A Biosensor for Detection of Indole Metabolites

Jiefei Wang[#], Chao Zhang[#], W. Seth Childers*

Department of Chemistry, University of Pittsburgh, 219 Parkman Avenue, 15260, United States

#Equal contributions.

Corresponding Author:

W. Seth Childers

Chevron Science Center, Room 801

219 Parkman Avenue

Pittsburgh, PA 15260

Phone Number: 412-624-3058

E-mail: wschild@pitt.edu

ABSTRACT

The lack of diverse health-related biosensors limits the progress towards our understanding of how the microbiome metabolism impacts health. Microbially produced indole-3-aldehyde (I3A) has been associated with reducing inflammation in diseases such as ulcerative colitis by stimulating the aryl hydrocarbon receptor (AhR) pathway. We mined the protein database for gut microbiome metabolites' sensors and developed a biosensor for I3A. We engineered *E. coli* embedded with a single plasmid carrying a chimeric two-component system that detects I3A. Our I3A receptor characterization identified residues that contribute to the sensor's high specificity in a range of 0.1-10 µM. The I3A biosensor opens the door to sensing indole metabolites produced at the host-microbe interface and provides new parts for synthetic biology applications.

KEY WORDS

Biosensor, Microbiome, Indole-3-aldehyde, Two-component system, Archaea, Indole.

INTRODUCTION

Synthetic biology has demonstrated the potential for probiotics to address a range of health issues, including the treatment of rare metabolic disorders[1] and pathogenic infections[2, 3]. The lack of diverse parts to sense health-related signals limits these applications. We mined the protein database for sensors that could detect signals associated with gut-microbiome interactions.

Amongst microbiome-produced metabolites, indole and its derivatives serve as important interkingdom signaling molecules[4]. In the gastrointestinal tract, tryptophan is metabolized by three major metabolic pathways[5]: monoamine neurotransmitter serotonin (5-hydroxytryptamine, 5-HT) production[6], nicotinic acid production through kynurenine (Kyn) pathway[7], and production of indole compounds triggering aryl hydrocarbon receptor (AhR) pathway[8]. *Lactobacillus*-generated indole-3-aldehyde (I3A), which is detected in gastric fluids and urine, stimulates AhR mediated transcription[8]. I3A stimulation of the AhR receptor may have broad health implications that include regulating gut barrier function[9], reprogramming T cells[10], reducing skin inflammation[11], central nervous system (CNS) inflammation[12], and inflammation associated with aging[13]. Similarly, pathogenic *E. coli* strains (EPEC, EHEC, and EAEC) secrete certain indole derivatives, including I3A and indole-3-acetic acid (I3AA), to kill *C. elegans*[14]. Therefore, I3A is a critical microbially produced biomarker that directly impacts health through AhR activation.

Compared to traditional detection methods, such as mass spectroscopy, cell-based biosensors take advantage of the natural ability of a cell to sense and respond to the environment. Thus cell-based biosensors provide the potential for longitudinal measurements and noninvasive *in vivo* clinical applications[2, 15, 16]. Because indole compounds are mainly produced through the metabolism of tryptophan by microbes, we seek to build a sensor for indole compounds in bacteria. While indole derivatives act on eukaryotic AhR[8, 17], those interactions are non-specific and function as a eukaryotic transcription factor, presenting challenges to repurpose AhR as an indole bacterial biosensor. Though indole acts upon bacterial sensor BaeS, CpxA, and EvgS[18, 19], it remains unclear if those indole-receptor interactions are direct or indirect through indole's impact upon membrane potenial[20, 21]. Therefore, the lack of biochemically characterized indole metabolite sensors prevents the use of this biomarker in health applications.

One of the primary biochemical systems microbes use to sense their surroundings is two-component systems (TCSs). TCSs provide a cornerstone of bacterial signaling and coordinate processes such as the cell cycle[22], chemotaxis[23], cell-cell communication[24], and bacterial virulence[25]. A TCS consists of a histidine kinase (HK) and a response regulator (RR), which are responsible for sensing environmental signals and regulating cellular processes, respectively (Figure 1A). TCSs are promising targets for engineering biosensors for four reasons[26].

First, TCSs provide a warehouse of sensors[27], including pH, metals, quorum sensing compounds, and secondary metabolites. For instance, TCSs sensing a broad spectrum of light are used as optogenetic tools in *E.coli*[28-32]. TCS sensing thiosulfate has been developed and applied in *E.coli* for diagnosing colon inflammation in mice[3, 33]. Also, TCS metabolic sensing has been developed to screen unique chassis producing small molecules and feedstock[33-37]. Despite significant HK characterization progress, the sensing signals remain unknown for most HKs[38-40].

3

The fact that signals remain unknown is partially due to the lack of crystal structures of sensors co-crystallized with their stimulating ligands.

Second, TCSs have remarkable plasticity in how signals flow within and between the HK and RR. On the one hand, the sensing domain can be swapped with homologous sensory domains to sense different signals[41, 42]. On the other hand, TCSs also exert diverse outputs that include regulation of transcription, translation, motility, and proteolysis[43, 44,45]. Additionally, the HK-RR protein-protein interaction interface can be engineered to alter phosphate transfer to other response regulators[46]. Therefore, the two-component systems' modularity allows for the straightforward and customized rewiring of a signal input to the desired signaling outputs.

Third, productive cross-talk amongst several TCSs can function as an integrated multi-kinase network (MKN)[47]. MKNs allow sensing multiple signals simultaneously and minimize the chance of off-target regulation in a physiological environment. For example, to improve the stringency of uranium sensing, two functionally independent TCSs, UzcRS, and UrpRS, were assembled as a AND gate[48].

Fourth, TCSs are transferable to eukaryotes[49, 50], which expand engineered biosensors' potential to be applied in prokaryotic and eukaryotic cells. Here we identified an I3A binding domain and engineered a chimeric HK that regulates mCherry expression upon the addition of I3A in *E. coli*. We engineered two one-plasmid systems, which carry the I3A binding domain within a chimeric histidine kinase and its cognate response regulator. This initial indole metabolite sensor provides a scaffold in synthetic biology for engineering sensors for a diverse repertoire of indole metabolites. This includes metabolites such as indole, which is known to induce virulence factors, promote resistance to pathogens, and promote homeostasis of gut epithelial function[5, 8, 51].

RESULTS & DISCUSSION

Identification of an indole-3-aldehyde binding PAS domain

We mined the protein database (PDB) for bacterial sensory domains that interact with signals in human microbiome environments. We focused our search upon Per-Arnt-Sim (PAS) domains, as they are essential sensing modules that commonly regulate histidine kinases. We found 1293 PAS domain entries and identified 38 entries that bind unique ligands (Table S1).

Amongst ligand-associated PAS domains, we identified a PAS domain co-crystallized with I3A (PDB: 3BWL)[52]. I3A plays a vital role in maintaining mucosal homeostasis by stimulating the AhR signaling pathway[5, 8]. The I3A-binding PAS domain belongs to the protein HtlD (halobacterial transducer of rhodopsin (HTR)-like protein, UniProt Q5V5P7)[53] encoded by *rrnAC0075* in halophilic Archaeon *Haloarcula marismortui*. Accordingly, we named the I3A binding PAS domain as HtlD[I3A-PAS] in this work and characterized the ability of HtlD[I3A-PAS] to serve in an I3A biosensor.

Engineer a two-component system in E.coli to sense indole-3-aldehyde

The HtlD signaling protein consists of a receiver domain (REC) followed by an HK domain (Figure 1B). However, there are no cognate RRs or histidine phosphotransfer proteins (Hpt) within the proximal genomic regions. The lack of knowledge of the downstream RR and the regulating set of genes prevented our ability to refactor the entire signaling system in *E. coli*. To overcome this problem, chimeric HKs have demonstrated the ability to exchange sensory domains for producing biosensors[41, 42]. One such example is the blue light-sensing chimeric

histidine kinase YF1[42, 54, 55]. The sensory domain of the blue light-sensing YF1 is also a PAS domain (YF1$^{\text{LOV-PAS}}$), which is structurally similar to the HtlD$^{\text{I3A-PAS}}$ (Figure S1A). Therefore, to exploit the I3A sensing capabilities of HtlD$^{\text{I3A-PAS}}$, we swapped the LOV sensory domain from YF1 with the HtlD$^{\text{I3A-PAS}}$ and tested if the chimeric HK could sense I3A.

