE-based mathematical simulations in iBioSim. The designed circuits and
 62   subsequent biosensor modelling presented here, thus, not only show the usefulness and
 63   importance of GDA tools, but also highlight their limitations and shortcomings that need to
 64   overcome in the future; thereby, providing a practical guidance for further improvement of
 65   such tools, so that they can be more effectively and routinely used in synthetic biology
 66   research.

67

68   **Keywords:** Whole-cell biosensor, genetic circuit, genetic design automation, iBioSim,
 69   groundwater contaminants

70

71

72

73

74

75

76

77 **Introduction**

78 The subsurface groundwater aquifers are enormous storehouses of water and one of the
79 Erath's most valuable resources. These aquifers provide groundwater in usable quantities for
80 world's vast majority of people, as their uses range from agricultural use such as irrigation for
81 crop production to industrial and public water supply use. Around 25% of the UK's public
82 water supply comes from groundwater [1]. In North America alone, around 150 million
83 people rely on groundwater for their drinking water supply [2]. Similarly, 85% of drinking
84 water supplies are provided by groundwater in developing countries such as India [3]. On
85 account of this, interest in the groundwater environment has heightened, and resources are
86 being administered to ensure and maintain its safety. Of these resources, many are focused on
87 the removal of contaminants which mainly originate from agricultural, industrial, and
88 municipal activities [4], [5]. Some of these contaminants have been reported to cause a range
89 of health problems, including cancer [6], diabetes [7], and birth defects [8]. Accordingly, the
90 detection and removal of these contaminants are of great importance for the safety and
91 security of not only these valuable groundwater aquifers but also of the environment and
92 public health in general.

93

94 Over the past decade, certain chemical compounds that were previously present at
95 insignificant levels in groundwater have become increasingly detectable. These chemical
96 compounds have been referred to as contaminants of emerging concerns (CECs) [9]. CECs
97 are commonly found in pesticides, herbicides, fungicides, and pharmaceutical compounds
98 [10]. The pharmaceutical CECs, such as tetracycline, erythromycin, and salicylic acid have
99 been shown to increase the rates of causing different types of cancers, organ damage, and
100 antibiotic resistance [11]–[13]. Similarly, the CECs commonly found in pesticides, including
101 atrazine, allethrin, and carbaryl have been shown to induce biodiversity damage, pest
102 resistance, and endocrine disruption [13]–[16]. In addition, Eriksen *et al* [17] observed a 30%
103 - 40% increase in prostate cancer due to the CEC, perfluorooctanoate. Further,
104 trihalomethanes present in groundwater were associated with an increased risk of
105 spontaneous abortions [18]. However, due to the limited number of studies and contradictions
106 in study outcomes, the connection between CECs and health risks have not yet been fully
107 established. Further research is required to fully elucidate this connection.

Previously, groundwater contaminants were detected by means of geophysical methods such as NMR geotomography [19] and spectroscopic techniques such as UV lasers [20]. In the past ten years, significant advances in synthetic biology research have been achieved [21]. These advances have resulted in the use of biosensors to detect contaminants such as arsenic [22], toluene [23], and xylene [24]. Biosensors utilise biological responses and physiochemical components to generate a measurable signal for detection of an object [25]. Depending on the transducing mechanism used, these devices can exist in many forms such as electrochemical biosensors, optical biosensors, and resonant biosensors [26]. However, the development cost and operating boundaries associated with these biosensors have prompted questions on their feasibility for environmental surveillance and monitoring applications. The advances in synthetic biology [27], together with the cost pressures and operating boundaries, have provoked notable interests in the development of whole-cell biosensors due to their low cost and versatility [28].

Whole-cell biosensors incorporate microorganisms to detect specific pollutants by receiving signals and processing these signals via genetic circuits to produce a quantitative or qualitative output. Genetic circuits can process signals by exploiting transcriptional regulation to orchestrate the gene expression [28]. Another technique is through the incorporation of Boolean logic gates in the cell [29]. Boolean logic gates make use of the logic functions such as AND, OR, NOT, and all of their combinations [30]. Once the input requirements of these logic functions are met in the gates, an output is executed. Whole-cell biosensors tend to detect a wider range of substances when compared to conventional biosensors [31]. Moreover, whole-cell biosensors can operate in much harsher pH values and temperatures than its conventional counterparts [31]. Although whole-cell biosensors boast great advantages, synthetic biologists may face challenges when designing feasible biosensors, as issues may arise when implementing these designs in the field. For this reason, the whole-cell biosensors have not yet reached their potential. Accordingly, attention has been directed towards the optimisation and practicality of whole-cell biosensor designs [32].

Preceding efforts to construct whole-cell biosensors to detect groundwater contaminants are similar to that of Liu *et al* [33]. A number of studies have further demonstrated the scalability of whole-cell biosensors, detecting a range of contaminants [22]–[24]. However,

the studies focused on single-input biosensors rather than considering the multi-input options. Multi-input biosensors are capable of handling complex environments such as groundwater aquifers, and this can be achieved through the integration of Boolean logic gates [34]. Moreover, biosensors for the detection of CECs have been poorly studied in comparison to other groundwater contaminants.

Recent advances in computational synthetic biology has allowed the development of computational models of genetic circuits to be implemented in a whole-cell biosensor. Over the past decade, various software tools have been developed to aid the advancement of synthetic biology designs. One such software is the genetic design automation (GDA) tool, iBioSim [35], [36]. iBioSim provides an environment whereby visualisation, modelling, and analysis of genetic circuits can be achieved; thus, permitting the analysis of various circuit design options in terms of robustness, operating conditions, and operating boundaries [35]. This study uses iBioSim to (1) design three genetic circuits capable of detecting three CECs: atrazine, salicylic acid, and tetracycline, both individually and simultaneously to produce measurable florescent outputs, (2) model and simulate these genetic circuits in the context of biological cells to emulate how the whole-cell biosensors work, and (3) analyse the results obtained from the modelling and simulation of the designed genetic circuits as biosensors. Furthermore, the factors to be considered for experimental implementation of the biosensor models will be discussed, and the drawbacks which may arise will be underscored.

