

(43) International Publication Date
17 November 2016 (17.11.2016)

- (51) International Patent Classification:
C12Q 1/04 (2006.01) C12N 1/21 (2006.01)
- (21) International Application Number:
PCT/US2016/030831
- (22) International Filing Date:
4 May 2016 (04.05.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
62/157,293 5 May 2015 (05.05.2015) US
- (71) Applicant: WILLIAM MARSH RICE UNIVERSITY
[US/US]; 6100 Main Street, Houston, TX 77005 (US).
- (72) Inventors: DYULGYAROV, Nikola; 6100 Main Street, Ms-142, Houston, TX 77005 (US). SCHMIDL, Sebastian; 6100 Main Street, Ms-142, Houston, TX 77005 (US). EKNES, Felix; 6100 Main Street, Ms-142, Houston, TX 77005 (US). SHETH, Ravi; 6100 Main Street, Ms-142, Houston, TX 77005 (US). LANDRY, Brian; 6100 Main Street, Ms-142, Houston, TX 77005 (US). TABOR, Jeffrey, J.; 6100 Main Street, Ms-142, Houston, TX 77005 (US).
- (74) Agents: BOULWARE, Margaret et al.; Boulware & Valoir, Three Riverway, Suite 950, Houston, TX 77056 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LI, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: IDENTIFYING LIGANDS FROM BACTERIAL SENSORS

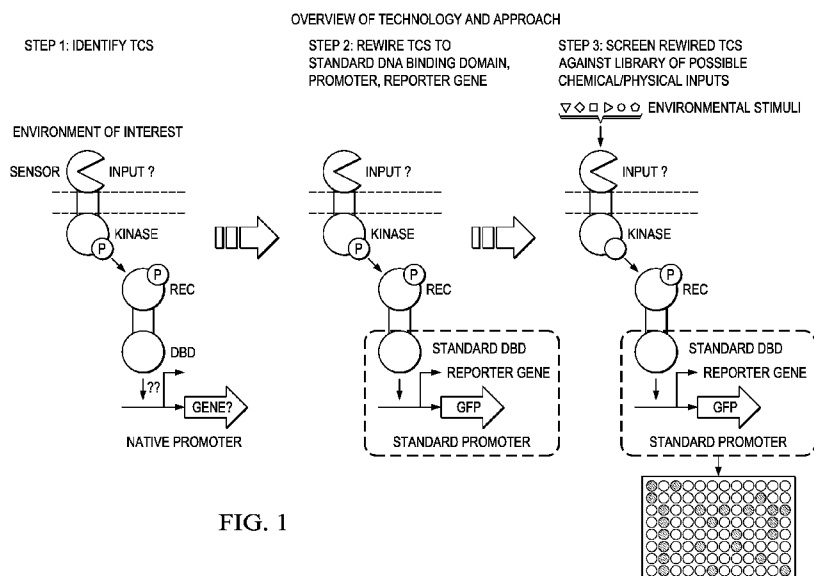


FIG. 1

(57) Abstract: Methods to create two component signal transduction systems by replace the DNA binding domains and output promoters in bacteria are described.

IDENTIFYING LIGANDS FROM BACTERIAL SENSORS

PRIOR RELATED APPLICATIONS

[0001] This application claims priority to US Serial No. 62/157,293, IDENTIFYING LIGANDS FROM BACTERIAL SENSORS, filed May 5, 2015 and incorporated by
5 reference herein in its entirety for all purposes.

FEDERALLY SPONSORED RESEARCH STATEMENT

[0002] This invention was made with government support under N00014-14-1-0487 awarded by the Office of Naval Research. The government has certain rights in the invention.

10

FIELD OF THE DISCLOSURE

[0003] The invention is a method to replace the DNA binding domains and output promoters bacterial two component signal transduction systems (a.k.a. two component systems, two component sensors, sensors, TCSs). The method enables TCSs to be transferred between different bacterial species despite incompatibilities that otherwise ‘silence’ their
15 ability to respond to inputs. The method can also be used to identify the inputs of novel or poorly characterized TCSs by transporting the TCSs from their native bacteria to non-native laboratory strains, encoding reporter genes as outputs, and performing screens wherein outputs are measured in the presence of different possible input signals. The method also enables TCSs to be engineered to function as biosensors with numerous applications in
20 medicine, industry, and basic science.

BACKGROUND OF THE DISCLOSURE

[0004] In the field of molecular biology, a two-component system serves as a basic stimulus-response coupling mechanism to allow organisms to sense and respond to changes in many different environmental conditions. A TCS typically consists of a membrane-bound
25 sensor histidine kinase (SK) that senses a specific environmental stimulus and a corresponding response regulator (RR) that mediates the cellular response, mostly through

differential expression of target genes. TCSs are widespread and highly evolutionarily diversified in the genomes of prokaryotes, but only a few TCSs have been identified in eukaryotic organisms.

[0005] Typically, signal transduction occurs through the transfer of phosphoryl groups from adenosine triphosphate (ATP) to a conserved histidine residue in the SK. This is an autophosphorylation reaction. The RRs are phosphorylated on a conserved aspartate residue and are protein phosphatases for the SKs. Phosphorylation causes a change in the RR conformation, usually activating an attached output domain, which then leads to the activation (or repression) of transcription of target genes. The phosphorylation state of an RR thereby controls its activity. Some SKs are bi-functional, catalyzing both the phosphorylation and dephosphorylation of their cognate RR. Inputs can regulate either the kinase or phosphatase activity of the bi-functional SK.

[0006] Signal transduction can also occur in a phosphorylation-independent manner. For example the SK can sequester the RR at the membrane in the absence of input. In the presence of input, the SK may then release the RR, which may then bind DNA and activate or repress transcription.

[0007] Two-component systems enable bacteria to sense, respond, and adapt to a wide range of environments, stressors, and growth conditions. Some bacteria can contain up to as many as 200 TCSs that have tight molecular specificity to prevent unwanted cross-talk. These pathways have been adapted to respond to a wide variety of stimuli, including nutrients, cellular redox state, changes in osmolarity, quorum sensing signals, antibiotics, temperature, chemoattractants, pH and more.

[0008] A few examples are provided:

[0009] In *Escherichia coli*, the EnvZ-OmpR osmoregulation system controls the differential expression of the outer membrane porin proteins OmpF and OmpC.

[0010] The KdpD-KdpE two-component regulatory system regulates the *kdpFABC* operon responsible for potassium transport in bacteria including *E. coli* and *Clostridium acetobutylicum*. The N-terminal domain of KdpD forms part of the cytoplasmic region of the protein, which may be the sensor domain responsible for sensing turgor pressure.

[0011] One variation of the two-component system is the phospho-relay system. In this system, a hybrid SK autophosphorylates and then transfers the phosphoryl group to an internal receiver domain, rather than to a separate RR protein. The phosphoryl group is then shuttled to histidine phosphotransferase (HPT) and subsequently to a terminal RR, which can
5 evoke the desired response.

[0012] Signal transducing SKs are the key elements in TCSs. Examples of SKs are EnvZ, which plays a central role in osmoregulation, and CheA, which plays a central role in the chemotaxis system.

[0013] SKs usually have an N-terminal ligand-binding domain and a C-terminal
10 kinase domain, but other domains may also be present. The N-terminal domain may also be a protein-protein interaction domain that enables activation of the SK by interaction with a third protein that binds the input, or another type of sensory domain. The kinase domain is responsible for the autophosphorylation of the histidine with ATP, the phosphotransfer from the kinase to an aspartate of the RR, and (with bi-functional enzymes) the phosphotransfer
15 from aspartyl phosphate back to water. The kinase core has a unique fold, distinct from that of the Ser/Thr/Tyr kinase superfamily.

[0014] SKs can be roughly divided into two classes: orthodox and hybrid. Most orthodox SKs, typified by the *E. coli* EnvZ protein, function as periplasmic membrane receptors and have a signal peptide and transmembrane segment(s) that separate the protein
20 into a periplasmic N-terminal sensing domain and a highly conserved cytoplasmic C-terminal kinase core. Members of this family, however, have an integral membrane sensor domain. Not all orthodox kinases are membrane bound, e.g., the nitrogen regulatory kinase NtrB (GlnL) is a soluble cytoplasmic SK.

[0015] Hybrid SKs contain multiple phosphodonor and phosphoacceptor sites and use
25 multi-step phospho-relay schemes instead of promoting a single phosphoryl transfer. In addition to the sensor domain and kinase core, they contain a CheY-like receiver domain and a His-containing phosphotransfer (HPT) domain.

