Orthogonal Regulatory Circuits for Escherichia coli Based on the gamma-Butyrolactone System of Streptomyces coelicolor

DOI: 10.1021/acssynbio.7b00425

Citation for published version (APA):

Published in:
ACS Synthetic Biology

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Orthogonal regulatory circuits for *E. coli* based on the γ-butyrolactone system of *Streptomyces coelicolor*

Author list: Marc Biarnes-Carrera¹, Chang-Kwon Lee²,⁴, Takuya Nihira²,³, Rainer Breitling¹ and Eriko Takano¹

Affiliation:

¹Manchester Centre for Synthetic Biology of Fine and Speciality Chemicals (SYNBIOCHEM), Manchester Institute of Biotechnology, School of Chemistry, Faculty of Science and Engineering, University of Manchester, 131 Princess Street, Manchester M1 7DN, United Kingdom.

²International Center for Biotechnology, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

³Mahidol University-Osaka University Collaborative Research Center for Bioscience and Biotechnology, Faculty of Science, Mahidol University, Rama VI Rd., Bangkok 10400, Thailand

⁴Present address: Mong-Go Food Co., Ltd., Changwon 641-847, Republic of Korea

Corresponding author: Eriko Takano, eriko.takano@manchester.ac.uk

Graphical Abstract

Abstract

Chemically inducible transcription factors are widely used to control gene expression of synthetic devices. The bacterial quorum sensing system is a popular tool to achieve such control. However, different quorum sensing systems have been found to cross-talk, both
between themselves and with the hosts of these devices, and they are leaky by nature. Here we evaluate the potential use of the γ-butyrolactone system from *Streptomyces coelicolor* A3(2) M145 as a complementary regulatory circuit. First, two additional genes responsible for the biosynthesis of γ-butyrolactones were identified in *S. coelicolor* M145 and then expressed in *E. coli* BL21 under various experimental conditions. Second, the γ-butyrolactone receptor ScbR was optimised for expression in *E. coli* BL21. Finally, signal and promoter crosstalk between the γ-butyrolactone system from *S. coelicolor* and quorum sensing systems from *Vibrio fischeri* and *Pseudomonas aeruginosa* was evaluated. The results show that the γ-butyrolactone system does not crosstalk with the quorum sensing systems and can be used to generate orthogonal synthetic circuits.

**Keywords**

γ-butyrolactone; *Streptomyces*; synthetic biology; orthogonal regulatory circuit; quorum sensing

Synthetic biology aims at the rational design of living organisms with novel functionalities. One major challenge is to develop new regulatory circuits that allow for tight regulation over a wide variety of conditions. Various tools are available to control transcript levels, translation rates or protein concentration. One popular tool is the use of chemically-inducible transcription factors to regulate the expression of genes of interest, such as the acyl-homoserine lactones (AHL) from the bacterial quorum sensing (QS) system originally derived from the microorganism *Vibrio fischerii*. This system has been previously used to generate devices such as oscillators or logic gates and is currently being seen as a promising tool to engineer microbial consortia, mainly due to the large diversity of well-characterised AHL systems reported. However, despite these promising applications and perspectives, different AHL devices have been found to cross-talk, either at the promoter or the signal level; a limitation that can result in undesirable or unpredictable outcomes. Furthermore, some AHL systems control the expression of virulence factors, such as the one from *Pseudomonas aeruginosa*, and consequently organisms have evolved to target these molecules using so-called quorum quenchers, which function as either lactonases or acylases. This can limit the application of QS systems in some hosts, such as mammalian cells. Finally, the promoter *luxp*, which drives expression of the output
signal in the AHL-based synthetic circuits, is leaky \(^{19}\), resulting in sometimes undesired high basal expression of the output signal.

Members of the genus *Streptomyces* have been widely researched due to their ability to produce a vast array of secondary metabolites, many of which have medical interest \(^{20}\). In *Streptomyces*, cell-to-cell communication is crucial to coordinate the onset of antibiotic production, as the final compound would be toxic to cells that have not developed the corresponding resistance. This coordination is achieved through the use of small diffusible molecules known as \(\gamma\)-butyrolactones (GBLs) \(^{21}\). These bacterial hormones are species-specific \(^{22}\), and although some recent publications \(^{23,24}\) have shown that different *Streptomyces* species can share a common GBL, currently there is no evidence of signal cross-talk between different GBL regulators at physiological concentrations \(^{25}\). In *Streptomyces*, GBLs promote a growth phase-dependent switch-like transition to antibiotic production by binding to their cognate receptor protein, a homodimer from the TetR family of repressors. These repressors are the master regulators of biosynthetic gene clusters and in some instances \(^{26-28}\) they also regulate their own transcription and that of a GBL synthase protein.

In this study, we evaluated the potential application of the GBL system from *S. coelicolor* as a regulatory tool for synthetic biology in *E. coli*. We show that this system can be used as a versatile and accessible tool to activate gene circuits in the heterologous host, with minimal crosstalk with the QS systems from *V. fischeri* and *P. aeruginosa*, and that it allows for tight control of genes of interest. To achieve this, we have generated a plasmid for production of GBL SCB2 in *E. coli* and a series of plasmids constitutively expressing an optimised version of ScbR and a library of *Streptomyces*-derived promoters: *scbRp, scbAp* or *cpkOp*, that are controlling the expression of *gfp*. 
**Results and Discussion**

*a) The SCO6264 and SCO6267 genes from S. coelicolor are essential for GBL production*

In *Streptomyces*, the biosynthesis of GBLs is reported to start with a condensation between dihydroxyacetone phosphate (DHAP) and a β-ketoacyl-acyl carrier protein (ACP) \(^{29}\), followed by an intramolecular aldol condensation that yields the corresponding butenolide (Fig. 1, compound 4). In *S. coelicolor*, the enzyme responsible for this catalytic step is known as ScbA \(^{26,32}\). Previous *in vitro* studies with homologues from *Streptomyces griseus* \(^{29}\) and *Streptomyces virginiae* \(^{30}\) showed that after this first condensation, the butenolide is enzymatically reduced in an NADPH-dependent manner to generate the reduced GBL (compound 5). In some *Streptomyces* species, this compound undergoes a second enzymatic NADPH-dependent reduction that stereo-specifically reduces the carbonyl in C\(_6\) into an (R)- or (S)-alcohol \(^{30}\) (compound 7). However, to our knowledge, no homologues of these enzymes have been identified and characterised in *S. coelicolor*. Using a BLASTp search we identified SCO6264 (hereafter *scbB*, 31% amino acid identity to BarS1 \(^{30}\)) as a putative 3-ketoacyl-ACP/Coenzyme A (CoA) reductase from the short-chain alcohol dehydrogenase superfamily and SCO6267 (hereafter *scbC*, 76% amino acid identity to BprA \(^{29}\) and 45% amino acid identity to BarS2 \(^{33}\)) as a putative butenolide phosphate reductase.