Materials and methods

Genetic Circuit Design and Construction

To increase consistency and replicability, all genetic circuit designs proposed in this study were designed, modelled, and simulated in iBioSim [37]. An example of the circuit design procedure has also been described in detail in the supplementary information.

A genetic circuit usually consists of several biological parts, including promoters, ribosome binding sites (RBS), proteins, and terminators. A promoter is the specific DNA region which initiates the transcription of a gene [38]. Translation is then initiated at the RBS whereby messenger RNA (mRNA) is decoded to produce a protein [39]. The terminator is a section of

DNA which stops the transcription process [40]. All biological parts were obtained from SynBioHub [41] and were designed following the BioBricks^{TM} standards [42]. Figure 1 illustrates a genetic circuit designed for the detection of salicylic acid using standard parts from the Registry of Standard Biological Parts (for example, BBa_B0015 is the BioBricks name of a double terminator) and expressed in Synthetic Biology Open Language Visual (SBOL Visual) format. SBOL is an open-source standardised data format for sharing and exchanging in silico genetic circuit designs [43]. SBOL Visual was developed as an extension of SBOL to promote coherency when expressing the function and structure of genetic designs [44]. In Figure 1, the coding sequence codes for the red fluorescent protein, RFP (BBa_I766101), where coding for this protein is activated by the salicylic acid-sensitive promoter, P_{SalI} (BBa_J61051). The ribosome binding site used to mediate the protein synthesis is RBS30 (BBa_K1789002), and the double terminator (BBa_B0015) was employed to end gene transcription. The various standard biological parts used from the parts registry to design the biosensors are summarised in Table 1. The detail SBOL Visual design for all biosensors can be found in the supplementary information.

Modelling biosensor circuits in iBioSim

The genetic circuit designs for simulating three different types of whole-cell biosensors in iBioSim are shown in Figure 2 (Fig. 2A is the ternion biosensor, Fig. 2B is the triple-input-AND-gated biosensor, and Fig. 2C is the genetic toggle switch biosensor). The ternion biosensor (Fig. 2A) is comprised of a three-cell consortium containing one sub-circuit for each cell. The sub-circuits are placed in Cell_1, Cell_2, and Cell_3 to detect atrazine, tetracycline, and salicylic acid for producing green fluorescent protein (GFP), yellow fluorescent protein (YFP), and RFP signals, respectively. The model also includes eight species (blue ovals in Fig. 2A) and four complex formation reactions (light green squares in Fig. 2A); species represent chemical components such as compounds, genes, and proteins, while complex formation reactions represent the chemical reactions transforming these species. These reactions are further described by four sub-models showing the formation of C1 TetR_production, C2 TetR_YFP, C3 Atrazine_GFP, and C4 Salicylic_acid_RFP (Fig. 2A). Further details of all sub-models are shown in the supplementary information. Next, events (dark green rectangles in Fig. 2), representing biosensor responses to varying concentration of atrazine, tetracycline, and salicylic acid, were added to the models. In the case of Cell_1, the atrazine concentration was changed through the events, Atrazine_High and Atrazine_Low to simulate high and low atrazine concentrations in the environment.

Similarly, tetracycline and salicylic acid concentration changes were simulated in Cell_2 and Cell_3, respectively (Fig. 2A). The triple-input-AND-gated (Fig. 2B) and the genetic toggle switch (Fig. 2C) biosensor models were also simulated in iBioSim following the same approach as described for the ternion biosensor model.

Mathematical basis of the designed biosensor circuits

Ordinary differential equations (ODEs) were used to simulate the biosensors in iBioSim. The Runge-Kutta-Fehlberg method was used to find numerical approximations to solve the ODEs [35]. To mathematically model the biosensors, the underlying reactions associated with gene expression must be deduced. The reactions in relation to the conversion of DNA to mRNA to protein can be written as [45]:

Where $k_{transcription}$ and $k_{translation}$ represent the rates of conversion of DNA to mRNA and mRNA to protein, respectively, while γ_m and γ_p are the degradation rates of mRNA and protein, respectively.

Thus, the rate of change of mRNA and protein level over time can be expressed using the following ODEs [45]:

$$\frac{d[mRNA]}{dt} = k_{transcription}[DNA] - \gamma_m[mRNA] \dots \dots [1]$$

$$\frac{d[Protein]}{dt} = k_{translation}[mRNA] - \gamma_p[Protein] \dots \dots [2]$$

Before using the above differential equations for simulating biosensor responses, it is essential to formulate a mathematically accurate depiction of how a transcription factor regulates the gene expression. When modelling genetic circuits, the approach commonly

taken emanates from the mass action kinetics [45]. According to this principle, the rate of a reaction in the genetic circuits is expressed as the product of a rate constant multiplied by the concentration of the inducer [46]. In order to use this kinetic scheme for a genetic circuit, the following assumptions are required [45], [47]:

1. Polymerase binding, ribosome binding, and other intricacies of molecular interactions are ignored to reduce the complexity of the model,
2. The transcriptional activity is considered at steady state to accommodate for rapid change of transcriptional activity by virtue of the input signals, and
3. The spatial effects are neglected to avoid partial differential equations.