To determine the optimal sensor-HK chimera fusion site, we exploited a set of conserved D-(A/I/V)-(T/S)-E residues at the C-terminus of the PAS domain fold. This conserved motif forms a network of hydrogen bonds with the central beta-sheets of the PAS domain fold[56, 57]. As demonstrated with the YF1 chimera construction[56], this C-terminal motif in PAS provides a conserved modular point to swap PAS sensory domains. Therefore, we swapped the YF1$^{\text{LOV-PAS}}$ to its DIT motif with the HtlD$^{\text{I3A-PAS}}$ up to its DIT motif and named the chimeric protein I3A-HK (Figure 1B, Figure S1B).

We introduced the I3A-HK into two plasmids, pDusk and pDawn[54], which contain well-characterized parts downstream of the sensor and exhibits low background in the absence of blue light[54]. We designed and constructed plasmid pDusk-I3A_p1. Within the plasmid pDusk-I3A_p1, a bicistronic operon, regulated by the promoter p1, controls the constitutive expression of the I3A-HK and the paired RR. I3A regulates phosphorylation of the RR, then triggers transcriptional repression of the mCherry reporter gene[54, 56] (Figure 1C). To invert the signal so that the introduction of I3A increases the reporter gene expression, we modified the pDawn plasmid[54] to incorporate the I3A-HK resulting in the plasmid pDawn-I3A (Figure S2A). In the engineered pDawn-I3A, I3A represses the expression of the λ phage repressor cI, which represses the expression of mCherry through the λ promoter pR[54, 58].

Upon addition of 10 μM I3A to strains containing the pDusk-I3A_p1 plasmid, we observed no impact upon cell growth (Figure S2B) and a 3-fold decrease in mCherry expression

relative to cultures supplemented with DMSO (Figure S2B). An I3A titration yielded a dose-response curve with an IC_{50} of 0.3 ± 0.1 µM (Figure 1D, Table 1). Similarly, upon the addition of 10 µM I3A to strains containing the pDawn-I3A_p1, we observed a 2.5-fold increase in mCherry expression. Further I3A titration studies yielded a dose-response curve with an EC_{50} of 11.5 ± 7.8 µM (Figure S2C, Table 1). Therefore, we have constructed two biosensors that repress or stimulate output gene expression upon the addition of I3A. Notably, the pDawn-I3A system exhibited a higher EC_{50} than the IC_{50} of the pDusk-I3A system (Table 1). We suspect that the difference is caused by the involvement of the λ phage repressor cI in pDawn-I3A. Other than the involvement of cI, the difference may be caused by the dynamic range of the pDusk-I3A_p1. Given that the dynamic range of the pDusk-I3A_p1 is narrow and the fluorescent intensity is low even without I3A, a high concentration of I3A may lead to cut-off distortion in detection.

Tuning the dynamic range of the I3A sensor

One approach to improving the dynamic range is to tune the ratio of receptor and substrate concentration[31]. To adjust the receptor to substrate ratio, we varied the promoter of I3A-HK using a set of well-characterized promoters from the BIOFAB kit[59] (Figure 1C, Table S2): p0, p2, and p3. We grew *E. coli* cultures carrying constructed plasmids in the presence of DMSO or I3A at various concentrations to determine the effect of I3A upon the mCherry expression (Figure 1D). The dose-response curves were fitted to the Hill equation, as shown in Table 1. Replacing the promoter enlarged the dynamic range, consistent with the strength of the promoter region (Table S2), with observed fold changes of 6.4-fold for p1, 25-fold for p2, and 35.5-fold for p3. The IC_{50} is 0.4 ± 0.1 µM for p2 and 0.7 ± 0.1 µM for p3, similar to the IC_{50} for p1. The weakest selected promoter (p0) displayed no observable changes in mCherry expression

7

upon the addition of I3A. Therefore, we demonstrated that the response to I3A could be tuned predictably by varying the promoter strength for the I3A-HK.

Our ability to accurately sense I3A could be affected by endogenous production of I3A by *E. coli* or the abundance of I3A in LB media. Indeed, co-crystallization of the sensor-I3A complex indicates that I3A likely interacts with HtlD[I3A-PAS] during protein expression[52]. Moreover, past experiments suggest that *E. coli* might produce I3A[14]. To interrogate how much I3A is in the growth culture, we measured the I3A concentration in the cell culture media via HPLC-MS. We observed that freshly prepared LB media contained 0.19 ± 0.07 µM of I3A, while minimal M9 media contained no detectable amounts of I3A.

We recovered 0.13 ± 0.03 µM of I3A from *E.coli* BL21 and 0.15 ± 0.05 µM of I3A from *E.coli* MG1655 wild-type strains. The I3A concentration from both BL21 and MG1655 wild-type strains are similar to the amounts of I3A in fresh LB media, 0.19 ± 0.07 µM (Figure 1E, Figure S2D-F). Therefore, we observed no additional I3A production by *E* coli under the tested growth conditions. However, it is possible that I3A production could be condition or strain-dependent[14]. Therefore, the endogenous I3A in fresh LB media provides a minor 0.19 ± 0.07 µM offset to our calculated IC_{50} and EC_{50}. The unexpected observation of I3A in the LB media (Figure 1E) may warrant its future investigation of I3A as a potential yeast secondary metabolite or the source of its contamination.

Interrogation of the I3A receptor binding site

A key question in part characterization is if structural features of the I3A chimeric sensor function as expected. PAS domains contain a conserved hinge motif D-(A/I/V)-(T/S) that couples the conformational change within the PAS to the conformational re-arrangement that

8

allows HK autophosphorylation[60-62]. Therefore, we mutated the DIT motif of the I3A-HK gene

to AIA (i.e., D507A and T509A) to interrogate if this hinge motif was also critical for I3A

sensing. We observed that the DIT to AIA mutation in the pDusk-I3A_p1 plasmid diminished

the responsiveness to I3A up to 10 µM (Figure 2A, Figure S2G). This indicates that the DIT

motif mutants disrupt signaling, which may be due to a lack of allostery between the sensor and

HK as observed in other HKs[56, 63, 64]. However, we cannot rule out that the double mutant caused

protein misfolding. The data indicate that the I3A stimulation of mCherry expression requires a

functional PAS sensory domain.

We next tested how changes to the I3A PAS domain binding site residues impacted I3A

sensing. We examined three interacting residues within 5 Å of the I3A ligand: D414, D449, and

F472 (PDB: 3BWL)[52] (Figure 2A-B). We mutated D414, D449, and F472 within the I3A-HK

gene individually to alanine in the pDusk-I3A_p1 plasmid using site-directed mutagenesis and

interrogated I3A-regulated mCherry expression (Figure 2C, Figure S2G). D414A and D449A

I3A-HK variants were not responsive to I3A up to a concentration of 10 µM. In comparison,

rather than reduction, the F472A I3A-HK variant resulted in 1.5-fold induction of mCherry. This

observed inversion of the response to I3A by the F472A variant, suggests that F472 may be

involved in allosterically coupling ligand binding to changes in the HK conformational state.

Overall, the interrogation of the I3A receptor binding site confirmed that D414, D449, and F472

are critical for I3A sensing. Future engineering of the amino acid positions present opportunities

to improve I3A-binding affinity for the sensor or alter the substrate specificity to other closely

related indole metabolites.