To confirm the involvement of ScbB and ScbC in GBL biosynthesis, insertion mutants of *scbB* and *scbC*, LW107 and LW108, respectively, were generated in *S. coelicolor* M145. The mutant strains were grown in SFM solid medium, their GBLs extracted and analysed using the kanamycin bioassay \(^{34}\) (Fig. 2). In this assay, only spiking of ethyl acetate extracts resuspended in methanol from strains that produce GBLs will be able to allow growth of reporter strain LW94 in a kanamycin-supplemented medium.

Production of GBLs was detected in the wild-type *S. coelicolor* M145 at all tested extract volumes, using the kanamycin bioassay. However, the indicator strain failed to grow in presence of either LW107 (*scbA/C*) or LW108 (*scbA/B*) extracts, suggesting that these mutants are not able to produce the GBLs, confirming that both enzymes are involved in the production of *S. coelicolor* GBLs.

*b) Expression of the S. coelicolor GBL system in E. coli BL21 results in production of SCB2*

After identifying the two new genes involved in *S. coelicolor* GBL production, we constructed a GBL production biosynthesis pathway to confirm their roles and to have an
easy production platform of GBLs in *E. coli*. An expression plasmid containing *scbA*, *scbB* and *scbC* under the control of the TetR repressor was constructed. A different TetR-dependent promoter (*tetA* for *scbA*<sup>9</sup>, Pb10 for *scbB* and Pb19 for *scbC*<sup>62</sup>) was added in front of each gene, as initial attempts to express the whole operon from a single promoter (*tetA*) were unsuccessful (data not shown). To facilitate expression of *S. coelicolor* genes in *E. coli* BL21, the first ten amino acids of each gene were codon optimised, and were His-tagged.

Expression of ScbA and/or ScbC was detected in the soluble and insoluble protein fraction, but not ScbB (Supplementary Fig. S1). In the kanamycin assay (Supplementary Fig. S2), addition of the extract from heterologously expressed ScbA/B/C (pTE1059) in *E. coli* to the reporter strain induced growth. This extract was then analysed by HPLC-MS/MS (Fig. 3 and Supplementary Fig. S3 and Supplementary Table S4), using an ethyl acetate-methanol extract from *S. coelicolor* M145 as positive control. In the latter, two peaks are identified in the extracted ion chromatogram (EIC) from the SCB2+Na adduct (theoretical m/z = 267.156678) at a retention time of approximately 19 min after injection. These were assigned as the isomers SCB1 and SCB2, respectively<sup>31,34</sup>. In the extract from *E. coli*, a peak was detected at the same retention time as observed for SCB2 (mass accuracy 0.68 ppm). The identity of the compound identified in the *E. coli* extract was confirmed through tandem MS showing the same fragmentation pattern as the obtained for SCB1 and SCB2 from *S. coelicolor* M145 (Supplementary Fig. S3). None of the other GBLs produced in *S. coelicolor* (SCB1 and SCB3) were detected. This can be explained by the fact that the β-ketoacid used in the first catalytic step towards the production of GBLs derives from fatty acid biosynthesis<sup>29,31</sup>. Fatty acid biosynthesis starts with the decarboxylative condensation of an acyl-CoA primer with malonyl-acyl carrier protein, catalysed by the gene product of *fabH*<sup>35</sup>. The *S. coelicolor* FabH homologue preferentially uses branched acyl-CoA primers for fatty acid production, such as isovaleryl-CoA and isobutyryl-CoA or metylbutyryl-CoA<sup>36,37</sup>, which then results in production of both branched (SCB1 and SCB3) and linear (SCB2) GBLs. However, FabH homologue from *E. coli* has a strong preference of linear acyl-CoA primers, such as acetyl-CoA or malonyl-CoA<sup>38</sup>; thus, if the GBL precursors are derived from fatty acid biosynthesis, this would result in exclusive production of linear GBLs (SCB2) in *E. coli*.

To further confirm the role of ScbB and ScbC, single knock-out mutants were generated in the GBL expression plasmid, transformed and expressed in *E. coli*, and their extracts analysed as described previously (Fig. 3 and Supplementary Fig. S3 and Supplementary Table S4). In the extract from *E. coli* harbouring *scbA/C* plasmid (pTE1060), a peak was
detected in the EIC from the A-factor+Na adduct (theoretical m/z = 265.141028) at approximately 21 min after injection, the same as A-factor from *S. griseus* (1.93 ppm). In both extracts from *S. griseus* and *E. coli* carrying the *scbA/C* plasmid (pTE1060), an additional broader peak is observed at approximately 20 min after injection. Further analysis through tandem MS showed that both peaks contained fragments consistent with A-factor (Supplementary Fig. S3). In the extract from *E. coli* carrying *scbA/B* plasmid (pTE1061), contrary to expectations, no peaks were detected at a mass corresponding to the non-reduced butenolide precursor from A-factor (compound 4), either phosphorylated or not (data not shown). However, an unidentified shunt metabolite was detected at a mass corresponding to A-factor+Na adduct but eluted at an earlier retention time than A-factor. These results show that by heterologous expression of the three GBL biosynthesis genes, GBLs with linear residue can be produced in *E. coli*.

c) *Production of SCB2 in E. coli BL21 is robust through different temperature and medium conditions*