These assumptions led to the final form of the equations used to model the synthetically constructed biosensors. In the example of the genetic toggle switch, this would be:

$$\frac{d[mRNA]}{dt} = \frac{k_{transcription}[DNA]}{1 + \left(\frac{[Repressor]}{\left(1 + \left(\frac{[Inducer]}{K_s}\right)^{n_1}\right) K_d} \right)^{n_1}} - \gamma_m[mRNA] \dots \dots [3]$$

$$\frac{d[Repressor]}{dt} = k_{translation}[mRNA] - \gamma_p[Repressor] \dots \dots [4]$$

Where $K_s^{n_1}$ and $K_d^{n_1}$ are constants. This permits the development of a system of ODEs for the genetic toggle switch leading to the following final solution:

$$[Protein] = \frac{Kk_{translation}}{\gamma_p\gamma_m} + \frac{Kk_{translation}}{\gamma_m^2 - \gamma_p\gamma_m} e^{-\gamma_m t} - \left(\frac{Kk_{translation}}{\gamma_p\gamma_m} + \frac{Kk_{translation}}{\gamma_m^2 - \gamma_p\gamma_m} \right) e^{-\gamma_p t} \dots \dots [5]$$

Where K can be defined as:

$$K = \frac{k_{transcription}[DNA]}{1 + \left(\frac{[Repressor]}{\left(1 + \left(\frac{[Inducer]}{K_s}\right)^{n_2}\right) K_d} \right)^{n_1}}$$

The full derivation of the genetic toggle switch is shown in the supplementary information.

Results and Discussion

Design and construction of biosensors

1. Ternion biosensor

A ternion whole-cell biosensor was modelled using a three-cell consortium to detect atrazine, salicylic acid, and tetracycline by producing green, red, and yellow fluorescent signals as outputs, respectively (Figure 3). While the natural signalling mechanisms of host cells were used to detect salicylic acid and tetracycline, a synthesised riboswitch was used to provoke the transcription and translation of GFP in the presence of atrazine. The schematics of different sub-circuits of the ternion biosensor are shown in Figures 3A, 3B, and 3C. For the atrazine genetic circuit (Fig. 3A), the constitutive promoter, BBa_J23115 is used to continuously activate the atrazine riboswitch. Initially, the RBS is blocked to prevent the translation of GFP. In the presence of atrazine, an aptamer-atrazine bond is created, which unblocks the RBS to allow the transcription and translation of GFP to produce green fluorescence signals (Fig. 3A). In contrast to the atrazine genetic circuit, the salicylic acid sensor utilises a salicylic acid-sensitive promoter, P_{SalI} (Fig 3B). In the presence of salicylic acid, the promoter activates the transcription and translation of RFP, producing red fluorescent signals. This incident demonstrates the detection of salicylic acid by the production of red signals as the output. The tetracycline genetic circuit (Fig. 3C) shows a constitutive promoter, BBa_J23106, continuously expressing the tetracycline-sensitive repressor, TetR. In the absence of tetracycline, TetR represses the tetracycline-sensitive promoter, P_{Tet} ; thereby, inhibiting the production of YFP. However, the presence of tetracycline results in a tetracycline-TetR complex formation leading to the reduction of TetR- P_{Tet} binding affinity. Consequently, P_{Tet} activates the transcription and translation of YFP to produce yellow fluorescent signals as the output; thus, achieving the detection of tetracycline by the ternion biosensor.

2. Triple-input-AND-gated biosensor

In situations such as groundwater contamination detection, demand for accuracy and specificity is prominent. Sensor systems with the ability to implement complex functions are

essential to detect various contaminants in parallel [29]. Taking this into account, a biosensor making use of the Boolean logic function, AND was designed and modelled. For the triple-input-AND-gated biosensor (Figure 4), a two-cell consortium containing three genetic circuits was utilised to produce a single green fluorescent output. The first cell includes two genetic circuits, sensing the detection of tetracycline and salicylic acid, while the second cell contains only the atrazine detecting circuit. The sensors are of a similar structure to that of the ternion biosensor, the only difference being their respective output signals. For this biosensor model, the outputs for the detection of tetracycline and salicylic acid are the production of hrpR and hrpS proteins, respectively. Both hrpR and hrpS are enhancer-binding proteins that act as positive regulatory factors for P_{hrpL} , the promoter for the alternative sigma factor, hrpL; thus, having the capability to activate it [48]. Upon activation of P_{hrpL} , the TetR protein is produced by the first cell. TetR has the ability to diffuse across cells to form a complex with tetracycline molecules [49]. The tetracycline-TetR complex is capable of activating the P_{Tet} promoter in the second cell. Thus, in the presence of atrazine, both P_{Tet} and the atrazine riboswitch are activated to produce GFP, generating green fluorescent signals as the output (Figure 4). Therefore, the generation of green signals by the triple-input-AND-gated biosensor indicates the presence of all three contaminants in the system.

3. Toggle switch biosensor

The genetic toggle switch biosensor is constructed using two repressors (TetR and Salicylic acid repressor), two constitutive promoters (P_{Tet} and P_{SalI}), and two reporter genes (GFP and RFP) (Figure 5). The promoters are able to inhibit each other through the repressor proteins they code for. This design was selected as it requires the fewest genes and cis-regulatory elements theoretically to achieve a robust bistable behaviour [50]. The bistability of a genetic toggle switch arises from the mutually inhibitory set up of the repressor genes. Although the previous two biosensor designs were able to detect all three CECs simultaneously, the genetic toggle switch can detect only two contaminants at a time. The toggle switch biosensor is, thus, modelled to detect tetracycline (Figure 5A) and salicylic acid (Figure 5B). Fig. 5A illustrates state 1 of the genetic toggle switch, when a low amount of RFP will be produced by the activation of P_{SalI} promoter. This state is altered in the presence of tetracycline molecules which binds to TetR. Consequently, the affinity between TetR and P_{Tet} reduces, resulting in the production of salicylic acid repressor proteins and GFP. Subsequently, the salicylic acid repressor proteins repress the P_{SalI} promoter, preventing further production of

TetR proteins. This results in a change in the state of toggle switch to state 2 as shown in Fig. 5B. In order to go back to state 1, salicylic acid must be detected by the sensor to prevent the repression of P_{SalI} ; thus, increasing the production of TetR and RFP. Production of red fluorescent signals, therefore, indicates salicylic acid detection, while tetracycline presence is indicated by green fluorescent signals.