Interestingly, in the presence of DMSO, the I3A binding site mutants (D414, D449, and

F472) exhibited a 3 to 10-fold increase in mCherry expression relative to the wild-type I3A-HK

9

(Figure S2G). This increased mCherry expression could be due to increased expression levels of the I3A-HK binding site mutants, similar to our observation of promoter tuning (Figure 1). Although, each I3A-HK variant used the same promoter. Alternatively, the observed I3A in the LB media by HPLC-MS (Figure 1E) may repress mCherry expression in the wild-type I3A-HK sensor but may be unable to stimulate the I3A-HK binding site mutants (Figure S2G). We envision the I3A biosensor will provide a powerful tool to identify the bacterial I3A biosynthetic pathway in bacteria suspected to produce I3A[8].

Selective I3A sensing through PAS domain

To examine the specificity of the I3A sensing, we treated the engineered *E. coli* against a panel of 11 closely related indole metabolites at a concentration of 10 µM and 100 µM. Many of these indole metabolites only differ by one functional group where the R_1-R_4 group is substituted from -CHO in I3A to other groups (Figure 3A). Amongst this set, only I3A could stimulate mCherry expression at a concentration of 10 µM (Figure 3B). We observed a similar specific I3A stimulation of mCherry expression from the pDawn-I3A plasmid (Figure S3A) as from pDusk-I3A-p3 plasmid (Figure S3B). Lack of stimulation by the other closely related indole metabolites could be due to a lack of binding to the I3A sensor or the inability of the indole metabolites to cross the cell wall. At a 100 µM ligand concentration, we observed a mild but statistically significant reduction of mCherry expression of less than 10% by the following: I3M, I3AA, IND, I7OH, and IMI (Figure S3B). These results indicate that I3A is a potent and specific stimulator of I3A-HK. While I3M, I3AA, IND, I7OH, and IMI weakly stimulate mCherry expression at concentrations 100-fold greater than I3A mediated stimulation of I3A-HK.

Notably, indole has been the most studied amongst these metabolites and reported at a much higher stimulating concentration[20]. We measured indole concentration via LC-MS and found *E. coli* BL21 strain produce 237.5 ± 1.1 µM indole in LB media (Figure S3C-D). However, the amount of indole endogenously produced by the *E. coli* BL21 strain expressing our chimeric I3A-HK did not stimulate complete repression of mCherry expression. Therefore, we tested higher exogenous addition of indole concentrations in the range of 500-1000 µM and observed repression of mCherry expression (Figure S3E). These data suggest that the I3A-HK has high specificity for I3A and a 100-1000-fold higher IC_{50} for indole. The indole detection could be through direct binding with the I3A sensor, or exogenous indole could stimulate biosynthesis of I3A in *E.coli*[14]. However, we also suspect that high indole concentrations used in the indole titration experiment could induce mild toxicity (Figure S3G) that impacts translation. Further genetic and biochemical studies will be needed to distinguish between direct and indirect indole stimulation models.

In addition, we observed that 100-1000 µM I3AA could be detected by the I3A-HK sensor (Figure S3F and H). The weak indole and I3AA repression indicate that one should cautiously consider conditions known to produce high levels of indole or I3AA[20], as it may provide a false-positive result for I3A detection. Critically, the mild sensing promiscuity suggests that enzyme engineering approaches upon the HtlD[I3A-PAS]-like PAS domains offer the potential to alter its specificity for I3A towards other indole metabolites such as IND or I3AA.

Homologs of the indole-3-aldehyde binding PAS domain

It is not clear why *Haloarcula marismortui* senses I3A and how the regulation through I3A benefits this organism. HtlD[I3A-PAS] homologs exist in archaea and bacteria (Figure S4A and

Table S3), and we investigated if I3A sensing was conserved amongst the HtlD$^{I3A-PAS}$ homologs PAS-1 to PAS-3 (Figure 4 and Figure 4B-D). We found each homologs' specificity corresponds to their sequence similarity to HtlD$^{I3A-PAS}$. We observed that PAS-2-HK could repress the expression of mCherry upon the addition of I3A with an IC$_{50}$ of 1.1 ± 0.4 µM (Figure S4E). Similarly, expression of mCherry could be weakly repressed by the PAS-3-HK upon the addition of I3A with an IC$_{50}$ of about 0.2 ± 0.1 µM (Figure S4E). Notably, both PAS-2-HK and PAS-3-HK exhibited a smaller dynamic range than their parent strain pDusk-I3A_p3 (Figure S4E). Critically, our validation of a direct I3A binding PAS domain that is conserved in a subset of archaea raises questions about the potential role of I3A signaling in archaea[65].

In summary, we have developed a biosensor to detect I3A with high specificity, demonstrating the PDB provides a ripe untapped source for biosensor design. We also identified critical residues surrounding the I3A binding pocket, which will provide the foundation for engineering changes in specificity to detect other health-associated biomarkers such as indole and I3AA. We envision this new I3A biosensor will open the door to new cell-based approaches to interrogate the indole signaling landscape in diverse microbiome environments. Moreover, the I3A biosensor could be a critical new part for cell-based therapeutics[66] that could respond to I3A or even alter the indole signaling landscape[67].

MATERIAL AND METHODS

Chemicals

Indole and its derivatives (Table S8) used in this work were purchased from ACROS organics^{TM} and TCI America^{TM}. Chemicals were stored in appropriate conditions after opening. DMSO was used to prepare a 100 mM stock solution for each ligand and frozen in a -20 ^{o}C

freezer until ready to use. In the assay for measuring dynamic range, 100 mM stock solution was diluted into a few lower concentration stocks in DMSO before adding to bacterial liquid culture.

Bacterial strains and growth

All *E.coli* strains, plasmids, plasmid construction methods, primers, and enzymes used in this study are listed in Table S4-6. Primers used for the construction of plasmids were designed using j5 software[68]. All primers were purchased from Integrated DNA Technologies (IDT, Coralville, IA). All PAS domains were synthesized as gBlocks from IDT and replaced the YF1$^{LOV-PAS}$ within the pDusk plasmid. The sequence for the chimeric site is shown in Figure S1B. For testing promoters, pDusk-I3A was used as templates and treated as p1. The 200 bp region upstream the I3A-HK in pDusk-I3A was replaced with the selected promoter and ribosomal binding site (RBS) regions. The selected regions were cloned from BIOFAB kit strains B7, B11, and B12 for p0, p2, and p3, respectively (Drew Endy Lab, Stanford University, USA)[59]. See Table S2 for sequences. All constructs above were assembled via Gibson Assembly using standard protocols[69]. The plasmids pDusk-I3A and pDusk-I3A_p3 were used as templates to generate single-site mutations. Mutagenesis was done with QuickChange. Plasmids were routinely isolated using the GeneJET Plasmid Miniprep kit (Thermofisher, USA).

E.coli strains (Table S6) were grown in Luria-Bertani (LB) media or on LB plates, including 1.5% w/v agar. Antibiotics were included at the following final concentrations: ampicillin, 100 mg/L, kanamycin, 30 mg/L or chloramphenicol, 30 mg/L.

Fluorescence reporter assay

Plasmids were transformed into chemical BL21(DE3) *E. coli* cells. Cells were plated onto selective LB media plates and grown overnight at 37 °C. From a single colony, 2 mL cultures were initiated and incubated at 37 °C with shaking at 250 rpm for 12 h. The saturated cultures were normalized to $OD_{600} = 1$ and were inoculated at 1% v/v to 1 mL selected LB media. The stock solution of ligands prepared as different concentrations in DMSO was added to the cultures at 1% v/v upon inoculation. Cultures were incubated at 37 °C with shaking at 250 rpm. After 20 h[54], 150 µL from each culture was aliquoted into a black-bottom 96-well plate (Greiner Bio-One, USA) for fluorescence measurements, and 50 µL were aliquoted in 100 µL LB media into a clear 96-well plate for measurements of optical density. All endpoint assays were measured in 96 well plates using the Infinite® M1000 plate reader (Tecan, Switzerland). mCherry fluorescence was measured using an excitation wavelength of 585 nm, an emission wavelength of 610 nm, a fixed gain of 60, and a bandwidth of 5 nm. 0.02mg/mL TAMRA SE dye (ThermoFisher, USA) was diluted and used as a standard instrumental control. The optical density was measured at 600 nm (OD_{600}). The ratio of fluorescent intensity to OD_{600} was recorded as normalized mCherry intensity. All the experiments were done with at least three replicates using cultures from different colonies. The mean and standard deviation of the normalized mCherry intensity were reported.