Synthetic genetic circuits need to be robust under different experimental conditions to allow for predictable rational design of organisms \(^3\). Important parameters to consider are the temperature and the media conditions at which the synthetic organism is grown \(^39\). We tested the qualitative analysis of SCB2 production in *E. coli* at different temperatures and in four different media. Cells containing the *scbA/B/C* (pTE1059) plasmid were grown in either M9 minimal medium, LB medium, 2xYT medium or terrific broth (TB) medium and in LB at 20, 25, 30 and 37 °C. SCB2 production was assayed using the kanamycin bioassay. Production was readily detected at assayed temperatures from 20-30 °C; however, the GBL production was not detected with extracts from cells grown at 37 °C (Fig. 4). This is consistent with previous *in vitro* enzyme activity results for AfsA \(^29\) and BarS1 \(^30\), where a reduction in activity was found for both enzymes after incubation at and above 35 °C. This could be a potential disadvantage when using the GBL system in *E. coli*. However, the thermal stability and catalytic activity of the GBL biosynthetic enzymes could be achieved through random mutagenesis, as has been previously shown for other enzymes \(^40,41\). When *E. coli* harbouring *scbA/B/C* was expressed in different media, GBLs were detected in all of the tested conditions, both in minimal and rich media. Therefore, SCB2 production is robust in different experimental conditions.
d) Engineering of ScbR to promote solubility in E. coli BL21

In S. coelicolor, GBLs promote a switch-like transition to antibiotic production upon binding their cognate receptor. Therefore, after characterising the production of SCB2 in E. coli, we proceeded to characterise the GBL receptor, ScbR, in the heterologous host. Previous reports identified ScbR, and other GBL receptors, as prone to aggregate and form insoluble inclusion bodies when expressed in an E. coli background. Furthermore, we, and other groups, have attempted to crystallise GBL receptors. However, only the structures of distant homologues, which are not known to bind GBLs, have been resolved so far. To try to obtain a more soluble version of ScbR, we rationally engineered the structure, based on the available structural information. Using the SWISS-MODEL suite, a model structure of ScbR was generated, using CprB crystal structure (PDB 1UI5) as template, and visualised using the UCSF Chimera package. CprB is a TetR-like DNA-binding protein from S. coelicolor M145 with homology to GBL receptor proteins but unable to bind to the GBLs (pseudo-receptor). The hydrophobic surface of both CprB and ScbR (33% amino acid identity to CprB) was modelled and compared (Supplementary Fig. S5). As expected, the highly conserved N-terminus of the protein, corresponding to the DNA-binding site of the protein, was similar. However, a hydrophobic patch was identified at the C-terminal end of ScbR, which was not present in CprB. This hydrophobic patch was also not found in the pseudo-receptor ScbR2 (35% amino acid identity to CprB) (Supplementary Fig. S5). To suggest possible roles of this region it was further interrogated using a docking analysis, which suggested that it might contain a potential GBL binding site (Supplementary Fig. S6). This suggestion would be consistent with an earlier observation that showed that the addition of GBLs into a crude extract containing ArpA enhanced the solubility of the protein from inclusion bodies. We therefore hypothesized that GBL receptors might form inclusion bodies due to interaction through this hydrophobic, putatively the GBL-binding, region. Thus, to reduce the aggregation propensity, a 6x arginine tag (Arg6) was added at the C-terminus of the protein. Addition of this tag, either at the N- or the C-terminus, has been shown previously to enhance solubility of proteins without affecting their function.

Recombinant ScbR-Arg6 was expressed from plasmid ScbR-Arg6 + luxP (pTE1067) and its expression was compared to that of recombinant ScbR expressed from ScbR + luxP (pTE1066) by Western Blot analysis (Fig. 5). In this, ScbR-Arg6 in pTE1067 was seen to be more soluble compared to ScbR. To confirm whether the soluble ScbR-Arg6 is functional, a
gel shift assay using protein crude extract was performed (Fig. 5). This shows that addition of increasing amounts of ScbR-Arg₆ to DNA, results in appearance of two shifted bands which presumably correspond to complexes of one or two ScbR-Arg₆ homodimers bound to the DNA fragment containing operator sequence Oₐ (Fig. 5). This result is consistent with previous reports where ScbR binds its cognate operator sequence as a dimer of homodimers and previous reported gel shift assays using recombinant ScbR. Addition of SCB₂ to ScbR-Arg₆ resulted in release of the repressor from its cognate operator sequence (see below), whereas it resulted in high variable results when SCB₂ was added to untagged ScbR (Supplementary Fig. S7). These results suggest that addition of the Arg₆ tag at the C-terminal of ScbR results in less aggregation-prone protein while retaining DNA- and GBL-binding activity.

e) Crosstalk evaluation between the GBL and the AHL systems

As previously mentioned, a property that synthetic gene circuits aim to achieve is high orthogonality, that is, that two independent genetic devices do not crosstalk. It has been shown that, at physiological conditions, GBLs from *Streptomyces* only interact with their cognate receptors, although some crosstalk can be achieved at high GBL concentrations, around 200 times higher than the physiological concentration. However, to our knowledge, it has not been previously explored whether the GBL system is orthogonal to the more well-established signalling system based on the AHLs from *V. fisheri* and *P. aeruginosa*. These systems are based on the AHLs binding to LuxR-like proteins. Upon binding to their cognate AHLs, the LuxR receptors act as activators and are able to bind to the corresponding operator sequence (e.g., *lux* box) and recruit the RNA polymerase to induce transcription of a target gene (Fig. 6). On the other hand, the GBL system is based on the binding of the GBLs to the ScbR-like repressors. The ScbR-like receptor binds the operator sequence (e.g., *Oₐ* box) and represses the expression of the target gene. Once bound to cognate GBLs, a conformational change occurs to the receptor and can no longer bind to the operator sequence, which results in the activation of the target gene transcript.

Signal and promoter crosstalk were evaluated with plasmid vector BC-A1-002 and derivatives pTE1063 to pTE1069. These plasmids contain either *luxR* or *scbR*-Arg₆ under the control of a strong constitutive promoter and *gfp* under the control of the *lux* promoter (*luxp*) (BC-A1-002 and pTE1066), the *scbA* promoter (*scbAp*) (pTE1064 and pTE1069), the *scbR*
promoter \((scbRp)\) \(^{26}\) (pTE1063 and pTE1068) or the promoter of the activator of the coelomycin biosynthetic cluster \((cpkOp)\) \(^{56}\) (pTE1065 and pTE1070). The GFP expression for each construct was measured under different concentrations of 3-oxo-C\(_6\)-HSL (3OC\(_6\)), 3-oxo-C\(_{12}\)-HSL (3OC\(_{12}\)) and purified SCB2, obtained from \(E.\ coli\) supernatant as described in Materials and Methods (Supplementary Fig. S9) to measure the relationship between the input and the response of the synthetic regulatory circuit (Fig. 7, 8).