In this model, the protein degradation rates for GFP and RFP are assumed to be much higher than the degradation rates of TetR and salicylic acid repressor proteins. This assumption was made to ensure that the retention time of the TetR and salicylic acid repressor molecules would be much higher in comparison to GFP and RFP. Therefore, sufficient amounts of relevant repressor proteins would be ever-present in the system to bind with the inducers. More research is needed to identify the salicylic acid and atrazine repressor proteins to develop the correct parts to incorporate salicylic acid and atrazine detection in the toggle switch.

Simulation and Analysis of Biosensor Models

1. Ternion biosensor

The ODE simulation results (Figure 6) of the ternion biosensor model are shown in Fig. 6A. Initially, the biosensor was at 'off' state as no contaminant was present in the system. Tetracycline was introduced to the system at 1,000 seconds. This gave rise to a sharp increase in the concentration of YFP which plateaued at a value of 60 molecules. Once tetracycline was removed from the system, the concentration of YFP decreased to 0 molecules, showing the dependency of YFP production on tetracycline levels. At 1,100 seconds a constant influx of atrazine was introduced into the system and was maintained for a period of 1,000 seconds. This resulted in a sharp rise in the concentration of GFP which eventually plateaued at 1,750 seconds with a concentration of 250 molecules. Once atrazine was removed from the system, the GFP protein levels began to diminish over time. This illustrates the proportionality of the GFP output to atrazine levels. The introduction of salicylic acid into the system at 1,200 seconds gave a similar response to that of atrazine. Once salicylic acid was introduced into the system, RFP was rapidly produced and started to level out and reach steady state around 1,875 seconds also at a concentration of 250 molecules. Once salicylic acid was removed, the RFP proteins began to diminish over time due to its degrading nature. This sensor response illustrates the proportionality of the RFP output to the salicylic acid levels. The simulation

results for the atrazine and salicylic acid sensors showed an equal amount of their respective reporter proteins being produced, however, the tetracycline sensor model produced much less of its reporter protein (Fig. 6A). This discrepancy may be due to the amount of TetR present in the system, brought about by the constitutive promoter BBa_J23106, repressing P_{Tet} . This is an indication of the different output responses which may occur due to the mechanisms by which the reporter proteins are produced.

2. Triple-input-AND-gated biosensor

Fig. 6B shows the simulation results for the triple-input-AND-gated biosensor model. As the biosensor produces only a single green fluorescent output from the detection of three CECs, no GFP production initially indicated at its 'off' state. Despite the introduction of tetracycline and salicylic acid at 1,000 seconds and 1,500 seconds, respectively in the system, GFP was not produced due to the AND-logic of the circuit was not satisfied. However, the addition of atrazine at 2,000 seconds, along with the presence of tetracycline and salicylic acid into the system allowed the transcription process to begin for the production of GFP. Thus, the presence of all three CECs resulted in a rapid production of GFP, which reached a maximum concentration of 600 molecules before decreasing steadily. At 3,000 seconds, atrazine was removed from the system, which resulted in an instant reduction of GFP production and the green fluorescent signals. This response behaviour supports the logic character of the biosensor, as GFP is only produced in the presence of all three contaminants in the system.

3. Toggle switch biosensor

The genetic toggle switch biosensor behaviour in response to CECs is shown in Fig. 6C. There was an initial rise of the concentration of GFP and RFP to 6 molecules before decreasing back to 0 molecules at 1,000 seconds was observed. This initial rise of reporter proteins occurred as a result of setting the promoter counts of P_{Tet} and P_{SalI} to be greater than zero to ensure that the toggle switch was capable of detecting either tetracycline or salicylic acid; thus, resembling a NOR-logic gate. Fig. 6C also shows a rapid increase of GFP, reaching a maximum of 67 molecules, when 50 molecules of tetracycline were added to the system at 1,000 seconds. This is caused by the complex formation reaction that takes place between the TetR proteins and tetracycline molecules. Once tetracycline leaves the system at 1,500 seconds, the GFP output rapidly decreases. Following this, salicylic acid was introduced into the system at 2,000 seconds. Once salicylic acid is in the system, it begins to

bind with the salicylic acid repressor proteins available, which consequently permits the transcription for RFP, producing a maximum molecular count of 57 RFP molecules. When salicylic acid is removed from the system at 2,500 seconds, the RFP diminishes over time. Although equal quantities of tetracycline and salicylic acid molecules were added, more GFP molecules were produced as compared to RFP molecules. This difference is due to the repression of P_{Sal1} by the traces of salicylic acid repressors already present in the system when salicylic acid was added.

The simulations of the biosensor models were accomplished using ODE analysis. This type of analysis makes two significant assumptions. One being that the reactions that are occurring discretely, are assumed to occur continually instead [50]. The other being that the concentration of molecules present in the system varies [50]. However, these assumptions vary greatly to what actually occurs in genetic circuits [50]. For this reason, the simulation results of the biosensor models may not fully represent the results that would occur in experimental testing and implementation of the biosensors. Thus, a better representation of experimental results would be possible through the implementation of stochastic chemical kinetics in biosensors models.

Factors to consider for experimental testing and field use of biosensors

Despite all the simulation results indicating the viability of the various biosensor designs, further experimental testing is needed to increase their reliability. In order to conduct experimental testing on these biosensors, several factors need to be considered. In the whole-cell biosensors, it is recommended that the cells used are of the prokaryotic nature and are derived from the bacteria, *Escherichia coli*. *E. coli* variants have been extensively used in previous experimental implementation of biosensors [51]–[53]. Temperature and pH levels are also important factors that need to be considered. The simulated biosensor models did not consider optimum operating conditions for the biosensors to detect various contaminants. Although whole-cell biosensors are able to operate over a broad range of temperatures and pH levels, the optimum conditions can only be obtained through their experimental testing [54]. In conjunction with temperature and pH levels, horizontal gene transfer (HGT) is also an important factor which needs to be considered during experimental testing of the designed biosensors. HGT is interpreted as the movement of genetic material across organisms and has been shown to occur in *E. coli* strains [55], [56]. Consequently, the biosensors may be

susceptible to HGT; thereby, affecting their competence and test results. HGT may also lead to serious environmental and health issues.