Fitting to the Hill equation

The measure fluorescent intensity was normalized by optical density for plotting and fitting in Prism9 (GraphPad Software, San Diego, USA). The normalized mCherry intensity at each ligand concentration were fitted to the Hill equation in Prism8 with default parameters, using [inhibitor] vs. response - variable slope function with Y = bottom + (top-

14

bottom)/(1+(IC50/X)^hillslope), or [agonist]] vs. response - variable slope function with Y = bottom + (X^hillslope)*(top-bottom)/(X^hillslope + EC50^hillslope). Bottom and top are minimal and maximal response, IC_{50} or EC_{50} is the concentration of ligands that gives a response halfway between top and bottom, and hillslope describes the steepness of curves.

Method for indole compounds quantification by LC-MS

Quantification of indole compounds was performed using the modified method previously published[70]. Individual *E. coli* colony was picked on an LB agar plate and inoculated into a 5 mL LB culture overnight. To extract, the overnight culture was centrifuged at 4000 g for 10 minutes at 4 °C, and the cell pellet and supernatant were separated. The pellet was washed with 2 mL of MilliQ water 3 times to remove residual compounds. Then 0.5 µL Lysonase Bioprocessing Reagent (MilliPore Sigma) was added to the pellet and incubate at room temperature for 5 minutes. The supernatant was transferred into another 15 mL conical tube, combined with lysed cells, mixed with 5 mL of ethyl acetate (FisherScientific, USA), and mixed vigorously. The mixture was then allowed to stand for the layers to separate. The extraction was repeated twice, and the organic layers from each extraction were combined. The organic layer was dried with anhydrous Na_2SO_4 (FisherScientific, USA) and evaporated to dryness using a rotary evaporator. The crude extract was dissolved in 1 mL LC-MS grade methanol (FisherScientific, USA), centrifuged to remove particles, filtered with PTFE membrane (FisherScientific, USA), and analyzed by LC-HRMS using SIM mode on Thermo Scientific Q-Exactive Orbitrap (ThermoFisher, USA). A gradient from 50-80% acetonitrile (FisherScientific, USA) over 30 min was used as a mobile phase using an Acclaim Polar Advantage II C18 column (Thermo Fisher Scientific, USA). For quantification of indole-3-aldehyde and indole, a standard

15

curve that correlates the area with the analyte concentration was used. To measure the extraction efficiency through this method, a known amount of analyte was added to water, followed by the above-described method. The efficiency is expressed as the amount calculated from the LC-MS quantification method over a known amount, which is $96.7 \pm 0.7\%$ for I3A and $97.6 \pm 0.3\%$ for IND. All biological samples were repeated at least three times as replicates.

Bioinformatics analysis

The alignment of HtlD[I3A-PAS] (UniProt: https://www.uniprot.org/uniprot/Q5V5P7)[53] was performed (12/2019) using XtalPred (http://xtalpred.godziklab.org/XtalPred-cgi/xtal.pl)[71, 72]. The alignment was queried against the UniProtKB database with default settings (Table S7). The result was downloaded and filtered to include 40% identity and 70% coverage, resulting in 25 proteins. The sequences of the 25 proteins were downloaded from the UniProtKB database. The alignment figure was generated using Jalview (http://www.jalview.org)[73] and WebLogo (http://weblogo.threeplusone.com)[74, 75].

16

ABBREVIATIONS

Abbreviation	Full name
5-HT	5-hydroxytryptamine
AhR	Aryl hydrocarbon receptor
CA	Catalytic domain
CNS	Central nervous system
DHp	Dimerization histidine phosphotransfer
DMSO	DMSO (Biological Grade)
HK	Histidine kinase
HtlD	HTR-like protein D
HTR	Halobacterial transducer of rhodopsin
I3A	Indole-3-aldehyde
I3AA	Indole-3-acetic acid
I3AM	3-aminoindole
I3BA	Indole-3-butyric acid
I3M	3-Methylindole
I3PA	indole-3-pyruvic acid
I3SP	3-Indoxyl sulfate
I5OH	5-Hydroxyindole
I7OH	7-Hydroxyindole
IACN	3-Indolylacetonitrile
IMI	1-Methylindole
IND	Indole
Kan	Kanamycin
Kyn	Kynurenine
LB	Luria-Bertani
MKN	Multi-kinase network
PAS	Per-Arnt-Sim
PDB	Protein database
RBS	Ribosomal binding site
REC	Receiver domain
RR	Response regulator

ACKNOWLEDGEMENT

AUTHOR CONTRIBUTIONS

W.S.C., C.Z., and J.W. conceptualized and designed the study. C.Z. and J.W. constructed plasmids and strains. C.Z. performed and collected preliminary data for I3A dose-response and ligands screening experiments on parent strains, pDawn-I3A, and pDusk-I3A_p1; J.W. performed the experiments and analyzed data. C. Z. performed and analyzed LC-MS experiments. W.S.C. and J.W. drafted the manuscript with input from C.Z. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

FIGURE LEGENDS

Figure 1. Design of an I3A biosensor.

(A) A diagram for a typical two-component system (TCSs). A TCS is composed of a sensor histidine kinase (HK) (grey) and a response regulator (RR) (purple). A conserved HK contains a dimerization histidine phosphotransfer (DHp) domain and a C-terminal ATP-binding catalytic domain (CA) domain. The RR includes a receiver (REC) domain and a DNA binding domain. The signal regulates autophosphorylation of a conserved histidine in the DHp domain. Further phosphotransfer from the DHp domain to the REC domain results in downstream gene regulation. (B) Design of an I3A sensing chimeric histidine kinase I3A-HK. YF1 is a chimeric protein that consists of a light-sensing PAS domain and FixL-HK domain, the conserved kinase core from FixL[56]. The HtlD$^{I3A-PAS}$ domain, an I3A binding PAS domain, is the second PAS domain in the HtlD as indicated (brown box). The chemical structure of I3A is shown (red). We swapped the light-sensing domain in YF1 to the HtlD$^{I3A-PAS}$ and generated the chimeric protein I3A-HK. (C) The construct pDusk-I3A_p0-p3 for I3A-repressed *mCherry* expression. I3A regulates the activity of I3A-HK in phosphorylating the FixJ RR, which stimulates transcription of *mCherry* at the *FixK2* promoter. I3A repressed expression of *mCherry*. (D) I3A dose-response curve for pDusk-I3A_p0-p3. Normalized fluorescent intensity to OD$_{600}$ was plotted against the concentration of I3A. See the curve fitting results in Table 1 for the three biological replicates. (E) The quantification of I3A extracted from culture media and saturated *E. coli* strains as determined by HPLC-MS. The I3A concentrations are plotted in the bar chart. Data represent the mean $\pm$ SD of at least 3 biological replicates. ns: not significant.

20

Figure 2. The I3A sensing through PAS domain.

(A) Structure of the I3A binding PAS domain, HtlD$^{I3A-PAS}$ (PDB: 3BWL) deposited by Osipiuk, J., Zhou, M., Freeman, L., and Joachimiak A. at the Midwest Center for Structural Genomics[52]. Dimer of HtlD$^{I3A-PAS}$ is shown (light orange and light blue). The structure of HtlD$^{I3A-PAS}$ is aligned to the light-sensing PAS domain of YF1 (PDB: 4GCZ) (grey). The signal transmission residues are located at C- terminal in the presented structure. The residues within 5 Å to the I3A are indicated (dashed box). (B) Conservation of the potential I3A binding residues are determined by sequence alignment and plotted using WebLogo[74, 75] as described in the method. (C) I3A binding site (D414A, D449A. F472A) and allosteric signal transmission residues (D-(A/I/V)-(T/S)-E as D507A-T509A) were mutated as indicated on the pDusk-I3A_p1 plasmid. Fold changes are plotted in log scale. Fold change is measured as the ratio of normalized mCherry intensity of I3A-treated vs. DMSO-treated. Ratio equals to 1 is labeled (red dash line). Data represent the mean $\pm$ SD of at least 3 biological replicates. *: $p < 0.05$, comparing I3A-treated vs. DMSO-treated samples.