In strains with plasmids containing the \(luxR\), addition of AHLs, either 3OC\(_6\) or 3OC\(_{12}\), resulted in activation of the \(luxp\) at concentrations above \(10^{-11}\) M or \(10^{-10}\) M, respectively, as previously reported \(^{10,54}\), highlighting the cross-talk potential between the C6 and C12 AHL systems. However, addition of purified SCB2 only activated \(luxp\) at concentrations above \(10^{-6}\) M (\(10^5\) times higher concentration of the signal compared to 3OC\(_6\)). On the other hand, no GFP expression was detected upon addition of any of the three signalling molecules when \(luxR\) was exchanged for \(scbR\)-Arg\(_6\). Although we have no clear evidence, this suggests that ScbR-Arg\(_6\) does not bind to the operator sequence of LuxR. If ScbR acted as an inducer, one would expect to see a decrease in GFP expression after the addition of SCB2, and if ScbR acted as a repressor, one would expect to see an increase in GFP expression after addition of SCB2, assuming that in both cases the DNA-binding is responsive to the SCB2 signal. We see neither of these occurring.

The three ScbR-dependent promoters, \(scbRp\), \(scbAp\) and \(cpkOp\), from \(S.\ coelicolor\), were assayed to evaluate further the potential crosstalk with the AHL system (plasmids pTE1068 – pTE1070). When \(gfp\) is placed downstream of any of these promoters in plasmids containing \(luxR\), constitutive expression of GFP was observed, with or without the addition of AHLs or GBL signals. This suggests that the LuxR protein does not interact with any of the tested operator sequences for ScbR, and therefore the promoters are always active (Fig. 7, 8). The GFP expression also corresponded to the strength of the promoters which has been shown before \(^{56,57}\) (e.g., \(scbRp\) the strongest promoter, followed by \(cpkOp\) and \(scbAp\)). When \(luxR\) was exchanged for \(scbR\)-Arg\(_6\), the cells only express \(gfp\) at SCB2 concentrations starting from \(10^{-9}\) M, reaching a maximum of induction at \(10^{-7} - 10^{-6}\) M. Addition of more SCB2 beyond this concentration resulted in a decrease in \(gfp\) expression for \(scbRp\) or \(scbAp\) (Fig. 7, 8). The potential toxicity of SCB2 in \(E.\ coli\) was discarded as the responsible of this issue, as no decrease in fluorescence was seen in the strains with LuxR and those with \(cpkOp\) (Fig. 7, 8) and the final \(OD_{600}\) was independent of the concentration of SCB2 in all samples (Supplementary Fig. S8). This narrow range of activity of the GBL system on these two
promoters has been previously seen in *S. coelicolor*, where exogenous addition of SCB1 results in precocious antibiotic production only at a narrow range of concentrations. Further research on the factors that influence this apparent narrow range of GBL induction would be of great interest. Addition of either AHLs to plasmids containing *scbR-Arg6* resulted in no expression of *gfp*, suggesting that ScbR does not bind to the AHLs and therefore the AHL and GBL signals do not cross-talk at any concentration. These results show that the GBL system from *S. coelicolor* does not crosstalk with the LuxR system from *V. fischerii* and that neither the AHLs from *V. fischerii* nor those from *P. aeruginosa* bind to ScbR. Furthermore, the GBL system is shown to be highly tightly regulated, resulting in *gfp* basal expression 4-fold lower than for *luxp*. This could help create synthetic circuit devices where tight control is required, such as multicellular decision making or multi-state systems. Moreover, the response obtained from the plasmids containing *Streptomyces*-derived promoters (*scbRp*, *scbAp* and *cpkOp*) results in relatively low maximum expression of GFP, as opposed to the high levels of expression obtained from *luxp*, with an approximately 2-fold difference of relative strength between them (Fig. 8). In synthetic microorganisms, especially in those designed to produce speciality chemicals, fine tuning and balancing of all proteins in the biosynthetic pathway has been previously shown to play a crucial role in the titres produced, and overly strong induction of the proteins involved in the biosynthetic pathway can result in an impairment of the intracellular metabolite fluxes. The combination of the *Streptomyces*-derived promoters with the *luxp* can provide the means to balance and regulate refactored pathways.

**f) Induction of the GBL genetic circuits**

The purification steps to obtain SCB2 in order to activate gene circuits can be costly and time-consuming, and require specialised equipment that may not be accessible to all. To overcome this limitation of the proposed system, we examined whether addition of crude extract containing SCB2, without the previous purification steps, would result in activation of the circuit. GBL producer strain was grown in LB medium under different temperatures, the GBLs were extracted by a simple ethyl acetate extraction of the culture supernatant, resuspended in different volumes of methanol and assayed against a strain containing plasmid pTE1068 (Fig. 9). Expression of GFP was achieved with extracts from cells grown at 20 °C enriched 100-fold, and with extracts from cells grown at 25 and 30 °C enriched 500-fold.
These results show that signalling molecule can be obtained to activate circuits without any need of prior complicated purification steps (e.g., by use of HPLC), making it an accessible and versatile tool.

Interestingly, as methanol extracts are added in a 1:100 dilution to the culture with strains containing the reporter plasmid pTE1068, addition of extracts enriched 100-fold corresponds to a concentration of SCB2 similar to that present in the natural supernatant of the producer strain before extraction. This suggests that the whole GBL system (receptor and synthases) could also be used in vivo, albeit some modifications to enhance thermal stability and enzymatic activity might be needed, depending on the application (e.g.; use in human cells). This would then allow to generate complex circuit systems by integrating the whole system inside a single cell to program a determined routine or in different cells (e.g., a sender and a receiver system) to generate a synthetic ecosystems.