Another important factor which needs to be considered is the undesired crosstalk between cells. Biological crosstalk refers to an instance where one or more components of one signal transduction pathway affect another [57]. Crosstalk may occur through the overexpression of particular proteins [58]. Additionally, crosstalk may occur through the overexpression of an unnatural response regulator [58]. Ninfa *et al* [59] observed the overexpression of a mutant version of nitrogen regulator II (NRII) in *E. coli*. This resulted in the phosphorylation of CheY, a non-cognate phosphorylation receptor and undesired interactions through which the swimming behaviour of *E. coli* was affected [59]. Effects of crosstalk may be minimised through the use of automation [60]. Additional challenges such as the time required to generate signals by reporter genes may arise when employing logic gated biosensors in the field. When compared to the ternion biosensor, the triple-input-AND-gated biosensor has more genetic circuit layers, thus, may require more time to transcribe the reporter gene further downstream. Moreover, the diffusion of TetR across the cell consortium may delay the production of the reporter proteins. Due to the structural intricacy of two-component systems, these sensors could also fail for reasons beyond present knowledge [58]. Furthermore, the stability of logic gated biosensors may give rise to problems in the field use. Thus, experimental testing is required to examine the stability of this type of biosensors. The mechanism for the regulation of P_{hrpL} promoter by *hrpS* and *hrpR* is not yet fully understood [48]. Hence, research in this area will permit the development of techniques to increase the efficiency of promoter activation.

Although iBioSim allowed us to model and simulate different biosensor designs, it lacks the ability to predict their experimental responses and behaviours. Also, the stability of the genetic circuits in the biosensor model is currently unexplainable through iBioSim. Furthermore, the promoter strength and specificity of genetic circuits are not considered in the iBioSim model simulations. Thus, to better indicate experimental results, software with the ability to integrate models and the phenotypic behaviour of the circuits must be developed [50].

Conclusion

Recent advances in synthetic biology have increased the potential for groundwater contamination detection through the use of whole-cell biosensors. One such advancement is the development of genetic circuits using standard biological parts that can perform various logic operations and regulatory functions using biological cells. This advancement is further aided by the development of computational synthetic biology tools. The availability of GDA software such as iBioSim now allows us to develop and analyse various design models of genetic circuits without any experimental implementation in a living system. Using iBioSim, three whole-cell biosensor models were designed and analysed to detect multiple groundwater CECs simultaneously. The viability of these biosensors has also been analysed and presented using mathematical models and simulation results. Despite these promising results, current use of whole-cell biosensors in environmental monitoring and field applications is significantly low. This is partly due to the lack of predictability and optimum operating conditions of whole-cell biosensors in field settings. Research gaps in transcriptional regulation of biosensor parts also need to be addressed to increase the predictability of these biosensors. Furthermore, computational synthetic biologists need to develop tools, which can imitate real life environments for more relevant simulations of biosensor models. With this framework in place, more reliable and robust whole-cell biosensor designs can be achieved for the detection of CECs in the future.

References

- [1] BGS, "Current UK groundwater use | Groundwater resources in the UK | Water resources and groundwater | British Geological Survey (BGS)," <https://www.bgs.ac.uk/research/groundwater/waterResources/GroundwaterInUK/2015.html>, accessed July 2020
- [2] P. D. Hynds, M. K. Thomas, and K. D. M. Pintar, "Contamination of groundwater systems in the US and Canada by enteric pathogens, 1990-2013: a review and pooled-analysis," *PLoS ONE*, vol. 9, no. 5. 2014, doi: 10.1371/journal.pone.0093301.
- [3] World Bank, "Deep wells and prudence: towards pragmatic action for addressing groundwater overexploitation in India," *Int. Bank Reconstr. Dev. World Bank*, p. 120, 2010, doi: 9504.
- [4] C. Postigo, D. E. Martinez, S. Grondona, and K. S. B. Miglioranza, "Groundwater pollution: sources, mechanisms, and prevention," in *Encyclopedia of the Anthropocene*, 2018, pp. 87–96.