21

Figure 3. The selectivity of the I3A sensor.

(A) Structure of indole derivatives that were used in the selectivity tests. I3A: indole-3-aldehyde, I3M: 3-methylindole, I3AM: 3-aminoindole, I3AA: Indole-3-acetic acid, I3PA: indole-3-pyruvic acid, I3BA: Indole-3-butyric acid, IACN: 3-Indoleacetonitrile, I3SP: indoxyl sulfate, IND: indole, I5OH: 5-Hydroxyindole, I7OH: 7-Hydroxyindole, IMI: 1-Methylindole (B) The selectivity of pDusk-I3A_p3. The fold-change is reported as a ratio of normalized mCherry intensity for ligand-treated vs. DMSO-treated samples. The fold-change is plotted as a heat map given the gradient as shown in the figure, with a max (1.1 fold, black) and a min (0 fold, red). Data represent the mean of at least 3 biological replicates. DMSO: dimethyl sulfoxide. See the corresponding bar charts and the results of the test for the pDawn-I3A in Figure S3A.

Figure 4. I3A detection is conserved amongst homologs of HtlD$^{\text{I3A-PAS}}$.

(A) Response to I3A of the homologs of HtlD$^{\text{I3A-PAS}}$. To test the homologs, we swapped the HtlD$^{\text{I3A-PAS}}$ in pDusk-I3A_p3 to each homolog and generated constructs for chimeric PAS-HKs. See the sequences of homologs in Table S3. The fold-change is reported as a ratio of normalized mCherry intensity for DMSO-treated vs. I3A-treated samples. Data represent the mean $\pm$ SD of at least 3 biological replicates. The mean is indicated in a column for each homolog. See a full phylogenetic tree in Figure S4A. (B) The selectivity of PAS-2. The fold-change is reported as a ratio of normalized mCherry intensity for ligand-treated vs. DMSO-treated samples. Data represent the mean $\pm$ SD of at least 3 biological replicates. The conditions where ligand-treated vs. DMSO-treated gives a $p > 0.05$ are indicated with ns. The ligands that give a >1.5-fold change in mCherry expression are I3A (red) and I3AA (yellow). DMSO: dimethyl sulfoxide. See the results for the selectivity of PAS-1 and PAS-3 in Figure S4.

TABLE LEGENDS

Table 1. Fitting for the dose-response curves.

Each curve is fitted to the Hill equation with default settings in Prism9. Bottom: predicted minimal mCherry; Top: predicted maximal mCherry. Standard deviation estimates are in parentheses. The mean and SD are reported in the table for at least 3 biological replicates.

ASSOCIATED CONTENT

Supporting Information:

Figure S1-S4: Structure of $HtlD^{I3A-PAS}$ and design of chimera I3A-HK, construction of I3A sensors, the selectivity of the I3A sensors, quantification of I3A and IND in the cell culture using LC-MS, and indole metabolite stimulation of homologs of $HtlD^{I3A-PAS}$. Tables S1-S8: PDB PAS domains with ligands, promoters used for constructs, selected homologs of $HtlD^{I3A-PAS}$, DNA oligos used in the study, Gibson cloning strategies for plasmids, *E.coli* strains used in the study, BLAST results for $HtlD^{I3A-PAS}$, and chemicals used in the study.
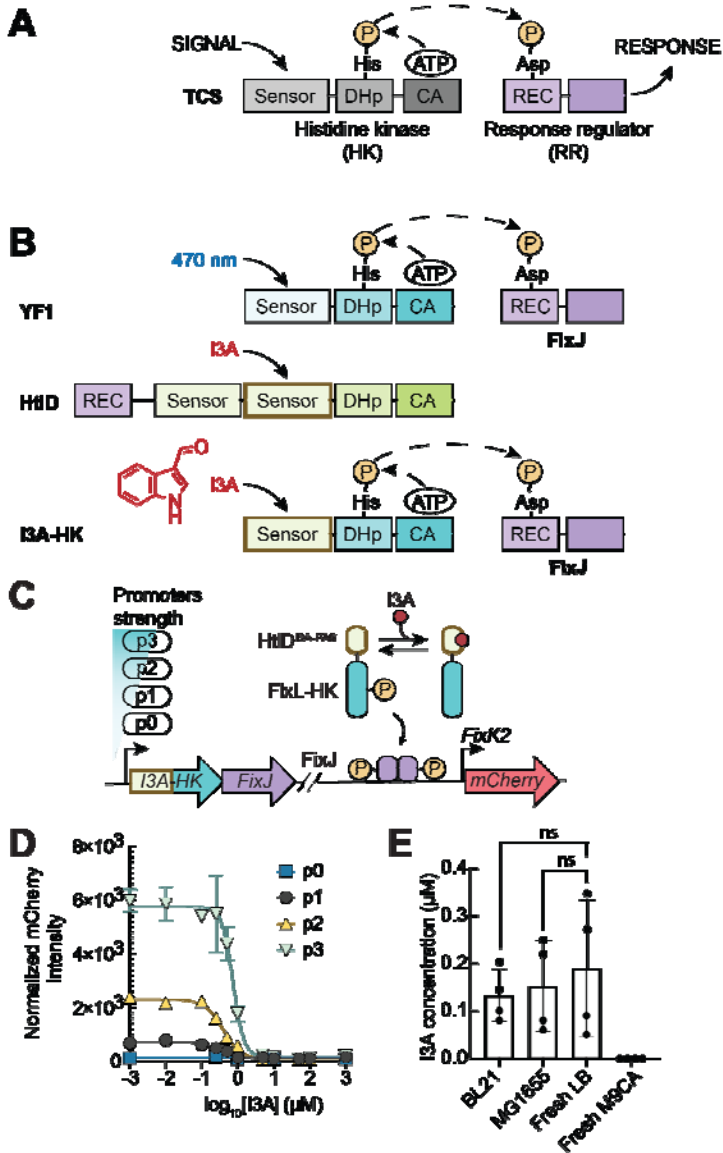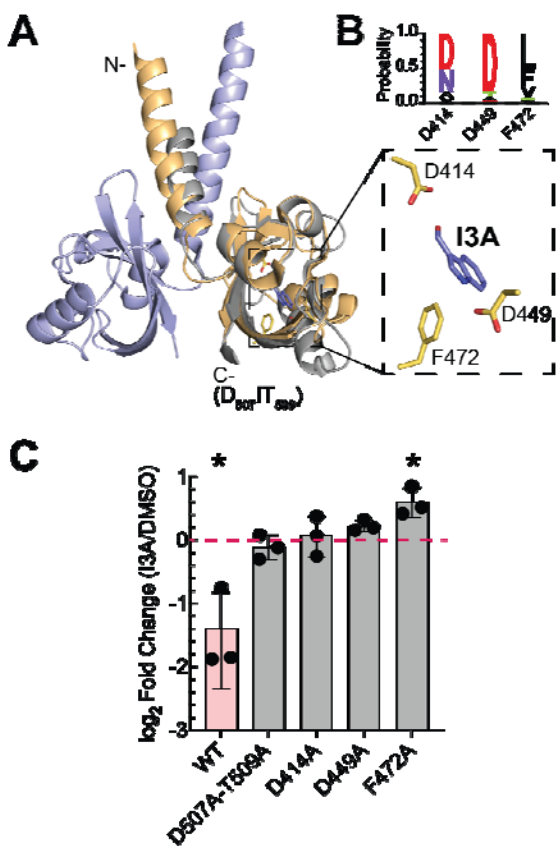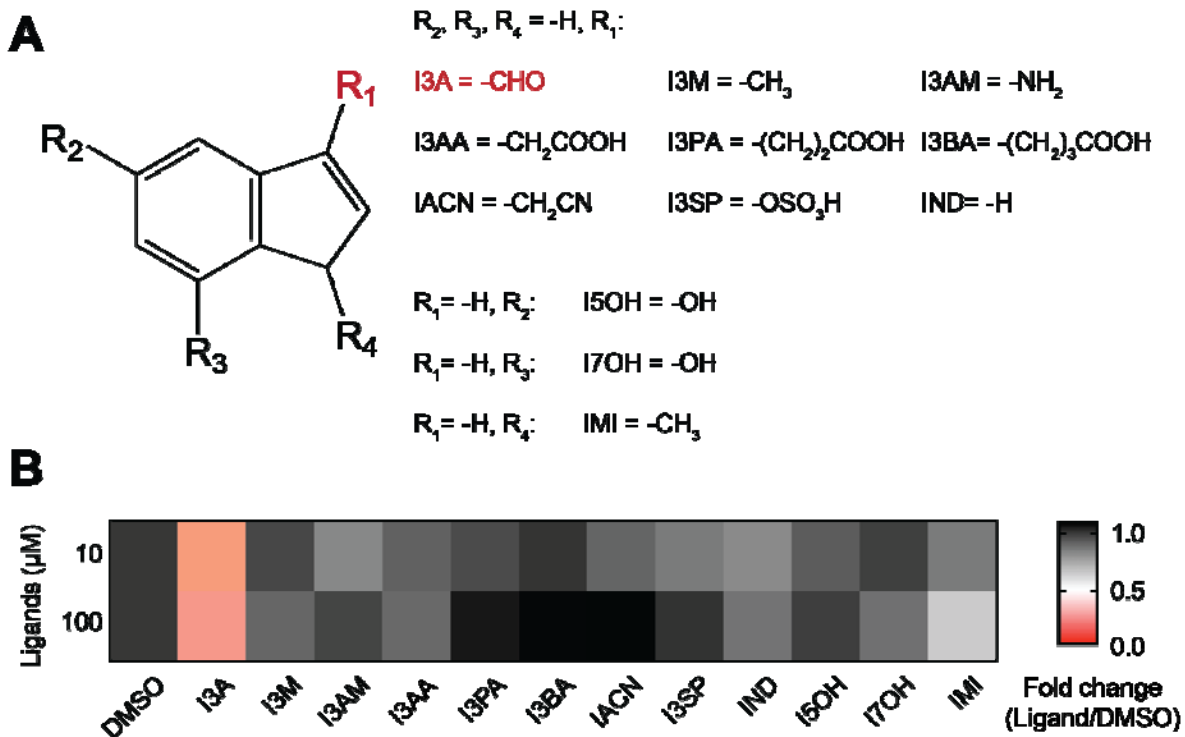
Figures