\textit{g) Towards establishing butyrolactone signalling circuits for synthetic biology}

The generation of novel regulatory circuits that allow the assembly of predictable genetic devices is one of the major challenges of synthetic biology. Currently, AHL-based circuits are one of the popular choices considered when designing such devices. Here, we report the first steps towards the design of the GBL system as a complementary tool for synthetic regulatory circuits. The results presented are the proof-of-concept that the GBL system from \textit{Streptomyces} can be used in \textit{E. coli} to control synthetic gene expression. An interesting area of future research will be the further optimisation of the GBL biosynthetic genes through directed evolution to obtain more active and thermally stable versions. This would allow expression of the whole system in mammalian cells, in which GBLs have already been found to be active \textsuperscript{60}. Also of great interest would be the diversification of the GBL toolbox by exploiting the modular nature of the GBL biosynthetic pathway, increasing the amount of orthogonal signal systems available to control gene expression. Finally, the interesting property that addition of high SCB2 concentrations results in inhibition of GFP expression could potentially be used to develop genetic band-pass filters that allow expression of synthetic devices only within a narrow range of inducer concentrations.
**Methods**

**a) Bacterial strains and Plasmids.**

Bacterial strains, oligonucleotides and plasmids used in this study are listed in Supplementary Tables S1-S3. *E. coli* DH5α strain was used as a host for plasmid construction and maintenance; *E. coli* BL21 was used for protein expression and GBL production. All oligonucleotides used for PCR amplification and sequencing were synthesized by IDT.

To construct a deletion mutant of *scbB* in *S. coelicolor*, a 3.6 kb *Sal*I-*Sph*I fragment containing *scbB* (*sco6264*) was subcloned into pUC19 to generate pCK1. pCK1 was digested with *Hinc*II and then 1.1 kb of thiostrepton-resistance gene (*tsr*) was inserted. The construct was confirmed to have the desired deletion by DNA sequencing and restriction enzyme digestion. From the resulting plasmid, the *Bam*HI-*Hind*III fragment was subcloned into the *Bam*HI-*Hind*III-digested conjugative vector pKC1132 to generate pCK2. The plasmid pCK2 was transferred into *S. coelicolor* by conjugation via *E. coli* donor ET12567 (pUZ8002) to generate the *scbB* mutant strain LW107 (*ΔscbB*). A deletion mutant of *scbC* (*sco6267*), LW108, in *S. coelicolor* was constructed by inserting *scbC* with an apramycin resistance gene and the oriT on cosmid AH10 using the PCR-targeting technology. Both the *scbB* and *scbC* deletion mutants have been verified by sequencing.

To construct plasmid pTE1059 (*scbA/B/C*), *scbA*, *scbB* and *scbC* were amplified from *S. coelicolor* M145 genomic DNA using primer pairs *scbA*_His_5/*scbA*_3, *scbB*_5/*scbB*_His_3 and *scbC*_5/*scbC*_His_3, respectively. Then these were modified with primer pairs *scbA*_mod_F/*scbA*_3, *scbB*_mod_F/*scbB*_His_3 and *scbC*_mod_F/*scbC*_His_3; and then assembled using Infusion (CloneTech) on plasmid pBbA2k, previously digested with EcoRI and XhoI, together with DNA fragments containing promoters Pb10 and Pb19, built through annealing of primers Pb10_5/3 and Pb19_5/3. Plasmids pTE1060 (*scbA/C*) and pTE1061 (*scbA/B*) were constructed by restriction digestion of pTE1059 with NotI and ApaI, respectively, followed by religation.

Strain *S. coelicolor* LW94 was built by integration of reporter plasmid pTE1062, which is plasmid pTE134 with *scbR* in the same orientation as the kanamycin resistant gene (*neo*).

Plasmid BC-A1-002 was a gift from Brian Chow (Addgene plasmid # 78689). To build plasmid pTE1066 or pTE1067, BC-A1-002 was digested with EcoRI and NotI and ligated with EcoRI and NotI digested PCR product containing *scbR* (primer pair: *scbR*_5/3) or *scbR*-Arg6 (primer pair: *scbR*_5 and *scbRarg6*_3). Exchange of promoter regions was performed...
either via whole plasmid amplification, followed by religation (plasmids with \textit{scbRp} and \textit{scbAp}) or by ligating a fragment generated through primer annealing to a HindIII and SpeI-digested reporter plasmids (plasmids with \textit{cpkOp}). All plasmids were confirmed by sequencing.

\textit{b) Culture media and conditions}

Luria Broth (LB) medium containing 10 g/L peptone, 5 g/L yeast extract, and 10 g/L NaCl was used for standard bacterial growth and GBL production. GBL input-response relationship was measured using M9 minimal medium (M9) containing 1X M9 salts (Sigma), 0.4 % glucose, 2 mM MgSO$_4$ and 0.1 mM CaCl$_2$, pH 7.0. To evaluate GBL production under different medium, LB and M9 medium were used as previously described; as well as Terrific Broth (TB) containing 12 g/L Tryptone, 24 g/L yeast extract, 0.4 %(v/v) glycerol, and 1X phosphate salts (0.17 M KH$_2$PO$_4$, 0.72 M K$_2$HPO$_4$, pH 7.0); and 2xYT medium (2xYT) containing 16 g/L Bacto Tryptone, 10 g/L Bacto Yeast Extract and 5 g/L NaCl, pH 7.0. Chloramphenicol and kanamycin were supplemented into the media, where appropriate, at concentrations of 30 µg/mL and 50 µg/mL, respectively.

\textit{c) Expression of the GBL biosynthetic pathway in \textit{E. coli}}

Single colonies of \textit{E. coli} BL21 strains carrying expression plasmids with the GBL synthases (pA15 origin of replication, ~5 copies per cell) were grown at 37 °C for 16 - 18 h in LB supplemented with 50 µg/mL of kanamycin. Samples were diluted to an OD$_{600nm}$ of 0.05 in LB medium without antibiotics and grown at 37 °C until they reached an OD$_{600nm}$ of 0.3 - 0.4, when protein expression was induced with 50 nM anhydrotetracycline (aTc) and performed for 5 h at 25 °C. Cells were pelleted at 10,000 xg and 4 °C for 10 min and resuspended in lysis buffer C (150 mM Tris-HCl pH 7.4, 200 mM NaCl and 1 mM DTT, cOmplete™ tablet (Sigma)) and homogenised by sonication. The soluble protein fraction was recovered after centrifugation at 17,000 xg and 4 °C for 20 min and the insoluble fraction was resuspended in MilliQ grade H$_2$O.