- 460 [5] I. V Muralikrishna and V. Manickam, "Chapter 11 - principles and design of water
461 treatment," in *Environmental Management*, 2017, pp. 209–248.
- 462 [6] M. J. C. G. D. Todd, O. M. Faroon, D. E. Jones, M. H. Lumpkin, J. Sticney,
463 "Toxicological profile for vinyl chloride," no. July, pp. 1–329, 2006, doi:
464 <http://www.atsdr.cdc.gov/toxprofiles/tp20.pdf>.
- 465 [7] C. H. Tseng, C. P. Tseng, H. Y. Chiou, Y. M. Hsueh, C. K. Chong, and C. J. Chen,
466 "Epidemiologic evidence of diabetogenic effect of arsenic," *Toxicol. Lett.*, vol. 133,
467 no. 1, pp. 69–76, 2002, doi: 10.1016/S0378-4274(02)00085-1.
- 468 [8] H. Ahmad Khan, "Benzene's toxicity: a consolidated short review of human and
469 animal studies," *Hum. Exp. Toxicol.*, vol. 26, no. 9, pp. 677–685, 2007, [Online].
470 Available: <http://10.0.4.153/0960327107083974>.
- 471 [9] M. Stuart, "Emerging contaminants briefing," p. 2011, 2011.
- 472 [10] K. Demnerová, M. Gavrilescu, J. Aamand, F. Fava, and S. Agathos, "Emerging
473 pollutants in the environment: present and future challenges in biomonitoring,
474 ecological risks and bioremediation," *N. Biotechnol.*, vol. 32, no. 1, pp. 147–156,
475 2014, doi: 10.1016/j.nbt.2014.01.001.
- 476 [11] M. Raghav, S. Eden, K. Mitchell, and B. Witte, "Contaminants of emerging concern in
477 water," *Water Resour. Res. Cent. Coll. Agric. Life Sci. Univ. Arizona*, pp. 1–12, 2013.
- 478 [12] F. Riva, S. Castiglioni, E. Fattore, A. Manenti, E. Davoli, and E. Zuccato, "Monitoring
479 emerging contaminants in the drinking water of Milan and assessment of the human
480 risk," *Int. J. Hyg. Environ. Health*, vol. 221, no. 3, pp. 451–457, 2018, doi:
481 10.1016/j.ijheh.2018.01.008.
- 482 [13] C. Barrios-Estrada, M. de Jesús Rostro-Alanis, B. D. Muñoz-Gutiérrez, H. M. N.
483 Iqbal, S. Kannan, and R. Parra-Saldívar, "Emergent contaminants: endocrine
484 disruptors and their laccase-assisted degradation – a review," *Science of the Total
485 Environment*, vol. 612, pp. 1516–1531, 2018, doi: 10.1016/j.scitotenv.2017.09.013.
- 486 [14] S. D. Richardson, "Water analysis: Emerging contaminants and current issues,"
487 *Analytical Chemistry*, vol. 81, no. 20, p. 8654, 2009, doi: 10.1021/ac9020648.
- 488 [15] F. P. Carvalho, "Pesticides, environment, and food safety," *Food and Energy Security*,
489 vol. 6, no. 2, pp. 48–60, 2017, doi: 10.1002/fes3.108.
- 490 [16] M. Petrović, S. Gonzalez, and D. Barceló, "Analysis and removal of emerging
491 contaminants in wastewater and drinking water," *TrAC - Trends Anal. Chem.*, vol. 22,
492 no. 10, pp. 685–696, 2003, doi: 10.1016/S0165-9936(03)01105-1.
- 493 [17] K. T. Eriksen *et al.*, "Perfluorooctanoate and perfluorooctanesulfonate plasma levels
494 and risk of cancer in the general danish population," *J. Natl. Cancer Inst.*, vol. 101, no.
495 8, pp. 605–609, 2009, doi: 10.1093/jnci/djp041.
- 496 [18] K. Waller, S. H. Swan, G. DeLorenze, and B. Hopkins, "Trihalomethanes in drinking
497 water and spontaneous abortion," *Epidemiology*, vol. 9, no. 2, pp. 134–140, 1998, doi:
498 10.1097/00001648-199803000-00006.
- 499 [19] V. M. Fomenko, O. A. Shushakov, and V. S. Kuskovskii, "Detection of groundwater
500 contamination with hydrocarbons," *Russ. Geol. Geophys.*, vol. 49, no. 3, pp. 183–186,
501 2008, doi: 10.1016/j.rgg.2008.02.003.
- 502 [20] W. A. Chudyk, M. M. Carrabba, and J. E. Kenny, "Remote detection of groundwater
503 contaminants using far-ultraviolet laser-induced fluorescence," *Anal. Chem.*, vol. 57,
504 no. 7, pp. 1237–1242, 1985, doi: 10.1021/ac00284a017.

- [21] D. Grieshaber, R. MacKenzie, J. Vörös, and E. Reimhult, “Electrochemical biosensors - sensor principles and architectures,” *Sensors*, vol. 8, no. January, pp. 1400–1458, 2008, doi: 10.3390/s8031400.
- [22] N. Joshi, X. Wang, L. Montgomery, A. Elfick, and C. E. French, “Novel approaches to biosensors for detection of arsenic in drinking water,” *Desalination*, vol. 248, no. 1–3, pp. 517–523, 2009, doi: 10.1016/j.desal.2008.05.096.
- [23] F. Behzadian, H. Barjeste, S. Hosseinkhani, and A. R. Zarei, “Construction and characterization of Escherichia coli whole-cell biosensors for toluene and related compounds,” *Curr. Microbiol.*, vol. 62, no. 2, pp. 690–696, 2011, doi: 10.1007/s00284-010-9764-5.
- [24] R. Tecon *et al.*, “Development of a multistrain bacterial bioreporter platform for the monitoring of hydrocarbon contaminants in marine environments,” *Environ. Sci. Technol.*, vol. 44, no. 3, pp. 1049–1055, 2010, doi: 10.1021/es902849w.
- [25] G. Luka *et al.*, “Microfluidics integrated biosensors: A leading technology towards lab-on-A-chip and sensing applications,” *Sensors (Switzerland)*, vol. 15, no. 12, pp. 30011–30031, 2015, doi: 10.3390/s151229783.
- [26] S. P. Mohanty and E. Kougiannos, “Biosensors: a tutorial review - North Texas University,” no. i, pp. 1–12, 2007.
- [27] P. Shapira, S. Kwon, and J. Youtie, “Tracking the emergence of synthetic biology,” *Scientometrics*, vol. 112, no. 3, pp. 1439–1469, 2017, doi: 10.1007/s11192-017-2452-5.
- [28] X. Liu, K. J. Germaine, D. Ryan, and D. N. Dowling, “Whole-cell fluorescent biosensors for bioavailability and biodegradation of polychlorinated biphenyls,” *Sensors*, vol. 10, no. 2, pp. 1377–1398, 2010, doi: 10.3390/s100201377.
- [29] B. Wang, M. Barahona, and M. Buck, “A modular cell-based biosensor using engineered genetic logic circuits to detect and integrate multiple environmental signals,” *Biosens. Bioelectron.*, vol. 40, no. 1, pp. 368–376, 2013, doi: 10.1016/j.bios.2012.08.011.
- [30] T. Miyamoto, S. Razavi, R. DeRose, and T. Inoue, “NIH Public Access,” *ACS Synth. Biol.*, vol. 2, no. 2, pp. 72–82, 2014, doi: 10.1021/sb3001112.Synthesizing.
- [31] Q. Gui, 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 (Switzerland)*, vol. 17, no. 7, pp. 1–17, 2017, doi: 10.3390/s17071623.
- [32] L. Bousse, “Whole cell biosensors,” *Sensors Actuators, B Chem.*, vol. 34, no. 1–3, pp. 270–275, 1996, doi: 10.1016/S0925-4005(96)01906-5.
- [33] X. Liu, K. J. Germaine, D. Ryan, and D. N. Dowling, “Development of a gfp-based biosensor for detecting the bioavailability and biodegradation of polychlorinated biphenyls (pcbs),” *J. Environ. Eng. Landsc. Manag.*, vol. 15, no. 4, pp. 261–268, 2007, doi: 10.1080/16486897.2007.9636939.
- [34] L. T. Bereza-Malcolm, G. Mann, and A. E. Franks, “Environmental sensing of heavy metals through whole cell microbial biosensors: a synthetic biology approach,” *ACS Synth. Biol.*, vol. 4, no. 5, pp. 535–546, 2015, doi: 10.1021/sb500286r.
- [35] C. Madsen, C. J. Myers, T. Patterson, N. Roehner, J. T. Stevens, and C. Winstead, “Design and test of genetic circuits using iBioSim,” *IEEE Des. Test Comput.*, vol. 29, no. 3, pp. 32–39, 2012, doi: 10.1109/MDT.2012.2187875.
- [36] N. Roehner and C. J. Myers, “Directed acyclic graph-based technology mapping of