Figure 1.

Figure 2.

Figure 3.

A

$R_2, R_3, R_4 = -H$, R_1:

R_1

R_2

I3A = -CHO   I3M = -CH$_3$ I3AM = -NH$_2$

I3AA = -CH$_2$COOH   I3PA = -(CH$_2$)$_2$COOH I3BA= -(CH$_2$)$_3$COOH

IACN = -CH$_2$CN   I3SP = -OSO$_3$H   IND= -H

R_1= -H, R_2:   I5OH = -OH

R_1= -H, R_3:   I7OH = -OH

R_1= -H, R_4:   IMI = -CH$_3$

R_3 R_4

B

Figure 4.
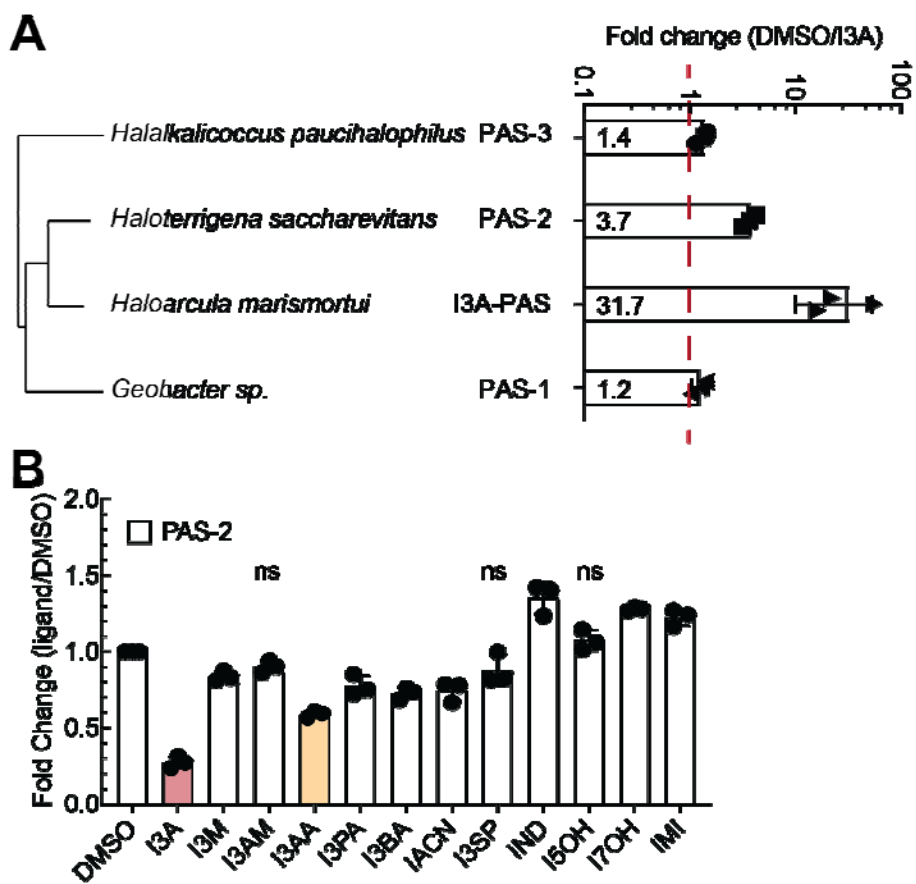
8

Table 1.

strain	Bottom	Top	IC50/EC50 (µM)	Hill Slope	fold change
pDawn	2959 (497)	33976 (2328)	11.5 (7.8)	0.7 (0.3)	11.8 (3)
pDusk_p1	116 (26)	739 (128)	0.3 (0.1)	-1.8 (0.6)	6.4 (0.5)
pDusk_p2	94 (11)	2325 (107)	0.4 (0.1)	-1.8 (0.4)	25.0 (4.2)
pDusk_p3	168 (40)	5804 (475)	0.7 (0.1)	-2.7 (1)	35.5 (6.4)

References:

[1] Crook, N., Ferreiro, A., Gasparrini, A. J., Pesesky, M. W., Gibson, M. K., Wang, B., Sun, X., Condiotte, Z., Dobrowolski, S., Peterson, D., and Dantas, G. (2019) Adaptive strategies of the candidate probiotic *E. coli* Nissle in the mammalian gut, *Cell Host Microbe 25*, 499-512 e498.

[2] Riglar, D. T., and Silver, P. A. (2018) Engineering bacteria for diagnostic and therapeutic applications, *Nat Rev Microbiol 16*, 214-225.

[3] Daeffler, K. N., Galley, J. D., Sheth, R. U., Ortiz-Velez, L. C., Bibb, C. O., Shroyer, N. F., Britton, R. A., and Tabor, J. J. (2017) Engineering bacterial thiosulfate and tetrathionate sensors for detecting gut inflammation, *Mol Syst Biol 13*, 923.

[4] Lee, J.-H., Wood, T. K., and Lee, J. (2015) Roles of indole as an interspecies and interkingdom signaling molecule, *Trends Microbiol 23*, 707-718.

[5] Agus, A., Planchais, J., and Sokol, H. (2018) Gut microbiota regulation of tryptophan metabolism in health and disease, *Cell Host Microbe 23*, 716-724.

[6] Yano, J. M., Yu, K., Donaldson, G. P., Shastri, G. G., Ann, P., Ma, L., Nagler, C. R., Ismagilov, R. F., Mazmanian, S. K., and Hsiao, E. Y. (2015) Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis, *Cell 161*, 264-276.