When evaluating the robustness of the GBL biosynthetic pathway under different experimental conditions, \textit{E. coli} cells containing the \textit{scbA/B/C} (pTE1059) plasmid and originated from the same colony, were grown at 25 °C in either M9 , LB, 2xYT or TB and in LB at 20, 25, 30 and 37 °C.
d) **GBL extraction and purification**

Extraction of GBLs from *S. coelicolor* strains grown in solid cultures was performed as previously described \(^{63}\). For the extraction of GBL or intermediate metabolites from *E. coli* BL21, single colonies of *E. coli* BL21 strains carrying the corresponding plasmids were grown and expressed as described in the previous section. After centrifugation, the supernatant was mixed with an equal volume of ethyl acetate and vigorously mixed. Organic phase was separated, dried with MgSO\(_4\) and the solvent was removed at room temperature and resuspended in 1:100 volume of methanol. In order to purify SCB2, 1.5 L of *E. coli* BL21/pTE1059 (*scbA/B/C*) were grown as previously described, with induction at 20 °C for 16 - 18 h. Purification was performed on a C\(_{18}\) reverse phase column (Kinetex 5 µm C18 100 Å 250 x 21.2 mm) on a preparative HPLC (Agilent 1250 Infinity) equipped with a UV detector set at 210 nm and 254 nm and with a flow rate of 5 mL/min. A maximum volume of 5 mL were loaded onto the column and eluted in a linear gradient of 5-100 % methanol + 0.1 % formic acid. Samples were collected every minute and subjected to bioassay. Positive samples were pooled, the solvent removed at room temperature and diluted in 5 mL of methanol. These were loaded again onto the column and eluted in a linear gradient of 5-100 % of acetonitrile + 0.1 % formic acid. Samples were collected every minute and subjected to bioassay. Active samples were pooled, the solvent removed at room temperature and diluted in 5 mL of methanol. This procedure was repeated four times and all samples resuspended in methanol pooled together. The solvent was removed and the samples were weighted, resulting in a total of 15.2 mg of extract obtained. To assess the purity of the sample, the extract was diluted in 1 mL of methanol and subjected to bioassay in serial dilutions of 2-fold. The minimum active dilution was considered to contain 0.05 µg/µL of SCB2, as previously determined by Hsiao et al. \(^{34}\). This yielded an active concentration of 3.2 mg, resulting in a yield of 2.2 mg/L culture.

e) **Kanamycin Bioassays**

Kanamycin bioassays were performed as previously described (\(^{63}\)). Briefly, a lawn of *S. coelicolor* LW94 was prepared in DNAgar plates supplemented with 5 µg/mL of kanamycin by diluting 2.6x10\(^6\) spores / 100 µL sterile deionised water and evenly spreading across the plate. The plates were allowed to dry for three minutes at room temperature and then 2-3 µL of methanol extract were spotted on the plate and allowed to dry at room temperature. Plates were incubated at 30 °C for 2 - 3 days and growth was monitored every 12 h.
f) HPLC and MS Analysis

A volume of 100 µL of sample (e.g., GBL extract) were diluted 1:2 with HPLC water (total volume 200 µL), centrifuged at 17,000 x \( g \) for 15 min to remove any aggregates formed and transferred 150 µL of the resulting product into a glass vial. A total of 15 µL were injected per sample. The solvents used were A: H\(_2\)O + 0.1% formic acid and B: MeOH + 0.1% formic acid. Both solvents were HPLC grade, from Sigma. The run conditions were 5 min isocratic A:B (95:5) then gradient of 25 min to A:B (5:95), followed by isocratic 5 min A:B (5:95) in an Accucore UHPLC C\(_{18}\) reverse phase column (Thermo Fisher) and a Q Exactive (Thermo Fisher). The ionization conditions were in positive ion mode at a spray voltage of 1.5 kV.

g) ScbR expression

Expression of ScbR was analysed using *E. coli* BL21 harbouring plasmids ScbR + luxp (pTE1066) or ScbR-Arg6 + luxp (pTE1067). Single colonies were grown for 16 - 18 h at 37 °C in LB medium supplemented with 30 µg/mL of chloramphenicol. Cells were diluted 1:100 in LB without antibiotics and grown at 30 °C for 6 h. Cells were centrifuged at 10,000 x \( g \) and 4 °C for 10 min, the supernatant was discarded and the pellet homogenised by sonication in buffer A \( ^{64} \) (50 mM sodium phosphate buffer pH 7.0, cOmplete\textsuperscript{TM} tablet (Sigma)). The soluble protein fraction was recovered from the supernatant after centrifugation for 20 min at 17,000 x \( g \) and 4 °C and the pelleted insoluble fraction was resuspended in MiliQ grade H\(_2\)O.

h) SDS-PAGE and Western Blot

To assess protein expression, crude extracts (either soluble or insoluble fractions) were resolved through SDS-PAGE (10 %(w/v), BioRad) according to Laemmli’s procedure \( ^{65} \). Resolved bands were visualised by Coomassie blue staining (Expedeon). For Western analysis, proteins resolved by SDS-PAGE gels were transferred to a polyvinylidene fluoride (PVDF) membrane by semi-dry blotting. Immunodetection of His-tagged ScbA, ScbB or ScbC was performed using mouse anti-His (Sigma H1029) as primary antibody and IRDye\textsuperscript{®}-conjugated anti-mouse IgG (Abcam ab216772) as secondary antibody, and visualised using LI-COR. Immunodetection of ScbR was carried out using rabbit antiserum raised against ScbR \( ^{66} \) as primary antibody and HRP-conjugated goat anti-rabbit IgG (BioRad) as secondary antibody. The substrate for chemiluminescent detection was Amersham ECL Prime (GE Healthcare) and was visualised using a GeneGnome (Syngene).
i) **Gel retardation Assays**

Gel shift assays were performed as previously described \(^6^2\), using the Roche DIG Gel Shift Kit (Roche). The scbRp was amplified from genomic DNA using primers scbRp_5/3, generating a 144 bp fragment that was labelled with DIG according to manufacturer’s protocol. For each reaction, 25 ng of labelled probe were used. Where appropriate, SCB2 extracts were added to the mixture prior incubation. DIG-labelled DNA fragments were immunodetected using antibody mouse anti-DIG (Abcam ab116590) as primary antibody and IRDye®-conjugated anti-mouse IgG (Abcam ab216772) as secondary antibody, and visualised using LI-COR.