genetic circuit models,” *ACS Synth. Biol.*, vol. 3, no. 8, pp. 543–555, 2014, doi: 10.1021/sb400135t.

[37] L. Watanabe *et al.*, “IBioSim 3: a tool for model-based genetic circuit design,” *ACS Synthetic Biology*, 2018.

[38] A. Kanhere and M. Bansal, “Structural properties of promoters: similarities and differences between prokaryotes and eukaryotes,” *Nucleic Acids Res.*, vol. 33, no. 10, pp. 3165–3175, 2005, doi: 10.1093/nar/gki627.

[39] H. M. Salis, E. A. Mirsky, and C. A. Voigt, “Automated design of synthetic ribosome binding sites to control protein expression,” *Nat. Biotechnol.*, vol. 27, no. 10, pp. 946–950, 2009, doi: 10.1038/nbt.1568.

[40] R. L. Miller and T. Ohkawa, “Chain terminator sequencing of double-stranded DNA with built-in error correction,” *J. Theor. Biol.*, vol. 161, no. 4, pp. 407–429, 1993, doi: 10.1006/jtbi.1993.1065.

[41] J. A. McLaughlin *et al.*, “SynBioHub: a standards-enabled design repository for synthetic biology,” *ACS Synth. Biol.*, 2018, doi: 10.1021/acssynbio.7b00403.

[42] R. P. Shetty, D. Endy, and T. Knight, “The BioBricks foundation: RFC,” *OpenWetWare.org*. 2009.

[43] M. Galdzicki *et al.*, “The synthetic biology open language (SBOL) provides a community standard for communicating designs in synthetic biology,” *Nature Biotechnology*, vol. 32, no. 6, pp. 545–550, 2014, doi: 10.1038/nbt.2891.

[44] R. S. Cox *et al.*, “Synthetic biology open language visual (SBOL Visual) version 2.0,” *J. Integr. Bioinform.*, vol. 15, no. 1, 2018, doi: 10.1515/jib-2017-0074.

[45] N. Csicsery and R. O. Laughlin, “BENG 221 – Mathematical methods in bioengineering: a mathematical model of a synthetically constructed genetic toggle switch,” pp. 1–29, 2013.

[46] P. Waage and C. M. Gulberg, “Studies concerning affinity,” *J. Chem. Educ.*, vol. 63, no. 12, p. 1044, 2009, doi: 10.1021/ed063p1044.

[47] U. Alon, “An introduction to systems biology: design principles of biological circuits,” *Chapman HallCRC Math. Comput. Biol. Ser.*, vol. 10, no. 10, p. 301, 2007, doi: citeulike-article-id:1314150.

[48] S. W. Hutcheson, J. Bretz, T. Sussan, S. Jin, and K. Pak, “Enhancer-binding proteins HrpR and HrpS interact to regulate hrp-encoded type III protein secretion in *Pseudomonas syringae* strains,” *J. Bacteriol.*, vol. 183, no. 19, pp. 5589–5598, 2001, doi: 10.1128/JB.183.19.5589-5598.2001.

[49] X. Zhang *et al.*, “The TetR family of transcriptional repressors,” *Microbiol. Mol. Biol. Rev.*, vol. 69, no. 2, pp. 326–356, 2005, doi: 10.1128/mmbr.69.2.326-356.2005.

[50] C. J. Myers, “Computational synthetic biology: progress and the road ahead,” *IEEE Trans. Multi-Scale Comput. Syst.*, vol. 1, no. 1, pp. 19–32, 2015, doi: 10.1109/TMSCS.2015.2478442.

[51] J. Cai and M. S. DuBow, “Use of a luminescent bacterial biosensor for biomonitoring and characterization of arsenic toxicity of chromated copper arsenate (CCA),” *Biodegradation*, vol. 8, no. 2, pp. 105–111, 1997, doi: 10.1023/A:1008281028594.

[52] B. Davidson *et al.*, “Development of a novel biosensor for the detection of arsenic in drinking water,” *IET Synth. Biol.*, vol. 1, no. 1, pp. 87–90, 2007, doi: 10.1049/iet-stb:20060002.