[7] Clarke, G., McKernan, D. P., Gaszner, G., Quigley, E. M., Cryan, J. F., and Dinan, T. G. (2012) A distinct profile of tryptophan metabolism along the kynurenine pathway downstream of Toll-like receptor activation in irritable bowel syndrome, *Front Pharmacol 3*, 90.

[8] Zelante, T., Iannitti, R. G., Cunha, C., De Luca, A., Giovannini, G., Pieraccini, G., Zecchi, R., D'Angelo, C., Massi-Benedetti, C., Fallarino, F., Carvalho, A., Puccetti, P., and Romani, L. (2013) Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22, *Immunity 39*, 372-385.

[9] Scott, S. A., Fu, J., and Chang, P. V. (2020) Microbial tryptophan metabolites regulate gut barrier function via the aryl hydrocarbon receptor, *Proc Natl Acad Sci USA 117*, 19376-19387.

[10] Cervantes-Barragan, L., Chai, J. N., Tianero, M. D., Di Luccia, B., Ahern, P. P., Merriman, J., Cortez, V. S., Caparon, M. G., Donia, M. S., Gilfillan, S., Cella, M., Gordon, J. I., Hsieh, C. S., and Colonna, M. (2017) *Lactobacillus reuteri* induces gut intraepithelial CD4(+)CD8alphaalpha(+) T cells, *Science 357*, 806-810.

[11] Yu, J., Luo, Y., Zhu, Z., Zhou, Y., Sun, L., Gao, J., Sun, J., Wang, G., Yao, X., and Li, W. (2019) A tryptophan metabolite of the skin microbiota attenuates inflammation in patients with atopic dermatitis through the aryl hydrocarbon receptor, *J Allergy Clin Immunol 143*, 2108-2119.e2112.

[12] Rothhammer, V., Mascanfroni, I. D., Bunse, L., Takenaka, M. C., Kenison, J. E., Mayo, L., Chao, C. C., Patel, B., Yan, R., Blain, M., Alvarez, J. I., Kebir, H., Anandasabapathy, N., Izquierdo, G., Jung, S., Obholzer, N., Pochet, N., Clish, C. B., Prinz, M., Prat, A., Antel, J., and Quintana, F. J. (2016) Type I interferons and microbial metabolites of tryptophan modulate astrocyte activity and central nervous system inflammation via the aryl hydrocarbon receptor, *Nat Med 22*, 586-597.

[13] Powell, D. N., Swimm, A., Sonowal, R., Bretin, A., Gewirtz, A. T., Jones, R. M., and Kalman, D. (2020) Indoles from the commensal microbiota act via the AHR and IL-10 to tune the

cellular composition of the colonic epithelium during aging, *Proc Natl Acad Sci USA 117*, 21519-21526.

[14] Bommarius, B., Anyanful, A., Izrayelit, Y., Bhatt, S., Cartwright, E., Wang, W., Swimm, A. I., Benian, G. M., Schroeder, F. C., and Kalman, D. (2013) A family of indoles regulate virulence and Shiga toxin production in pathogenic *E. coli*, *PLoS One 8*, e54456.

[15] Hicks, M., Bachmann, T. T., and Wang, B. (2020) Synthetic biology enables programmable cell-based biosensors, *ChemPhysChem 21*, 132-144.

[16] Naydich, A. D., Nangle, S. N., Bues, J. J., Trivedi, D., Nissar, N., Inniss, M. C., Niederhuber, M. J., Way, J. C., Silver, P. A., and Riglar, D. T. (2019) Synthetic gene circuits enable systems-level biosensor trigger discovery at the host-microbe interface, *MSystems 4*, e00125-00119.

[17] Hubbard, T. D., Murray, I. A., Bisson, W. H., Lahoti, T. S., Gowda, K., Amin, S. G., Patterson, A. D., and Perdew, G. H. (2015) Adaptation of the human aryl hydrocarbon receptor to sense microbiota-derived indoles, *Sci Rep 5*, 12689.

[18] Hirakawa, H., Inazumi, Y., Masaki, T., Hirata, T., and Yamaguchi, A. (2005) Indole induces the expression of multidrug exporter genes in *Escherichia coli*, *Mol. Microbiol. 55*, 1113-1126.

[19] Boon, N., Kaur, M., Aziz, A., Bradnick, M., Shibayama, K., Eguchi, Y., and Lund, P. A. (2020) The signaling molecule indole inhibits induction of the AR2 acid resistance system in *Escherichia coli*, *Front Microbiol 11*, 474.

[20] Zarkan, A., Liu, J., Matuszewska, M., Gaimster, H., and Summers, D. K. (2020) Local and universal action: the paradoxes of indole signalling in bacteria, *Trends Microbiol 28*, 566-577.

[21] Gaimster, H., Cama, J., Hernández-Ainsa, S., Keyser, U. F., and Summers, D. K. (2014) The indole pulse: a new perspective on indole signalling in *Escherichia coli*, *PLoS One 9*, e93168.

[22] Mann, T. H., Seth Childers, W., Blair, J. A., Eckart, M. R., and Shapiro, L. (2016) A cell cycle kinase with tandem sensory PAS domains integrates cell fate cues, *Nat Commun 7*, 11454.

[23] Porter, S. L., Wadhams, G. H., and Armitage, J. P. (2011) Signal processing in complex chemotaxis pathways, *Nat Rev Microbiol 9*, 153-165.

[24] Henke, J. M., and Bassler, B. L. (2004) Three parallel quorum-sensing systems regulate gene expression in *Vibrio harveyi*, *J Bacteriol 186*, 6902-6914.

[25] Beier, D., and Gross, R. (2006) Regulation of bacterial virulence by two-component systems, *Curr Opin Microbiol 9*, 143-152.

[26] Kowallis, K. A., Duvall, S. W., Zhao, W., and Childers, W. S. (2020) Manipulation of bacterial signaling using engineered histidine kinases, In *Histidine Phosphorylation*, pp 141-163, Springer.

[27] Krell, T., Lacal, J., Busch, A., Silva-Jimenez, H., Guazzaroni, M. E., and Ramos, J. L. (2010) Bacterial sensor kinases: diversity in the recognition of environmental signals, *Annu Rev Microbiol 64*, 539-559.

[28] Schmidl, S. R., Sheth, R. U., Wu, A., and Tabor, J. J. (2014) Refactoring and optimization of light-switchable *Escherichia coli* two-component systems, *ACS Synth Biol 3*, 820-831.

[29] Ramakrishnan, P., and Tabor, J. J. (2016) Repurposing *Synechocystis* PCC6803 UirS–UirR as a UV-violet/green photoreversible transcriptional regulatory tool in *E. coli*, *ACS Synth Biol 5*, 733-740.

[30] Ong, N. T., Olson, E. J., and Tabor, J. J. (2018) Engineering an *E. coli* near-infrared light sensor, *ACS Synth Biol 7*, 240-248.

[31] Ong, N. T., and Tabor, J. J. (2018) A miniaturized *Escherichia coli* green light sensor with high dynamic range, *ChemBioChem 19*, 1255-1258.

[32] Olson, E. J., Hartsough, L. A., Landry, B. P., Shroff, R., and Tabor, J. J. (2014) Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals, *Nat methods 11*, 449.

[33] Riglar, D. T., Giessen, T. W., Baym, M., Kerns, S. J., Niederhuber, M. J., Bronson, R. T., Kotula, J. W., Gerber, G. K., Way, J. C., and Silver, P. A. (2017) Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation, *Nat. Biotechnol 35*, 653-658.

[34] Ganesh, I., Ravikumar, S., Yoo, I. K., and Hong, S. H. (2015) Construction of malate-sensing *Escherichia coli* by introduction of a novel chimeric two-component system, *Bioprocess Biosyst Eng 38*, 797-804.