j) **Relationship between signal input and response of GBL and AHL receptors**

Single colonies of *E. coli* BL21 cells containing appropriate plasmids were grown for 16 - 18 h in LB medium with 30 µg/mL of chloramphenicol and then diluted 1:100 in M9 without antibiotics and grown at 37 °C until an OD\(_{600}\) 0.3 - 0.4, where they were aliquoted into 500 µL aliquots and supplemented with 1 % (v/v) of the serial dilutions of either 3OC\(_6\)-HSL (Sigma-Aldrich), 3OC\(_{12}\)-HSL (Sigma-Aldrich) or SCB2 (this study) in methanol. Samples were grown at 30 °C for 20 h and then transferred into a 96-well plate in triplicate (150 µL sample/well). OD\(_{600}\) and GFP fluorescence (excitation, 466 nm; emission, 511 nm) were measured in triplicate in a ClarioStar (BMG Labtech) plate reader. Each measurement was performed in biological triplicate.

k) **Structural model of ScbR and docking**

The model structure of ScbR was generated using the SWISS-MODEL suite \(^4^7\) and CprB crystal structure (PDB 1UJ5 \(^4^8\)) as template. The resulting structure was exported as PDB file and visualised using the UCSF Chimera package \(^4^9\). The protein surface was generated using default settings and the hydrophobic regions highlighted with the kdHydrophobicity command line option. Dockings were performed using Autodock Vina software under standard configuration and a grid previously defined with Autodock Tools \(^5^0\), and were either a broad grid covering the whole protein surface or a more constraint grid covering only the C-terminus end of the protein.
Supporting Information. Table S1: Bacterial strains; Table S2: Plasmids; Table S3: Oligonucleotides; Figure S1: Western Blot analysis of ScbA, ScbB and ScbC protein expression; Figure S2: Kanamycin bioassay to assess production of GBLs in *E. coli* containing plasmids for the expression of ScbA/B/C; Table S4: MS1 adducts of SCB2 and A-factor; Figure S3: Tandem MS of SCB2 and A-factor; Figure S4: Structural Model of ScbR; Figure S5: Docking results; and Figure S6: Purified SCB2 HPLC-MS analysis.

Author information

Author list:

Marc Biarnes-Carrera¹, Chang-Kwon Lee²,⁴, Takuya Nihira²,³, Rainer Breitling¹ and Eriko Takano¹

Affiliation:

¹Manchester Centre for Synthetic Biology of Fine and Speciality Chemicals (SYNBIOCHEM), Manchester Institute of Biotechnology, School of Chemistry, Faculty of Science and Engineering, University of Manchester, 131 Princess Street, Manchester M1 7DN, United Kingdom.

²International Center for Biotechnology, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

³Mahidol University-Osaka University Collaborative Research Center for Bioscience and Biotechnology, Faculty of Science, Mahidol University, Rama VI Rd., Bangkok 10400, Thailand

⁴Present address: Mong-Go Food Co., Ltd., Changwon 641-847, Republic of Korea

Author Contributions

MBC and CKL performed the experiments. MBC, TN and ET designed the experiments. MBC, TN, RB and ET analysed the results and wrote the manuscript. All authors have read, edited and approved the manuscript.

Corresponding Author

*E-mail: eriko.takano@manchester.ac.uk.*
Notes
The authors declare no competing financial interest.

Acknowledgments:
MBC was supported by the School of Chemistry, Faculty of Science and Engineering, University of Manchester. Plasmid BC-A1-002 was a gift from Brian Chow. This is a contribution from the Manchester Centre for Synthetic Biology of Fine and Specialty Chemicals (SYNBIOCHEM) and acknowledges the Biotechnology and Biological Sciences Research Council (BBSRC) and Engineering and Physical Sciences Research Council (EPSRC) for financial support (Grant No. BB/M017702/1).