[0016] TCSs are highly evolutionarily diversified and have been shown to sense
30 diverse chemical and physical inputs such as ions, sugars, polysaccharides, oxygen, antimicrobial peptides, human hormones, light, and so on. TCSs also regulate a wide range of

different gene outputs and theoretically can regulate any gene output. In addition, a single bacterial cell can have hundreds of different TCSs, enabling multiplexed sensing.

[0017] It is trivial to identify TCSs from bacterial genome sequences by computational methods, such as homology and/or domain searching and the fact that most SKs reside <1000 base pairs from the RRs with which they communicate (i.e. their cognate RRs). However, such TCSs typically sense **unknown** inputs and control **unknown** output genes. In particular, output genes are more difficult to identify from genome sequences because they do not have a conserved sequence or domain structure that can be predicted from the sequence of the SK or RR. Moreover, output genes more often reside greater distances away from the SK and RR on the genome, making their identification difficult. Because output genes cannot be easily identified, they cannot be easily measured in response to different possible input signals, making inputs difficult to identify. Furthermore, the microbes containing most TCSs are un-culturable or difficult to genetically manipulate in the laboratory, further making inputs difficult to identify. Moreover, many TCSs are recalcitrant to transfer from their native strains to non-native bacteria such as laboratory strains due to incompatibilities between the transcription regulating DNA binding domain or output promoter and the transcriptional machinery (e.g. RNA Polymerase) of the non-native strain. Therefore, while TCSs have tremendous medical, industrial and basic research applications, various technical challenges have kept them from being fully exploited.

[0018] There have been a very small number of limited publications where small scale DNA binding domain (DBD) replacement has been shown for a small number of TCSs. However, those reports do **not** use DBD replacement to identify the inputs of the TCSs, but rather to study the fundamental structural and functional biology of the RR. These studies also do not demonstrate DBD replacement for multiple members of the OmpR-PhoB family, nor any members of the NarL-FixJ family, nor identify general primary or secondary structural “rules” for rewiring the DBD of those structural sub-families, as we do. They also do not use the method to transport TCSs to new species of bacteria where they can be studied without complicating and potentially obfuscating cross-regulation from poorly understood native regulatory networks. Importantly, they do not use the method to engineer biosensors.

[0019] In a related, but fundamentally different approach, Skerker et al. (2008) swap the SK–RR interaction interface of very closely related TCSs. The SK-RR interface is a

different site than is rewired herein. Their data show molecular interaction and phosphotransfer activity *in vitro* and *in vivo*, but they do not show that the rewired TCS responds to an input. Thus, in direct contrast to the rewired TCSs described herein, the rewired TCSs in Skerker are not fully functional. Accordingly, one could not utilize the Skerker technology to identify the input of a TCS. Cheng (2014) takes a similar approach to Skerker, but all results are theoretical, not experimental. Furthermore, their focus is on the nature of the molecular interactions rather than the applications of rewired TCSs.

[0020] Wang et al. (2013) describes an idea for DBD rewiring for the NtrC subfamily of RRs, not the OmpR-PhoB and NarL-FixJ families. OmpR-PhoB and NarL-FixJ are much more widespread than NtrC, thus our method has broader scope for our applications. However, Wang et al. (2013) show no data, and thus there is no reasonable expectation of success in the absence of any proof of concept. Finally, Wang et al. claim that one should “design linkers” (i.e. non-native linkers) between REC and DBD. Non-native or exogenous linkers are not needed herein—we use only the native linker sequences. Stated another way, we literally cut and paste regions of existing RR sequences together, we add no unnatural sequences to our chimeric RRs. Wang also specifically states they are not interested in engineering “one input/one output” TCSs. On the other hand, we are precisely interested in engineering “one input/one output” TCSs. Furthermore, they do not say their method could be used to move TCSs between species, nor that their method could be used to identify the ligands sensed by TCSs. Rather, they want to use TCSs with known signals to control non-natural genes in order to study the natural genes that are regulated by the TCSs.

[0021] There are examples of rewiring in the OmpR-PhoB family. For example, Allen (2001) replaces the REC domain of the OmpR-PhoB family *E. coli* RR PhoB with that of the chemotaxis responsive RR CheY. They use mutant *E. coli* strains where the chemotaxis pathway is strongly activated or strongly de-activated to show that one of the CheY-PhoB chimeras activates a PhoB output promoter only in the former strain.

[0022] Walthers (2003) also construct chimeras between the REC domains of OmpR and PhoB and the DBDs of OmpR and PhoB. PhoB-OmpR chimeras fail to activate gene expression, but OmpR-PhoB chimeras did activate gene expression.

[0023] Howell (2003) rewires the DNA binding domain (DBD) of PhoP (OmpR-PhoB family) with that of YycF (OmpR-PhoB), both from *B. subtilis*, and all experiments are

in *B. subtilis*. In this experiment, they replace the DBD of a TCS with a known input and output, with the DBD of a TCS with an unknown output. Then they use the known input for TCS 1 to identify the unknown outputs of TCS 2—the goal being to understand the fundamental biology of the gene regulatory outputs of a TCS.

5 [0024] Tapparel (2006) replace the REC domain of *E. coli* CpxR (OmpR-PhoB family) with several others, but all chimeras are constitutively active. That is, they fail to achieve functional, switchable chimeras.

[0025] None of the above use DBD rewiring to identify unknown inputs for the TCS pair, nor can they then construct biosensors for those identified inputs using the new
10 constructions.

[0026] Thus, what is needed in the art are simple, reliable methods of identifying two component system inputs and engineering them to function as biosensors. It would also be beneficial to identify several useful crossover points for making functional hybrids, and to be able to transport two component systems between bacterial species, ensuring transcriptional
15 compatibility of the transported two component system with the new bacterium.

SUMMARY OF THE DISCLOSURE

[0027] The general purpose of this disclosure is to re-engineer naturally evolved bacterial two component signal transduction systems (a.k.a. two component systems or two component sensors, TCSs) to transport them between different bacterial species, discover the
20 inputs (chemicals, metabolites, hormones, environmental pollutants, industrial compounds, other bacteria, mechanical, or physical stimuli, etc.) that they sense, and enable their use as biosensors for scientific, industrial, medical, defense, environmental and other applications.

[0028] TCSs are a family of protein-based signaling pathways and the primary means by which bacteria sense and respond to the environment. Canonical TCSs consist of a
25 “sensor histidine kinase” or “SK” and a “response regulator” or “RR”. The canonical SK is embedded in the inner bacterial membrane. It contains a “sensor domain” that faces the extracellular environment (or periplasm in the case of gram negative bacteria), and a “transmembrane domain” that transmits information into an intracellular (cytoplasmic) histidine kinase “signaling domain.” In the presence of the cognate input stimulus

(hereinafter “input” or “input signal” or “ligand”), the sensor domain changes conformation. This conformational change is transmitted to the signaling domain by way of the transmembrane domain. The signaling domain is then activated, causing it to phosphorylate (chemically modify) itself on a specific histidine residue. The phosphoryl group is then transferred to a specific (cognate) non-membrane associated (cytoplasmic) RR, which then changes the expression of its target proteins. Some SKs can have cytoplasmic sensor domains. Others can be completely cytoplasmic (i.e. not membrane associated).

[0029] For canonical RRs, each RR has two domains, an N-terminal “REC” or “receiver” domain, which is phosphorylated by the SK, and a C-terminal DNA binding domain (“DBD”) that binds to the output promoter(s). The phosphorylated RR (known as RR~P) changes (activates or deactivates) expression of one or more specific gene(s). In this way, bacteria can sense and then respond to stimuli in the environment.

[0030] We have developed a combined computational/experimental method to discover what virtually any TCS senses, or at least the orthodox TCSs. We have also shown that we can use this method to transport TCSs between species of bacteria while retaining their sensing, signaling and transcriptional regulatory functions—which facilitates the identification of their inputs and their use as biosensors. This method also allows us to engineer novel biosensors from the newly characterized TCSs.

[0031] First, we use modern synthetic biology methods such as DNA synthesis or PCR and gene assembly methods to express computationally identified (by homology search and analysis of organization of TCS genes within bacterial genomes) TCSs in standard laboratory bacteria, such as the gram-negative organism *E. coli* and the gram-positive organism *B. subtilis*.