[35] Ganesh, I., Ravikumar, S., Lee, S. H., Park, S. J., and Hong, S. H. (2013) Engineered fumarate sensing *Escherichia coli* based on novel chimeric two-component system, *J Biotechnol 168*, 560-566.

[36] Ravikumar, S., David, Y., Park, S. J., and Choi, J. I. (2018) A chimeric two-component regulatory system-based *Escherichia coli* biosensor engineered to detect glutamate, *Appl Biochem Biotechnol 186*, 335-349.

[37] Selvamani, V., Ganesh, I., Maruthamuthu, M. K., Eom, G. T., and Hong, S. H. (2017) Engineering chimeric two-component system into *Escherichia coli* from *Paracoccus denitrifican*s to sense methanol, *Biotechnol. Bioprocess Eng. 22*, 225-230.

[38] A, P. T., and Fussenegger, M. (2019) Engineering mammalian cells for disease diagnosis and treatment, *Curr Opin Biotechnol 55*, 87-94.

[39] Schmidl, S. R., Ekness, F., Sofjan, K., Daeffler, K. N., Brink, K. R., Landry, B. P., Gerhardt, K. P., Dyulgyarov, N., Sheth, R. U., and Tabor, J. J. (2019) Rewiring bacterial two-component systems by modular DNA-binding domain swapping, *Nat Chem Biol 15*, 690-698.

[40] Galperin, M. Y., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2018) Phyletic distribution and lineage-specific domain architectures of archaeal two-component signal transduction systems, *J Bacteriol 200*.

[41] Levskaya, A., Chevalier, A. A., Tabor, J. J., Simpson, Z. B., Lavery, L. A., Levy, M., Davidson, E. A., Scouras, A., Ellington, A. D., Marcotte, E. M., and Voigt, C. A. (2005) Synthetic biology: engineering *Escherichia coli* to see light, *Nature 438*, 441-442.

[42] Wang, B., Barahona, M., Buck, M., and Schumacher, J. (2013) Rewiring cell signalling through chimaeric regulatory protein engineering, *Biochem Soc Trans 41*, 1195-1200.

[43] Papon, N., and Stock, A. M. (2019) Two-component systems, *Curr. Biol. 29*, R724-R725.

[44] Gao, R., and Stock, A. M. (2009) Biological insights from structures of two-component proteins, *Annu Rev Microbiol 63*, 133-154.

[45] Gao, R., Mack, T. R., and Stock, A. M. (2007) Bacterial response regulators: versatile regulatory strategies from common domains, *Trends Biochem Sci 32*, 225-234.

[46] Whitaker, W. R., Davis, S. A., Arkin, A. P., and Dueber, J. E. (2012) Engineering robust control of two-component system phosphotransfer using modular scaffolds, *Proc Natl Acad Sci USA 109*, 18090.

[47] Francis, V. I., and Porter, S. L. (2019) Multikinase networks: two-component signaling networks integrating multiple stimuli, *Annu Rev Microbiol 73*, 199-223.

[48] Park, D. M., and Taffet, M. J. (2019) Combinatorial sensor design in *Caulobacter crescentus* for selective environmental uranium detection, *ACS Synth Biol 8*, 807-817.

[49] Hansen, J., Mailand, E., Swaminathan, K. K., Schreiber, J., Angelici, B., and Benenson, Y. (2014) Transplantation of prokaryotic two-component signaling pathways into mammalian cells, *Proc Natl Acad Sci USA 111*, 15705-15710.

[50] Maze, A., and Benenson, Y. (2019) Artificial signaling in mammalian cells enabled by prokaryotic two-component system, *Nat Chem Biol 16*, 179-187.

[51] Behnsen, J., and Raffatellu, M. (2013) Keeping the peace: aryl hydrocarbon receptor signaling modulates the mucosal microbiota, *Immunity 39*, 206-207.

[52] Osipiuk, J., Zhou, M., Freeman, L., Joachimiak, A., Midwest Center for Structural Genomics (MCSG). (2008) Crystal structure of PAS domain of HTR-like protein from *Haloarcula marismortui*.

[53] UniProt, C. (2019) UniProt: a worldwide hub of protein knowledge, *Nucleic Acids Res 47*, D506-D515.

[54] Ohlendorf, R., Vidavski, R. R., Eldar, A., Moffat, K., and Moglich, A. (2012) From dusk till dawn: one-plasmid systems for light-regulated gene expression, *J Mol Biol 416*, 534-542.

[55] Krueger, B., Friedrich, T., Förster, F., Bernhardt, J., Gross, R., and Dandekar, T. (2012) Different evolutionary modifications as a guide to rewire two-component systems, *Bioinform. Biol. Insights 6*, BBI. S9356.

[56] Moglich, A., Ayers, R. A., and Moffat, K. (2009) Design and signaling mechanism of light-regulated histidine kinases, *J Mol Biol 385*, 1433-1444.

[57] Dubey, B. N., Lori, C., Ozaki, S., Fucile, G., Plaza-Menacho, I., Jenal, U., and Schirmer, T. (2016) Cyclic di-GMP mediates a histidine kinase/phosphatase switch by noncovalent domain cross-linking, *Sci Adv 2*, e1600823.

[58] Ptashne, M. (2004) *A genetic switch: phage lambda revisited*, CSHL press.

[59] Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P., and Endy, D. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements, *Nat Methods 10*, 354-360.

[60] Moglich, A., and Moffat, K. (2007) Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA, *J Mol Biol 373*, 112-126.

[61] Engelhard, C., Diensthuber, R. P., Moglich, A., and Bittl, R. (2017) Blue-light reception through quaternary transitions, *Sci Rep 7*, 1385.

[62] Diensthuber, R. P., Bommer, M., Gleichmann, T., and Moglich, A. (2013) Full-length structure of a sensor histidine kinase pinpoints coaxial coiled coils as signal transducers and modulators, *Structure 21*, 1127-1136.

[63] Duvall, S. W., and Seth Childers, W. (2020) Design of a histidine kinase FRET sensor to detect complex signal integration within living bacteria, *ACS sensors 5*, 1589-1596.

[64] Kowallis, K. A., Silfani, E. M., Kasumu, A. P., Rong, G., So, V., and Seth Childers, W. (2020) Synthetic control of signal flow within a bacterial multi-kinase network, *ACS Synth Biol 9*, 1705-1713.

[65] Megaw, J., and Gilmore, B. F. (2017) Archaeal persisters: persister cell formation as a stress response in *Haloferax volcanii, Front Microbiol 8*.

[66] Fischbach, M. A., Bluestone, J. A., and Lim, W. A. (2013) Cell-based therapeutics: the next pillar of medicine, *Sci Transl Med 5*, 179ps177.

[67] Lee, J., Attila, C., Cirillo, S. L., Cirillo, J. D., and Wood, T. K. (2009) Indole and 7-hydroxyindole diminish *Pseudomonas aeruginosa* virulence, *Microb Biotechnol 2*, 75-90.

[68] Hillson, N. J., Rosengarten, R. D., and Keasling, J. D. (2012) j5 DNA assembly design automation software, *ACS Synth Biol 1*, 14-21.

[69] Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nat methods 6*, 343-345.

[70] Ittiamornkul, K., Zhu, Q., Gkotsi, D. S., Smith, D. R. M., Hillwig, M. L., Nightingale, N., Goss, R. J. M., and Liu, X. (2015) Promiscuous indolyl vinyl isonitrile synthases in the biogenesis and diversification of hapalindole-type alkaloids, *Chem Sci 6*, 6836-6840.

[71] Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I. A., Lesley, S. A., and Godzik, A. (2007) XtalPred: a web server for prediction of protein crystallizability, *Bioinformatics 23*, 3403-3405.

[72] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, *J Mol Biol 215*, 403-410.

[73] Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009) Jalview Version 2 - a multiple sequence alignment editor and analysis workbench, *Bioinformatics 25*, 1189-1191.

[74] Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator, *Genome Res 14*, 1188-1190.

[75] Schneider, T. D., and Stephens, R. M. (1990) Sequence logos: a new way to display consensus sequences, *Nucleic Acids Res 18*, 6097-6100.