- [53] H. C. Flynn *et al.*, “Assessment of bioavailable arsenic and copper in soils and sediments from the Antofagasta region of northern Chile,” *Sci. Total Environ.*, vol. 286, no. 1–3, pp. 51–59, 2002, doi: 10.1016/S0048-9697(01)00962-7.
- [54] R. M. Maier, “Bacterial Growth,” *Environ. Microbiol.*, pp. 37–54, 2009, doi: 10.1016/B978-0-12-370519-8.00003-1.
- [55] P. J. Keeling and J. D. Palmer, “Horizontal gene transfer in eukaryotic evolution,” *Nature Reviews Genetics*, vol. 9, no. 8, pp. 605–618, 2008, doi: 10.1038/nrg2386.
- [56] M. Javadi, S. Bouzari, and M. Oloomi, “Horizontal gene transfer and the diversity of *Escherichia coli*,” in *Escherichia coli - Recent Advances on Physiology, Pathogenesis and Biotechnological Applications*, 2017.
- [57] G. Vert and J. Chory, “Crosstalk in cellular signaling: background noise or the real thing?,” *Dev. Cell*, vol. 21, no. 6, p. 985, 2011, doi: 10.1016/j.devcel.2011.11.006.
- [58] F. Zhang and J. Keasling, “Biosensors and their applications in microbial metabolic engineering,” *Trends Microbiol.*, vol. 19, no. 7, pp. 323–329, 2011, doi: 10.1016/j.tim.2011.05.003.
- [59] A. J. Ninfa, E. G. Ninfa, A. N. Lupas, A. Stock, B. Magasanik, and J. Stock, “Crosstalk between bacterial chemotaxis signal transduction proteins and regulators of transcription of the Ntr regulon: evidence that nitrogen assimilation and chemotaxis are controlled by a common phosphotransfer mechanism,” *Proc. Natl. Acad. Sci.*, vol. 85, pp. 5492–5496, 1988, [Online]. <http://www.pnas.org/content/85/15/5492>.
- [60] L. Goers, C. Ainsworth, C. H. Goey, C. Kontoravdi, P. S. Freemont, and K. M. Polizzi, “Whole-cell *Escherichia coli* lactate biosensor for monitoring mammalian cell cultures during biopharmaceutical production,” *Biotechnol. Bioeng.*, vol. 114, no. 6, pp. 1290–1300, 2017, doi: 10.1002/bit.26254.

FIGURES

Figure 1. SBOL Visual design of a genetic circuit to detect salicylic acid. P_{SalI} (BBa_J61051) is the promoter, RBS30 (BBa_K1789002) is the ribosome binding site, mCherry/RFP (BBa_I766101) is the reporter protein, and BBa_B0015 is the terminator.

A

B

C

Figure 2. iBioSim models of the designed genetic circuits. (A) iBioSim model of the ternion biosensor circuit. The model includes eight species, four sub-models, six events, and

663 two complex formation reactions. (B) iBioSim model of the triple-input-AND-gated
 664 biosensor circuit. This model has nine species, one sub-model, four complex formation
 665 reactions, four events, and three grey SBOL parts. (C) iBioSim model of the toggle switch
 666 biosensor circuit. This model includes eight species, two sub-models, four complex formation
 667 reactions, and four events. Species, sub-models, complex formation reactions, and events are
 668 represented by blue ovals, light green squares, dark green rectangles, and dashed arrows,
 669 respectively.
 670

Figure 3. Schematic of the genetic circuits for the ternion biosensor model. The model contains the atrazine, salicylic acid, and tetracycline sensors shown in (A), (B), and (C) to generate green, red, and yellow signals by producing green fluorescent protein (GFP), red fluorescent protein (RFP), and yellow fluorescent protein (YFP), respectively.

Figure 4. Schematic of the genetic circuits for the triple-input-AND-gated biosensor model. Two specific input signals (tetracycline and salicylic acid) are detected by the cell consortium 1 releasing TetR as an output. The cell consortium 2 detects the TetR released from the cell consortium 1 whilst also detecting atrazine to produce a GFP output.

Figure 5. Schematic of the genetic circuits for the toggle switch biosensor model. (A) shows state one of the genetic toggle switch, in which TetR and RFP are produced. TetR represses P_{Tet} promoter preventing production of salicylic acid repressor gene and GFP. (B) shows state two of the genetic toggle switch, in which salicylic acid repressors and GFP are

689 produced. Salicylic acid repressor represses the P_{SalI} promoter preventing further production
690 of TetR and RFP.
691

A

B

C

Figure 6. ODE simulation of biosensor models. (A) ODE simulation results of the ternion biosensor, (B) ODE simulation results of the triple-input-AND-gated biosensor, and (C) ODE simulation results for the toggle switch biosensor.

Table 1: Summary of standard biological parts used to design different biosensor circuits

Biosensor circuits	Biosensor circuit part	Genetic identifier of individual parts
Ternion biosensor	Atrazine Sensor	BBa_J23115- BBa_K1442030- BBa_K2027047- BBa_K1789002- BBa_E0040- BBa_B0015
	Tetracycline Sensor	BBa_J23106- BBa_B0034- TetR- BBa_B0015- BBa_R0040- BBa_B0034- BBa_K592101- BBa_B0015
	Salicylic Acid Sensor	BBa_J61051- BBa_K1789002- BBa_I766101- BBa_B0015
Triple-input-AND-gated biosensor	Atrazine Sensor	BBa_R0040- BBa_K1442030- BBa_K2027047- BBa_K1789002- BBa_E0040- BBa_B0015
	Tetracycline Sensor	BBa_J23106- BBa_B0034- TetR- BBa_B0015- BBa_R0040- BBa_B0034- BBa_K1014001- BBa_B0015
	Salicylic Acid Sensor	BBa_J61051- BBa_K1789002- BBa_K1014001- BBa_B0015
	hrpL Sensor	BBa_K1014002- BBa_B0034- TetR- BBa_B0015
Genetic toggle switch biosensor	Tetracycline Sensor	BBa_R0040- BBa_B0034- BBa_I732100- BBa_B0034- BBa_E0040- BBa_B0015
	Salicylic Acid Sensor	BBa_K1638000- BBa_B0034- TetR- BBa_B0034- BBa_I766101- BBa_B0015