[0032] We have identified several specific amino acid residues in the two most widespread structural families, OmpR-PhoB (aka winged Helix-Turn-Helix; wHTH) and NarL-FixJ (aka Helix-Turn-Helix; HTH), of DNA-binding RRs wherein the DBD can be “swapped” for a well characterized DBD with a known output promoter (See **FIG. 23, FIG. 24, FIG. 25, FIG. 26, and FIG. 27**). We have engineered a suite of well-characterized DBDs and promoters that work in *E. coli* and *B. subtilis* in our laboratory, providing multiple options for rewiring the DBDs of any given TCS from the wHTH and HTH families.

[0033] Although the input of a novel TCS initially remains unknown, we can “rewire” that TCS to control a known output promoter. Then, by expressing any number of standard “reporter genes”, such as Green Fluorescent Protein (GFP), from the known output promoter, we can monitor the activity of any TCS in *E. coli* or *B. subtilis* or other bacteria by standard laboratory methods. If the TCS is non-functional in the new organism, we can rewire its DBD to make it functional. See FIG. 19.

[0034] There are a great variety of reporter genes that can be used herein, and GFP is only one convenient reporter. The amount or activity of the reporter protein produced is taken as a proxy for the TCS response to the target. Ideal reporter proteins are easy to detect and quantify (preferably noninvasively), highly sensitive and, ideally, not present in the native organism. They can be set up to detect either activation or deactivation. Several currently popular reporter proteins and their characteristics are listed in TABLE 1.

Table 1 | Common spectroscopically active reporter proteins and their detection

Reporter protein	Reporter genes	Origin	Substrate	Detection method	Comments	Refs
Bacterial luciferase	luxAB* or luxCDABE	Bioluminescent bacteria*	O ₂ , FMN, and long-chain aldehydes	Bioluminescence	Requires O ₂ ; aldehyde addition is required if only luxAB is used	94,95
Firefly luciferase	lucFF	Firefly (<i>Photinus pyralis</i>)	O ₂ , ATP and luciferin	Bioluminescence	Requires O ₂	96
Click beetle luciferase	lucGR	Click beetle (<i>Pyrophorus plagiopthalmus</i>)	O ₂ , ATP and phthalin	Bioluminescence	Requires O ₂	97
Renilla luciferase	Rluc	<i>Renilla reniformis</i>	Cosubstrate and Ca ²⁺	Bioluminescence	Requires O ₂	98
β -Galactosidase	lacZ	<i>Escherichia coli</i>	Galactopyranosides [†]	Chemiluminescence, colorimetry, electrochemistry and fluorescence	External substrate addition (may require cell permeabilization)	1
Fluorescent proteins	gfp, etc.	<i>Aequorea victoria</i> and additional marine invertebrates	N/A	Fluorescence	O ₂ is required for maturation; different colour varieties exist	99-101
Spheroidene monooxygenase	ortA	<i>Rhodovulum sulfatophilum</i>	Spheroidene	Colorimetry	None	102
Infrared fluorescent proteins	Various	Bacteriophytochrome family	N/A	Fluorescence	None	103
FMN-based fluorescent proteins	Various	Engineered from <i>Bacillus subtilis</i> and <i>Pseudomonas putida</i>	None	Fluorescence	Functional in both oxic and anoxic conditions; requires endogenous FMN	104

N/A, not applicable. *Most commonly used species include *Aliivibrio fischeri* (also known as *Vibrio fischeri*), *Vibrio harveyi* and *Photobacterium luminescens*. [†]For example, *O*-nitrophenyl- β -D-galactoside (ONPG), 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal), 4-methylumbelliferyl- β -D-galactopyranoside, 4-aminophenyl- β -D-galactopyranoside and D-luciferin- β -D-galactopyranoside.

[0035] Using the amount of reporter gene expression as a readout, and using standard high throughput screening methods, such as fluorimetry or flow cytometry, we can screen the novel TCS against virtually **any chemical or physical input**, and very easily measure those chemicals or input signals that it senses, using standard, high throughput laboratory assays.

[0033] Although the input of a novel TCS initially remains unknown, we can “rewire” that TCS to control a known output promoter. Then, by expressing any number of standard “reporter genes”, such as Green Fluorescent Protein (GFP), from the known output promoter, we can monitor the activity of any TCS in *E. coli* or *B. subtilis* or other bacteria by standard laboratory methods. If the TCS is non-functional in the new organism, we can rewire its DBD to make it functional. See **FIG. 19**.

[0034] There are a great variety of reporter genes that can be used herein, and GFP is only one convenient reporter. The amount or activity of the reporter protein produced is taken as a proxy for the TCS response to the target. Ideal reporter proteins are easy to detect and quantify (preferably noninvasively), highly sensitive and, ideally, not present in the native organism. They can be set up to detect either activation or deactivation. Several currently popular reporter proteins and their characteristics are listed in TABLE 1.

Table 1 Common spectroscopically active reporter proteins and their detection						
Reporter protein	Reporter genes	Origin	Substrate	Detection method	Comments	Refs
Bacterial luciferase	<i>luxAB*</i> or <i>luxCDABE</i>	Bioluminescent bacteria*	O ₂ , FMNH ₂ and long-chain aldehydes	Bioluminescence	Requires O ₂ ; aldehyde addition is required if only <i>luxAB</i> is used	94,95
Firefly luciferase	<i>lucFF</i>	Firefly (<i>photinus pyralis</i>)	O ₂ , ATP and luciferin	Bioluminescence	Requires O ₂	96
Click beetle luciferase	<i>lucGR</i>	Click beetle (<i>Pyrophorus plagiophthalmus</i>)	O ₂ , ATP and pholasin	Bioluminescence	Requires O ₂	97
<i>Renilla</i> luciferase	<i>Rluc</i>	<i>Renilla reniformis</i>	Coelenterazine and Ca ²⁺	Bioluminescence	Requires O ₂	98
β-Galactosidase	<i>lacZ</i>	<i>Escherichia coli</i>	Galactopyranosides†	Chemiluminescence, colorimetry, electrochemistry and fluorescence	External substrate addition (may require cell permeabilization)	1
Fluorescent proteins	<i>Gfp</i> , etc.	<i>Aequorea victoria</i> and additional marine invertebrates	N/A	Fluorescence	O ₂ is required for maturation; different colour varieties exist	99-101
Spheroidene monooxygenase	<i>crtA</i>	<i>Rhodovulum sulfidophilum</i>	Spheroidene	Colorimetry	None	102
Infrared fluorescent proteins	Various	Bacteriophytochrome family	N/A	Fluorescence	None	103
FMN-based fluorescent proteins	Various	Engineered from <i>Bacillus subtilis</i> and <i>Pseudomonas putida</i>	None	Fluorescence	Functional in both oxic and anoxic conditions; requires endogenous FMN	104

N/A, not applicable. *Most commonly used species include *Aliivibrio fischeri* (also known as *Vibrio fischeri*), *Vibrio harveyi* and *Photobacterium luminescens*. †For example, *O*-nitrophenyl-β-D-galactoside (ONPG), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), 4-methylumbelliferyl-β-D-galactopyranoside, 4-aminophenyl-β-D-galactopyranoside and D-luciferin-O-β-galactopyranoside.

[0035] Using the amount of reporter gene expression as a readout, and using standard high throughput screening methods, such as fluorimetry or flow cytometry, we can screen the novel TCS against virtually **any chemical or physical input**, and very easily measure those chemicals or input signals that it senses, using standard, high throughput laboratory assays.

[0036] As used herein, a “two component system”, “two component signal transduction system”, “two component sensor”, “two component sensor system”, or “TCS” is understood to be a two protein system including a sensor kinase and a response regulator, wherein the sensor kinase changes activity in response to a cognate input, resulting in a change in phosphorylation of the cognate response regulator by the sensor kinase, which then activates or de-activates transcription from the cognate output promoter(s) and thereby expression of relevant downstream proteins.

[0037] As used herein, “cognate” means the two (or more) parts function together. Non-cognate, by contrast, means a component that, under normal circumstances, would not function together with a given SK, although we subsequently rewire non-cognate DBD to be recognized by a SK by DBD swapping. Nonetheless, “non-cognate” indicates its origins from a non-cognate system.

[0038] As used herein, a sensor kinase is a protein understood to have a ligand binding domain (“LBD”) or similar input mechanism operably coupled to a kinase domain (KD), such that when the LBD binds its cognate ligand or sensor input, the kinase is activated.