References


21
Figure 1. Schematic representation of the GBL biosynthesis gene cluster from *S. coelicolor* A3(2) M145 and the possible biosynthetic pathway leading towards production of GBLs: SCB1, SCB2 or SCB3 in *S. coelicolor*; based on the biosynthetic pathways proposed by Kato et al. (29) for A-factor production in *S. griseus* and by Shikura et al. (30) for virginiae butanolides in *S. virginiae*. The biosynthesis of GBLs has been proposed to start with the condensation of DHAP with a β-ketoacid by (29, 31) ScbA in *S. coelicolor* A3(2), leading to product 3, which putatively undergoes a spontaneous intramolecular Claisen condensation, resulting in butenolide 4. This compound is reduced by a butenolide phosphate reductase, ScbC, yielding 5, which could be hydrolysed, yielding compound 8. The phosphate group is lost, resulting in the A-factor-like GBL 6, which can also be hydrolysed to render the open lactone 9. The final GBLs (7) are obtained after stereo-specific reduction through a 3-ketoacyl-ACP/CoA reductase, ScbB. During transition phase, GBLs accumulate in the environment and bind to the receptor ScbR, resulting in a switch-like transition towards antibiotic production.
Figure 2. Kanamycin bioassay plates with different concentrations (1 µL, 2.5 µL, 5 µL and 10 µL) of GBL ethyl acetate extracts from the different *Streptomyces* strains as depicted above the plates. Growth in the presence of kanamycin is detected in all concentrations of *S. coelicolor* M145 extract, whereas no growth is seen for extracts from LW107 (scbA/C) and LW108 (scbA/B) mutants, suggesting these genes are involved in the production of GBLs.
Figure 3. Analysis of intermediate metabolites in GBL biosynthesis through LC-MS. Extracted Ion Chromatograms (EIC) of extracts from *E. coli* cells expressing pTE1059 (*scbA/B/C*), pTE1060 (*scbA/C*) or pTE1061 (*scbA/B*). (A) When expressing the three genes using pTE1059, a EIC peak is detected that elutes at the same retention time as SCB2 from *S. coelicolor* M145 extract. This peak was further shown to be SCB2 by MS/MS analysis (Supplementary Fig. S3). (B) After deleting *scbB* in pTE1060, this peak is no longer detectable; instead two peaks are detected at a mass corresponding to A-factor, which elute at the same retention time as the two peaks of A-factor from *S. griseus*, suggesting that *scbB* is responsible for the stereospecific reduction of compound 6. (C) Deletion of *scbC* in pTE1061 results in an unidentified detectable peak at a mass corresponding to A-factor, which elutes at a distinctive retention time. The concentration of the anhydrotetracycline (aTc) and the positive control extracts from *S. coelicolor* or *S. griseus* is denoted at the side. (D) Overlap of peaks identified in *E. coli* ethyl acetate extracts showed in panels A - C, highlighting the distinctive retention time between them.
Figure 4. Qualitative production of SCB2 at different temperatures and media.
Production of SCB2 in *E. coli* was assessed under different temperatures from 20 °C to 37 °C and different media conditions, minimal to rich media. (A) Kanamycin bioassay with extracts from *E. coli* producer cells grown at different temperatures. Expression of SCB2 is detected at assayed temperatures from 20 – 30 °C, but not at 37 °C. (B) Kanamycin bioassay with extracts from *E. coli* harbouring *scbA/B/C* (pTE1059) grown under different media. All media allowed production of SCB2.
Figure 5. Western Blot and gel retardation analysis of ScbR-Arg₆. (A) Western Blot analysis of ScbR and ScbR-Arg₆ in ScbR + luxp plasmid (pTE1066) and ScbR-Arg₆ + luxp plasmid (pTE1067), respectively, at stationary phase in the soluble and insoluble fractions, and compared to plasmid BC-A1-002, which does not contain ScbR. A faint band corresponding to ScbR (black arrow) can be seen in pTE1066 soluble fraction, which is stronger in pTE1067. Addition of the Arginine tag improved the solubility of ScbR. (B) Gel retardation analysis of ScbR-Arg₆ against a DNA fragment containing the operator sequence OR (26). Addition of increasing amounts of ScbR-Arg₆ results in formation of complex ScbR-Arg₆:OR and 2 ScbR-Arg₆:OR, showing that addition of the Arg₆ tag does not affect DNA-binding properties of ScbR.
Figure 6. Signal and promoter crosstalk between the AHL and the GBL systems: constructs. Schematic representation of the plasmids built to evaluate signal and promoter crosstalk between the AHL and the GBL systems. In plasmid BC-A1-002 (34), constitutive expression of the luxR results in production of LuxR, which binds to the LUX box when the concentration of AHLs reaches a concentration threshold. There, it induces gfp expression by RNA polymerase recruitment. Exchange of the luxp and the LUX box from BC-A1-002 for scbRp and ScbR cognate operator site (OR) generated plasmid pTE1063, used to evaluate whether LuxR could interact with ScbR OR. Parallel construction of pTE1067 was created by replacing luxR for scbR-Arg6 and was used to evaluate whether ScbR could interact with the LUX box. Finally, plasmid pTE1058 replaced both the luxp by scbRp and luxR by scbR-Arg6 and should respond only to the addition of SCB2.
Figure 7. Signal and promoter crosstalk between the AHL and the GBL systems. Normalised GFP/OD$_{600}$ output after addition of different concentrations of autoinducer concentrations ([Al]/M), either 3OC$_6$HSL (blue line), 3OC$_{12}$HSL (green line) or SCB2 (red line); to strains harbouring plasmids BC-A1-002 (LuxR+luxp), pTE1063 (LuxR+scbRp), pTE1067 (ScbR-Arg$_6$+luxp) and pTE1068 (ScbR-Arg$_6$+scbRp) (Fig 6). Measurements were made ~20 h after induction and growth at 30 °C. Expression of $gfp$ is at the maximum with concentrations of $10^{-9}$ M of 3OC$_6$HSL and $10^{-8}$ M of 3OC$_{12}$HSL, when added to BC-A1-002. However, no GFP expression is seen when AHLs are added to the cells with the ScbR-Arg$_6$+luxp (pTE1067). On the other hand, addition of any signalling molecules results in no change in $gfp$ expression in LuxR+scbRp (pTE1063), where scbRp is always active. GFP expression is induced upon the addition of SCB2 with the ScbR-Arg$_6$+scbRp (pTE1068). Interestingly, addition of an excess of SCB2 in ScbR-Arg$_6$+scbRp results in repression of the system. No induction of GFP was observed with addition of AHLs to ScbR-Arg$_6$+scbRp (pTE1068).
Figure 8. Expression of GFP with different ScbR-dependent promoters results in a gradient of output signal. Normalised GFP/OD$_{600}$ output after addition of different concentrations of autoinducer ([AI]/M) 3OClHSL (blue line) or SCB2 (red line); to strains harbouring pTE1063 (LuxR + scbRp), pTE1064 (LuxR + scbAp), pTE1065 (LuxR + cpkOp), pTE1068 (ScbR-Arg$_6$ + scbRp), pTE1069 (ScbR-Arg$_6$ + scbAp) or pTE1070 (ScbR-Arg$_6$ + cpkOp). Measurements were made ~20 h after induction and growth at 30 °C. As seen in Fig 7, plasmids containing LuxR are active at all concentrations of signals, suggesting that LuxR does not interact with any of the SebR operator sequences to repress the promoters. Exchange of luxR for scbR-Arg$_6$ results in repression of the promoters until around 10$^{-9}$ M of SCB2. Interestingly, both scbRp and scbAp are only active in concentrations between 10$^{-8}$ to 10$^{-4}$ M of SCB2, respectively. However, cpkOp is active at all concentrations above 10$^{-9}$ M, and the activity was not shut down.
Figure 9. Response analysis of plasmid pTE1068 under different extracts obtained from 10 mL LB culture of *E. coli* producer cells. The results match the previously observed characteristics, where lower temperatures seem to be favourable for SCB2 production (Fig. 3). Interestingly, addition of a 100-fold concentrated ethyl acetate extracts of the liquid culture supernatant to the reporter cells (shaded box) resulted in full activation of *gfp* expression at 20 °C and about half activation of *gfp* expression at 25 and 30 °C, suggesting that the system could be used at different temperatures to fine tune gene expression.