[0039] As used herein, a “response regulator” or “RR” typically has a “receiver” or “REC” domain that is activated by the active kinase of the SK. Typically, the REC domain is operably coupled to a DNA binding domain or DBD, which thus can bind to and turn on relevant downstream protein expression.

[0040] As used herein, a “non-cognate DBD” means a DBD that comes from another protein, not the response regulator that the REC domain comes from. Typically, the DBD then binds to the DNA it is targeted to, which is itself coupled to a reporter gene that can easily be detected.

[0041] As used herein, an OmpR-PhoB family TCS is a TCS containing an RR which has 40% or greater amino acid sequence identity to OmpR (**FIG. 9**).

[0042] As used herein, a NarL-FixJ family of TCS is a TCS containing an RR, which has 40% or greater amino acid sequence identity to NarL (**FIG. 9**).

[0043] As used herein, a “crossover site” is a site where in the two domains (REC & DBD) of the RR can be successfully separated and a non-cognate DBD fused to the REC

domain. Exemplary crossover sites are shown in **FIG. 10, FIG. 23, FIG. 24, FIG. 25, FIG. 26, and FIG. 27.**

[0044] Preferably, the REC and DBD domains are separated at the crossover site, and recoupled in frame to a non-cognate DBD directly, e.g., **without** the use of added linker peptides. Linker peptides can be used, but we show herein that they are not needed and they may even be detrimental. Also, preferred, the non-cognate DBD is cut at the same crossover point, or reasonably close by, e.g., within 1-10 aa, preferably 1-3 aa. Some small amount of leeway may be accommodated, providing the 3D structure of the protein is largely retained. Of course, it is understood that the gene fragments must be fused in frame for operability.

[0045] As used herein, a “rewired” RR is a “chimeric” or “hybrid” RR to RRs with DBDs swapped for those of another family member.

[0046] As used herein, an “input” or input signal” refers to the incoming chemical or environmental condition that activates the SK of a TCS. An “output” on the other hand, refers to those genes, or promoters thereof, being regulated by the RR.

[0047] As used herein, “heterologous” means a component from another species.

[0048] As used herein, the expressions “bacterium”, “bacteria”, “microorganism”, “microbe”, “strain”, “species” and the like may be used interchangeably and all such designations include their progeny. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Mutant progeny that have the same function or biological activity as screened for in the originally transformed cell are included. Where distinct designations are intended, it will be clear from the context.

[0049] As used herein, reference to a “cell” is generally understood to include a culture of such cells, as the work described herein is done in cultures having 10^9 - 10^{15} cells.

[0050] As used herein, “growing” cells used in its art accepted manner, referring to exponential growth of a culture of cells, not the few cells that may not have completed their cell cycle at stationary phase or have not yet died in the death phase or after harvesting.

[0051] As used in the claims, “homolog” means an enzyme with at least 40% identity to one of the listed sequences and also having the same general catalytic activity, although of

course K_m , K_{cat} and the like can vary. While higher identity (60%, 70%, 80%) and the like may be preferred, it is typical for bacterial sequences to diverge significantly (40, 50, 60%), yet still be identifiable as homologs, while mammalian species tend to diverge less (80-90%).

[0052] Reference to proteins herein can be understood to include reference to the gene encoding such protein. Thus, a claimed “permease” protein can include the related gene encoding that permease. However, it is preferred herein to refer to the protein by standard name per *ecoliwiki* or HUGO since both enzymatic and gene names have varied widely, especially in the prokaryotic arts.

[0053] Once an exemplary protein is obtained, many additional examples of proteins with similar activity can be identified by BLAST search. Further, every protein record is linked to a gene record, making it easy to design overexpression vectors. Many of the needed enzymes are already available in vectors, and can often be obtained from cell depositories or from the researchers who cloned them. But, if necessary, new clones can be prepared based on available sequence information using RT-PCR techniques. Thus, it should be easily possible to obtain all of the needed enzymes/genes for overexpression.

[0054] Another way of finding suitable enzymes/genes for use in the invention is to consider other enzymes with the same EC number, since these numbers are assigned based on the reactions performed by a given enzyme. An enzyme that thus be obtained, e.g., from AddGene or from the author of the work describing that enzyme, and tested for functionality as described herein. In addition, many sites provide lists of proteins that all catalyze the same reaction.

[0055] Understanding the inherent degeneracy of the genetic code allows one of ordinary skill in the art to design multiple nucleotides that encode the same amino acid sequence. NCBI™ provides codon usage databases for optimizing DNA sequences for protein expression in various species. Using such databases, a gene or cDNA may be “optimized” for expression in *E. coli*, yeast, algal or other species using the codon bias for the species in which the gene will be expressed.

[0056] Initial cloning experiments have proceeded in *E. coli* for convenience since most of the required genes are already available in plasmids suitable for bacterial expression, but the addition of genes to bacteria is of nearly universal applicability. Indeed, since

recombinant methods were invented in the 70's and are now so commonplace, even school children perform genetic engineering experiments using bacteria. Such species include e.g., *Bacillus*, *Streptomyces*, *Azotobacter*, *Trichoderma*, *Rhizobium*, *Pseudomonas*, *Micrococcus*, *Nitrobacter*, *Proteus*, *Lactobacillus*, *Pediococcus*, *Lactococcus*, *Salmonella*, *Streptococcus*,
5 *Paracoccus*, *Methanosarcina*, and *Methylococcus*, or any of the completely sequenced bacterial species. Indeed, hundreds of bacterial genomes have been completely sequenced, and this information greatly simplifies both the generation of vectors encoding the needed genes, as well as the planning of a recombinant engineering protocol. Such species are listed along with links at http://en.wikipedia.org/wiki/List_of_sequenced_bacterial_genomes.

10 **[0057]** Furthermore, a number of databases include vector information and/or a repository of vectors and can be used to choose vectors suitable for the chosen host species. *See e.g.*, AddGene.org which provides both a repository and a searchable database allowing vectors to be easily located and obtained from colleagues. *See also* Plasmid Information Database (PlasmID) and DNASU having over 191,000 plasmids. A collection of cloning
15 vectors of *E. coli* is also kept at the National Institute of Genetics as a resource for the biological research community. Furthermore, vectors (including particular ORFs therein) are usually available from colleagues.

[0058] The proteins can be added to the genome or via one or more expression vectors (preferably inducible), as desired. Preferably, multiple proteins are expressed in one
20 vector or multiple enzymes can be combined into one operon by adding the needed signals between coding regions. Further improvements can be had by overexpressing one or more, or even all of the proteins, e.g., by adding extra copies to the cell via plasmid or other vector. Initial experiments may employ expression plasmids hosting one or more ORFs for convenience, but it may be preferred to insert operons or individual genes into the genome
25 for long-term stability. It may be especially preferred to integrate a reporter gene construct into the genome, as that cell could then be used as the basis for many different biosensors, merely by switching out the TCS components.

[0059] In calculating “% identity” the unaligned terminal portions of the query sequence are not included in the calculation. The identity is calculated over the entire length
30 of the reference sequence, thus short local alignments with a query sequence are not relevant (e.g., % identity = number of aligned residues in the query sequence/length of reference

sequence). Alignments are performed using BLAST homology alignment as described by Tatusova TA & Madden TL (1999) FEMS Microbiol. Lett. 174:247-250, and available through the NCBI website. The default parameters were used, except the filters were turned OFF.

5 [0060] “Operably associated,” “operably linked”, or “operably coupled” as used herein, refer to functionally coupled nucleic acid or amino acid sequences.

[0061] “Recombinant” is relating to, derived from, or containing genetically engineered material. In other words, the genetic material of an organism was intentionally manipulated by the hand-of-man in some way.

10 [0062] “Overexpression” or “overexpressed” is defined herein to be at least 150% of protein activity as compared with an appropriate control species, or any detectable expression in a species that lacks the activity altogether. Preferably, the activity is increased 100–500% or even ten fold. Overexpression can be achieved by mutating the protein to produce a more active form or a form that is resistant to inhibition, by removing inhibitors, or adding
15 activators, and the like. Overexpression can also be achieved by removing repressors, adding multiple copies of the gene to the cell, or up-regulating the endogenous gene, and the like. All overexpressed genes or proteins are signified herein by “+”.

[0063] In certain species, it is possible to genetically engineer an endogenous protein to be overexpressed by changing the regulatory sequences or removing repressors. However,
20 overexpressing the gene by inclusion on selectable plasmids or other vectors that exist in hundreds of copies in the cell may be preferred due to its simplicity and ease of exerting external controls, although permanent modifications to the genome may be preferred in the long term for stability reasons.

[0064] The term “endogenous” means that a gene originated from the species in
25 question, although that gene may be naturally or intentionally mutated, or placed under the control of a promoter that results in overexpression or controlled expression of said gene. Thus, genes from *Clostridia* would not be endogenous to *Escherichia*, but a plasmid expressing a gene from *E. coli* or would be considered to be endogenous to any *E. coli* strain, even though it may now be overexpressed. In contrast, a “wild type” gene or protein means
30 the gene coding regions and have not been substantively changed, nor the activity, and the

amino acid sequence corresponds to one found in nature. A “wild type endogenous” gene would thus be the same gene found in that species, without any substantive mutations to the coding regions.

5 [0065] “Expression vectors” are used in accordance with the art-accepted definition of a plasmid, virus or other propagatable sequence designed for protein expression in cells. There are thousands of such vectors commercially available, and typically each has an origin of replication (ori); a multiple cloning site; a selectable marker; ribosome binding sites; a promoter and often enhancers; and the needed termination sequences. Most expression vectors are inducible, although constitutive expression vectors also exist.

10 [0066] As used herein, “inducible” means that gene expression can be controlled by the hand-of-man, by adding e.g., a ligand to induce expression from an inducible promoter. Exemplary inducible promoters include the lac operon, inducible by IPTG, the strong LAC4 promoter inducible with lactate, and the like. Low level of constitutive protein synthesis may occur even in expression vectors with tightly controlled promoters.

15 [0067] As used herein, an “integrated sequence” means the sequence has been integrated into the host genome, as opposed to being maintained on an expression vector. It will still be expressible, and preferably is inducible as well.

[0068] The use of the word “a” or “an” when used in conjunction with the term “comprising” in the claims or the specification means one or more than one, unless the
20 context dictates otherwise.

[0069] The term “about” means the stated value plus or minus the margin of error of measurement or plus or minus 10% if no method of measurement is indicated.

[0070] The use of the term “or” in the claims is used to mean “and/or” unless explicitly indicated to refer to alternatives only or if the alternatives are mutually exclusive.

25 [0071] The terms “comprise”, “have”, “include” and “contain” (and their variants) are open-ended linking verbs and allow the addition of other elements when used in a claim.

[0072] The phrase “consisting of” is closed, and excludes all additional elements.

[0073] The phrase “consisting essentially of” excludes additional material elements, but allows the inclusions of non-material elements that do not substantially change the nature

of the invention, such as instructions for use, buffers, background mutations that do not effect the invention, and the like.

[0074] The following abbreviations are used herein:

ABBREVIATION	TERM
aa	Amino acid
ACCN	Accession number
aTc	Anhydrotetracycline
BAD_0568	RR from <i>Bifidobacterium adolescentis</i> (see Kegg entry for BAD_0568).
BAD_0569	SK from <i>B. adolescentis</i> . Phosphorylates BAD-0568.
BceR	RR from <i>B. subtilis</i> , ACCN: WP_004399109
CcaR	RR from <i>Synechocystis</i> PCC6803, ACCN: WP_010874216.
CcaS	Green light activated, red light repressed SK from <i>Synechocystis</i> PCC 6803 engineered to work in <i>E. coli</i> . Phosphorylates CcaR.
CopR	RR from <i>Synechocystis</i> PCC 6803, ACCN: WP_010873936 SEQ ID NO: 3
Cph8	Engineered red/far red light switchable SK in <i>E. coli</i> . Phosphorylates OmpR.
DBD	DNA binding domain
FusR	RR from <i>E. coli</i> , ACCN: AAG54714 SEQ ID NO: 29, the cognate SK is fusK
GFP	Green fluorescent protein
KD	Kinase domain
LBD	Ligand binding domain
LiaR	RR from <i>B. subtilis</i> , ACCN: WP_003243201 SEQ ID NO: 28
ManR	RR from <i>Synechocystis</i> PCC6803, ACCN: WP_010872074
mCherry	mCherry fluorescent protein
NarL	RR from <i>Escherichia coli</i> . ACCN: NP_415739
NarX	Nitrate sensing SK from <i>Escherichia coli</i> . Phosphorylates NarL.
OmpR	RR from <i>E. coli</i> . ACCN: NP_417864
PydJ	<i>B. subtilis</i> Promoter activated by the phosphorylated form of the response regulator Ydfl.
PdcuS	Phosphorylated NarL binds to the repressible PdcuS promoter from <i>E. coli</i> , regulating the expression of sfGFP. Low nitrate conditions result in high sfGFP expression, high nitrate conditions repress transcription. The sensitivity of the sensor is in the sub-mM to mM range for nitrate.
PcpcG2	Inverted Green light inducible pCpcG2 promoter, from pJT122 plasmid constructed by Tabor et al. (2010). It is positively regulated by the two component system CcaS/R, which exhibits a maximum response in 535 nm and is inactivated in 650 nm light. Light intensities must be carefully regulated to achieve successful gene expression. The sequence was inverted to ease DNA synthesis.
PsdR	RR from <i>B. subtilis</i> , ACCN: WP_003244535
REC	Receiver domain
RR	Response regulator
SK	Sensor kinase or sensor histidine kinase
TCS	Two component sensor system including a SK and an RR
UhpA	RR from <i>E. coli</i> , ACCN: NP_418125
UhpB	<i>E. coli</i> SK that interacts with UhpC, a periplasmic protein that binds Glucose-6-phosphate. UhpB phosphorylates UhpA.
UhpC	<i>E. coli</i> periplasmic protein that binds Glucose-6-phosphate and then stimulates the SK UhpB.
Ydfl	RR from <i>B. subtilis</i> , ACCN: WP_003244318

YxdJ	RR from <i>B. subtilis</i> , ACCN: WP_003243527
------	---

BRIEF DESCRIPTION OF FIGURES

[0075] **FIG. 1.** Overview of new technology and approach.

[0076] **FIG. 2.** Rewiring the REC domain of the *B. adolescentis* response regulator
5 BAD_0568 to the known output DBD of *Synechocystis* PCC6803 response regulator CcaR in
E. coli.

[0077] **FIG. 3.** BAD_0568-CcaR activates GFP expression from the standard PcpG2
promoter in *E. coli*. BAD_0568 is an RR from *Bifidobacterium adolescentis* (see Kegg entry
for BAD_0568). CcaR is an RR from *Synechocystis* PCC 6803 that we have transported to *E.*
10 *coli*. It induces transcription from PcpG2 in *Synechocystis* PCC6803 and *E. coli*.
Anhydrotetracycline is used to induce BAD_0568/CcaR expression. Superfolder GFP is being
expressed from the PcpG2-172 promoter (a variant of PcpG2) in *E. coli*. The cognate
BAD_0569 SK is absent, but BAD_0568/CcaR is being phosphorylated from another source,
such as acetyl phosphate or a non-cognate SK. A plasmid vector with the ColE1 origin of
15 replication and chloramphenicol resistance maker is used to carry the SK, RR constructs, and
the reporter gene construct was described in Tabor (2010) and Tabor (2011).

[0078] **FIG. 4** Validation of proper BAD_0568-CcaR chimera function: activation of
PcpG2-172 depends on the conserved REC domain phosphorylation site.

[0079] **FIG. 5** Validation of chimeric pathway activity. Expression of the sensor
20 kinase BAD_0569 activates transcription from PcpG2-172 in the presence of
BAD_0568/CcaR chimera.

[0080] **FIG. 6** Validation of pathway activity. Mutation of BAD_0569 catalytic
histidine (H298) to non-catalytic asparagine (N) residue (i.e. H298N) as shown here, greatly
reduced if not eliminated reporter gene expression (see **FIG. 7**).

25 [0081] **FIG. 7** High throughput screen of rewired BAD_0569/BAD_0568 (a.k.a.
BAD_0569/8)-CcaR pathway to panel of mammalian gut metabolites in *E. coli* with negative
control H298N showing no response and positive control of TAZ/OmpR-CcaR system

(wherein the SK TAZ senses only aspartate and phosphorylates OmpR-CcaR which activates transcription from PcpG2-172) responding only to aspartate.

[0082] **FIG. 8** Methionine Dose Response of rewired BAD_0569/BAD_0568-CcaR TCS with sfGFP output from PcpG2-172 in *E. coli*.

5 [0083] **FIG. 9.** Amino acid sequences for the native RRs and hybrid RRs tested herein.

[0084] **FIG. 10.** Identification of crossover points permitting the rewiring of OmpR-CcaR to make chimeric or hybrid RRs. The 705 nm light activated, 650 nm light de-activated sensor histidine kinase (SK) Cph8 phosphorylates our novel OmpR-CcaR hybrids. The
10 phosphorylated OmpR-CcaR hybrids activate transcription from the CcaR activated PcpG2-172 promoter in *E. coli*. PcpG2-172 activity is measured using sfGFP reporter gene and flow cytometry.

[0085] **FIG. 11.** OmpR-ManR rewire. Dark and 705 nm light both stimulate the SK Cph8—the latter shown in **FIG. 10**. Thus, dark results in phosphorylation of OmpR-ManR,
15 causing repression of the ManR output promoter in *E. coli*.

[0086] **FIG. 12.** OmpR-PsdR rewire. Data collected in *E. coli* expressing Cph8. PsdR activates its output promoter, thus dark results in increased transcription and 650 nm results in decreased transcription in this chimera.

[0087] **FIG. 13.** OmpR-YxdJ. Data collected in *E. coli* expressing Cph8. YxdJ
20 activates its output promoter.

[0088] **FIG. 14.** CcaR-CopR rewire. The SK CcaS is activated by green light (520 nm) and repressed by red (650 nm) in the presence of the chromophore phycocyanobilin (PCB), but not in its absence. These data were collected in *E. coli* expressing CcaS with and without PCB. CcaS phosphorylates CcaR-CopR, causing repression of its output promoter.
25 This effect is absent without chromophore.

[0089] **FIG. 15.** CcaR-ManR rewire. Experiments are as described in **FIG. 14**, unless stated otherwise.

[0090] **FIG. 16.** CcaR-BceR rewire.

[0091] **FIG. 17.** CcaR-PsdR rewire.

[0092] **FIG. 18.** CcaR-YxdJ rewire.

[0093] **FIG. 19.** NarL-YdfI rewire. The wild-type TCS NarX-NarL with PdcuS promoter output was transported from *E. coli* to *B. subtilis*. The bacteria were treated with and without the NarX inducer nitrate. No response is seen (Left). When NarL DBD is replaced with YdfI and the PydfJ output promoter, which is activated by the YdfI DBD, the nitrate response is recovered. This data shows that DBD rewiring allows transport of TCSs between organisms with otherwise incompatible transcriptional regulation systems.

[0094] **FIG. 20.** UhpA-YdfI chimera in *E. coli*. The SK UhpB is activated by Glucose-6-phosphate. UhpB phosphorylates UhpA-YdfI, which activates the PydfJ output promoter. The dose response curve to the inducer is shown.

[0095] **FIG. 21.** UhpA-LiaR rewire. The experiment is as described in **FIG. 20** unless stated otherwise.

[0096] **FIG. 22.** FusR-YdfI rewire. In the presence of the SK (kinase) FusK, FusR-YdfI (RR) is activated, activating the PydfJ output promoter. In the absence of FusK, the RR is less activated. In the absence of RR, the output promoter is not activated. Experiments done in *E. coli*.

[0097] **FIG. 23** Alignment of amino acid sequences, numbered according to the numbering of the OmpR RR, and showing all the various sites tested for successful cleavage of the REC domain from the DBD.

[0098] **FIG. 24.** The various crossover sites tested in the OmpR family, showing the best results are seen with crossover sites in the 120-140 range.

[0099] **FIG. 25.** OmpR 3D structure with the optimal crossover point from the OmpR-CcaR crossover survey highlighted.

[00100] **FIG. 26.** Alignment of NarL-FixJ family members with crossover points tested and best identified crossover points for specific cross-overs indicated.

[00101] **FIG. 27.** NarL 3D structure with the optimal crossover point from the NarL-YdfI crossover study highlighted.

[00102] **FIG. 28.** A partial listing of embodiments of the invention, any one of which can be combined with any one or more embodiments, or portions thereof.

DETAILED DESCRIPTION

[00103] The steps of the method include one or more of the following, although control experiments may be varied and certain steps can be omitted, depending on the state of research:

5 [00104] 1) Identify a TCS (SK and RR) of interest from published literature or computational search of DNA or preferably amino acid sequence. The SK and its cognate RR are typically within 200-1000 base pairs of one another on the genome, encoded on the same or opposite strands, making the pairing of the cognate SK and RR fairly simple to determine based on homology search and domain identification.

10 [00105] 2) Swap the naturally occurring DBD of the RR with that of a well-characterized non-cognate DBD of an RR with a known output promoter, such as CcaR (or a list of numerous others that we have developed herein or can be identified from the literature).

[00106] 3) Express the SK and the modified RR in a model laboratory bacterium such
15 as *E. coli* or *Bacillus subtilis*, and the like. Constitutive or inducible promoters may be used, but inducible promoters enable rapid identification of SK and RR expression levels resulting in proper input-dependent TCS response. Improper SK or RR expression levels can reduce or eliminate input dependent TCS response.

[00107] The functionality of the chimeric RR alone can be validated by inducing its
20 expression over a wide range and measuring the response of the output promoter that the chimeric RR is intended to regulate with a GFP reporter gene. In a typical case, the chimeric RR may activate the desired output promoter. Thus induction of the RR over a wide range will result in activation of the new promoter. Activation without the input (or even SK) present can result from alternative sources of phosphorylation of the RR in the cell (e.g. from
25 acetyl-phosphate or non-cognate SKs) or low-affinity binding of the promoter, which nonetheless becomes significant at high levels of RR expression. By mutating the conserved RR aspartate to a non-phosphorylatable residue, the chimera can be further validated (this mutant should not activate the desired output promoter as strongly, or at all).

[00108] 4) Co-express a reporter gene, such as a fluorescent protein, a chromogenic
30 enzyme (beta-galactosidase) or an mRNA that can be quantified under the known output

promoter of the replacement DBD. The three components (SK, modified RR and reporter gene) can be provided on the same expression vector, or multiple expression vectors. The reporter gene can also be integrated into the genome, such that the same cell line can be used for a number of different SK/RR pairs. The SK and RR can also be integrated, but do not have to be.

[00109] 6) Screen the cells expressing the SK and modified RR against any input signals that may activate the SK by growing the bacteria at different concentrations of the candidate input signals.

[00110] 7) Identify those input signals that trigger a change in the reporter gene expression.

[00111] Preferably, these experiments are followed up by validation experiments, to confirm proper signaling. Alternatively, these experiments can be performed simultaneously, but fewer sample runs are needed with sequential experiments since only those positive inputs are then tested by these control experiments.

[00112] 8) Validate those input signals by mutating the conserved histidine on the SK or phosphorylated aspartate on the RR and repeating the treatment with the input. These mutations should reduce if not eliminate the signal.

[00113] 9) Demonstrate that the input does not activate the RR or output promoter or reporter gene non-specifically by expressing a second reporter gene (e.g. mCherry) from a constitutively active promoter (not regulated) in the same bacterium. A true input will result in a change in the ratio of the pathway-specific reporter to the constitutive reporter, but not when the histidine or aspartate is mutated.

[00114] If a TCS cannot be expressed (e.g. does not fold) in a standard laboratory bacterium, this can be overcome by a number of standard means for increasing solubility (lower temperature, fusion to maltose binding protein, chaperone overexpression). Other model organisms, including other bacteria and yeasts, could also be used to expand work around possible failures.

[00115] If a TCS is not completely self contained (e.g. requires an additional gene, protein, or cofactor in the native organism that is not present in the model organism), this could be overcome by expressing the additional gene (if known) or libraries of genes from

the genome of the native organism alongside the DBD-replaced TCS in the laboratory organism, or using other model organisms. Alternatively, the additional gene may be simple to identify because it resides adjacent to the TCS on the genome and is homologous to known additional genes that are required for signaling (e.g. UhpC resides next to UhpB and UhpA on the *E. coli* genome and UhpC is a transporter like protein shown to bind Glucose-6-phosphate for UhpB mediated signaling to UhpA and UhpC homologs are found next to other TCSs as well).

[00116] We have demonstrated proof of concept for the method using a novel OmpR-PhoB family TCS, BAD_0569-BAD_0568 (aka BAD_0569/8), which is taken from the species *Bifidobacterium adolescentis*. Homologs of BAD_0569/8 have also been discovered to be enriched in the colonic bacteria of obese humans, and thus this particular TCS is of general interest in obesity research.

[00117] The general experimental outline for this proof of concept experiment is shown in **FIG. 2**. Briefly, the wild type BAD_0569 SK of *Bifidobacterium adolescentis* was expressed in *E. coli*. A modified cognate response regulator BAD_0568 was simultaneously expressed in *E. coli*, wherein the native DBD from the *Bifidobacterium adolescentis* RR was replaced with that of the DBD from the well-characterized OmpR-PhoB family RR CcaR. When phosphorylated, native CcaR activates transcription from PcpG2, or its derivatives. A superfolder GFP reporter gene (a.k.a. GFP) was expressed from PcpG2 in the same *E. coli*. Anydrotetracycline (aTc) induces expression of the BAD_0568-CcaR hybrid, which activates PcpG2 transcription and thus GFP expression.

[00118] Generally speaking herein, SK was expressed from a ColE1 plasmid, the RR was expressed from a p15A plasmid, and the output promoter and reporter were expressed from a p15A or a pSC101 plasmid. However, this is a matter of convenience only, and two or all three components could be co-expressed from a single plasmid, and/or one or more components can be integrated.

[00119] **FIG. 3** shows the functionality of the system. Three different BAD_0568-CcaR chimeras were made and induced with aTc as in **FIG. 2**. GFP expression increases with induction of the BAD_0568-CcaR chimeras, likely due to spontaneous phosphorylation of the chimera by acetyl phosphate or non-cognate SKs in *E. coli* (BAD_0569 is absent). This experiment shows that the chimera can activate the non-native PcpG2 output promoter.

Functionality of the chimera was also validated in several ways. For example, in **FIG. 4**, we show that response completely depends on the conserved REC domain phosphorylation site. **FIG. 5** shows that co-expression of the BAD_0569 SK (induced by IPTG in this case), which phosphorylates the BAD_0568 REC domain to some extent even in the absence of its input increases PcpG2 activity. In **FIG. 6**, we show that mutation of catalytic histidine to non-catalytic asparagine (N) residue eliminates the activating effect of BAD_0569, validating that the effect is due to phosphotransfer (see **FIG. 7** for exemplary data).

[00120] We then tested *E. coli* expressing the hybrid BAD_0569/8 TCS (i.e. BAD_0569 and BAD_0568-CcaR with sfGFP expressed from PcpG2) in the presence of over 40 chemicals previously found to be enriched in the large intestines of germ-free mice made to carry gut bacteria of obese humans, but not found to be enriched in those mice carrying the gut bacteria of their lean twins. From this list, we have found several possible candidate molecules that are sensed by BAD_0569/8. **FIG. 7**.

[00121] Although we engineered our system in order to identify the inputs for BAD-0569, once the input was identified, our bacteria could then be used as a biosensor for that input. The molecule yielding the largest response was methionine, an amino acid, which may be a biomarker of an unbalanced gut microbiome that can potentially cause obesity (based on published mouse studies and human observations), or at least be a marker for same. Thus, we have used DBD rewiring to engineer a bacterial sensor of pre-obesity or obesity conditions in the human gut.

[00122] Of note, *B. adolescentis* is a gram-positive bacterium (lacks an outer membrane), while *E. coli* is a gram negative bacterium (has an outer and inner membrane, and the sensor domain of the SK is in the middle periplasmic region). Thus, we demonstrate herein that the TCS can be moved between these two types of bacteria via DBD rewiring. In particular, the output promoter of wild-type BAD_0568 is not known, which precludes movement of a functional pathway into a new organism. Thus, DBD rewiring overcame this limitation, enabling facile movement into *E. coli*.

[00123] We next showed that the proof of concept could be applied more generally to other TCSs—or at least those of the OmpR-PhoB and NarL-FixJ families. **FIG. 10** shows some 15 rewired sequences that we have tested using the methods described herein.

[00124] We have identified multiple amino acid positions wherein the DBDs of RRs with OmpR-PhoB (a.k.a. winged Helix-turn-Helix; wHTH) and NarL-FixJ (a.k.a. Helix-turn-Helix; HTH) family DBDs can be rewired. Specifically, we have used a set of 7 wHTH RRs from *E. coli*, *Synechocystis* PCC 6803, and *B. subtilis*, and 5 HTH RRs from *E. coli* and *B. subtilis*. We have created 15 wHTH rewires (i.e. the DBD of a wHTH RR is replaced with that of a different wHTH RR) from this set and demonstrated their functionality using input (visible light or chemical ligand)-induced transcriptional activation and a superfolder GFP (sfGFP) reporter gene and flow cytometry in *E. coli* (see **FIG. 11-22**).

[00125] We have created 4 successful HTH family rewires and similarly demonstrated their functionality using chemical ligand induced transcriptional activation and sfGFP and flow cytometry in *E. coli* and *B. subtilis*. In particular, we have moved the nitrate activated TCS NarX-NarL from *E. coli* into *B. subtilis* by replacing the *E. coli*-derived NarL DBD with that of the *B. subtilis* derived YdfI DBD.

[00126] **FIG. 22** shows a compilation of results obtained by testing a wide range of crossover sites in OmpR family RRs. As can be seen, successful separation occurs in the range of amino acid 110-151. **FIG. 23** provides additional detail, and it can be seen that the best separation sites are about 120-140, especially, 122, 137, 138 and 139. All amino acid sequences are numbered according to their alignment with OmpR.

[00127] We have also shown that certain TCSs cannot be transported from one bacterium into different species due to incompatibilities of the native DBD with the transcriptional apparatus in the target or host bacterium. However, we have shown that DBD rewiring overcomes this challenge, enabling TCSs to be transported between otherwise incompatible species. In particular, replacement of a native DBD from the native organism (wherein the TCS evolved) with that of a DBD from the target organism (where the DBD is to be moved) enables successful transport. Thus, use of a host DBD of known output overcomes expression incompatibilities.

[00128] We have also used rewiring to move the nitrate activated TCS NarX-NarL from the gram-negative *E. coli* to the gram-positive *B. subtilis*. In particular, we showed that nitrate does not alter gene expression from the *E. coli* output promoter PdcuS in *B. subtilis*, likely due to transcriptional incompatibilities. However, we then replaced the DBD of NarL with that of the *B. subtilis* RR YdfI, and expressed sfGFP from the YdfI-activated PydfI

output promoter. In this system, nitrate strongly activates sfGFP expression, demonstrating that we have used DBD rewiring to transport this TCS from *E. coli* to *B. subtilis*.

[00129] The fact that we can now move modified TCSs between gram positive and gram-negative bacteria dramatically increases the utility and breadth of the method, allowing
5 us to characterize most computationally identifiable TCSs in the laboratory.

[00130] Because we have demonstrated that DBD rewiring is general, and that TCSs can be moved between the major classes of bacteria, this method has potential to be used very widely to determine the inputs sensed by virtually any computationally (or otherwise) identifiable TCS. Our high-throughput approach to developing novel bacterial sensors has
10 wide utility across biotechnology and medicine.

[00131] Some major uses are:

[00132] 1. To better understand the biology of the human microbiome by revealing what the bacteria in and upon our bodies are sensing.

[00133] 2. To engineer novel diagnostic agents for a wide range of chemical and
15 physical signals within or upon the human body.

[00134] 3. To engineer novel biosensors for a wide range of chemical and physical signals in the environment (e.g. pollutants, toxins, chemical weapons, pathogenic bacteria, etc.)

[00135] 4. To engineer "smart probiotics" that secrete therapeutic molecules in the
20 body only in the presence of biomarkers that indicate a diseased state.

[00136] 5. To engineer metabolic sensors for "smart" fermentation strains that can detect diverse ranges of feedstocks, metabolic intermediates, fermenter conditions, and so on, and tune gene expression for optimal product yields in response.

[00137] 6. To understand existing complex bacterial regulatory machinery—by
25 replacing native TCSs with unknown outputs with our known TCSs using our technique, we could identify and uncover existing modes of regulation with relevance to elucidating antibiotic resistance or novel therapeutic strategies.

[00138] Most evolved bacteria in which TCSs naturally occur cannot be cultured nor genetically manipulated in the laboratory. This makes studying their function highly difficult.

Additionally, it is computationally difficult to identify the output promoters of most TCSs, meaning one must perform screens to identify their input signals, yet screening methods are very difficult where there are no culturing methods available. Our method overcomes these problems for virtually all TCSs, rendering identification of their inputs much easier. Our
5 method also inherently generates an engineered biosensor for the molecule sensed by the new TCS.

[00139] The above experiments are repeated in *Bacillus subtilis*. The same genes can be used, especially since *Bacillus* has no significant codon bias. A protease-deficient strain like WB800N is preferably used for greater stability of heterologous protein. The *E. coli* - *B.*
10 *subtilis* shuttle vector pMTLBS72 exhibiting full structural stability can be used to move the genes easily to a more suitable vector for *Bacillus*. Alternatively, two vectors pHT01 and pHT43 allow high-level expression of recombinant proteins within the cytoplasm. As yet another alternative, plasmids using the theta-mode of replication such as those derived from the natural plasmids pAM β 1 and pBS72 can be used. Several other suitable expression
15 systems are available.

[00140] Our future plans include use of the method to characterize other TCSs from the human microbiome. We also plan to use the method to characterize other TCSs from marine and other environmental bacteria. We hope to characterize hundreds of novel TCSs overall in the coming three years and generate a large number of biological sensors for a
20 variety of chemicals and other inputs.

[00141] Each of the following citations is incorporated by reference herein in its entirety for all purposes.

[00142] US20030049799 Engineered stimulus-responsive switches.

[00143] US9062320 Biological systems input-output response system and plant
25 sentinels

[00144] Allen, et al. Genetic Evidence that the α 5 Helix of the Receiver Domain of PhoB Is Involved in Interdomain Interactions, J. Bacteriology 183(72001): 2204–2211 (2001).

- [00145] Castillo-Hair, S.M., et al., How to train your microbe: methods for dynamically characterizing gene networks, *Current opinion in microbiology* 24, 113-123 (2015).
- [00146] da Silva D.P. et al., Studies on synthetic LuxR solo hybrids, *Front. Cell. Infect. Microbiol.*, Vl. 5, Art. 52 (2015).
- [00147] Howell, et al, Genes controlled by the essential YycG/YycF two component system of *Bacillus subtilis* revealed through a novel hybrid regulator approach, *Molecular Microbiology*, 49(6) 1639–1655 (2003).
- [00148] Kohanski, M.A., & Collins J.J., Rewiring Bacteria, Two Components at a Time, *Cell* 133: 947-948 (2008).
- [00149] Levskaya, A., et al., Synthetic biology: engineering *Escherichia coli* to see light, *Nature* 438 (7067), 441-442 (2005).
- [00150] Ryan, R. et al., Toward rationally redesigning bacterial two -component signaling systems using coevolutionary information, *PNAS* 111(5): E563–E571 (2014).
- [00151] Schmidl, S.R., et al., Refactoring and optimization of light-switchable *Escherichia coli* two-component systems, *ACS synthetic biology* 3 (11), 820-831 (2014).
- [00152] Skerker, J.M., et al., Rewiring the Specificity of Two-Component Signal Transduction Systems, *Cell*. 133(6): 1043–1054 (2008).
- [00153] Tabor, J.J., et al., A synthetic genetic edge detection program, *Cell* 137 (7), 1272-1281 (2009).
- [00154] Tabor, J.J., et al., Performance characteristics for sensors and circuits used to program *E. coli*, *Systems Biology and Biotechnology of Escherichia coli*, 401-4397 (2009).
- [00155] Tabor JJ, et al., Multichromatic Control of Gene Expression in *Escherichia coli*. *J Mol Biol* 405:315-324 (2010).
- [00156] Tapparel et al. The DNA-binding domain of the *Escherichia coli* CpxR two-component response regulator is constitutively active and cannot be fully attenuated by fused adjacent heterologous regulatory domains, *Microbiology* 152: 431–441 (2006).

[00157] Walthers et al. Interdomain Linkers of Homologous Response Regulators Determine Their Mechanism of Action, J. Bacteriology Jan. 185(1): 317–324 (2003).

[00158] Wang, B., et al., Rewiring cell signalling through chimaeric regulatory protein engineering, Biochem. Soc. Trans. (2013) 41, 1195–1200 (2013).

5 [00159] While a number of exemplary aspects and embodiments have been discussed above, those of skill in the art will recognize certain modifications, permutations, additions and sub-combinations thereof. It is therefore intended that the following appended claims and claims hereafter introduced are interpreted to include all such modifications, permutations, additions and sub-combinations as are within their true spirit and scope. It is also intended
10 that any detail anywhere in the claims or anywhere in the specification can be combined with any other detail herein, even if not yet expressly so combined, as the specification would be of inordinate length if we were to recite all possible combinations of DBDs, reporter genes, promoters, host species, and the like.

CLAIMS

- 1) A genetically engineered bacteria, comprising:
 - a) a modified two component sensor system (TCS), said TCS comprising:
 - i) a wild type sensor kinase (SK) comprising a ligand binding domain operably coupled to a kinase domain, and
 - ii) a modified response regulator (RR) that is cognate to said SK, said RR comprising a cognate receiver domain (REC) operably coupled to a non-cognate DNA binding domain (DBD) of known functionality, and
 - b) a reporter gene under the control of a promoter containing an operator site that is bound by said DBD, such that said reporter gene is activated or repressed when said SK signals to said modified RR and said DBD binds to said DNA binding site.
- 2) The bacteria of claim 1, wherein said SK and RR are members of an OmpR-PhoB family of TCSs or a member of a NarL-FixJ family of TCSs.
- 3) The bacteria of claim 2, wherein:
 - a) said TCS is a member of the OmpR-PhoB family and said REC is separated from its wild type DBD at a crossover site between amino acids 110 and 151, said amino acids numbered according to alignment with wild type OmpR, or
 - b) said TCS is a member of the NarL-FixJ family and said REC is separated from its wild type DBD at a crossover site between amino acids 110 and 155, said amino acids numbered according to alignment with wild type NarL.
- 4) The bacteria of claim 2, wherein
 - a) said TCS is of the OmpR-PhoB family and said REC is separated from its wild type DBD at a crossover site at amino acid 122, 137, 138 or 139, said amino acids numbered according to alignment with wild type OmpR; or
 - b) said TCS is of the NarL-FixJ family, and said REC is separated from its wild type DBD at a crossover site at amino acid 113, 127, 130, 132, 142 or 154, said amino acids numbered according to alignment with wild type NarL.

- 5) The bacteria of claims 1-4, where said bacteria is gram-positive and said TCS is from a gram-negative species, or *vice versa*.
- 6) The bacteria of claims 1-5, where said bacteria is the same bacterium wherein which said TCS evolved.
- 5 7) The bacteria of claims 1-6, having one or more inducible expression vectors encoding said SK and said modified RR.
- 8) The bacteria of claims 1-7, said reporter gene being encoded on an expression vector.
- 9) The bacteria of claims 1-7, said reporter gene being integrated into a genome of said bacteria.
- 10 10) A genetically engineered bacteria, comprising:
- a) a modified two component sensor system (TCS), said TCS being a member of a OmpR-PhoB or a NarL-FixJ family of two component sensor kinases, said TCS comprising:
 - 15 i) a wild type sensor kinase (SK) comprising a ligand binding domain of unknown input operably coupled to a kinase domain, and
 - ii) a modified response regulator (RR) that is cognate to said SK, said RR comprising a cognate receiver domain (REC) operably coupled to a non-cognate DNA binding domain (DBD) of known functionality, and
 - b) a reporter gene under the control of a DNA binding site that binds said DBD,
20 such that said reporter gene is activated or repressed when said SK signals to said modified RR and said DBD binds to said DNA binding site;
 - c) wherein said REC is separated from its wild type DBD at a crossover site between amino acids 110-155, said amino acids numbered according to alignment with either wild type OmpR or wild type NarL, depending whether the TCS belongs to
25 the OmpR-PhoB or the NarL-FixJ family, respectively; and,
 - d) wherein no exogenous linker peptide is used between said REC and said non-cognate DBD.

- 11) The bacteria of claim 10, where said bacteria is gram-positive and said non-cognate DBD is from a gram-positive bacteria and said TCS is from a gram-negative species, or *vice versa*.
- 12) The bacteria of claim 10, where said bacteria is the same bacterium wherein which the SK and RR evolved.
- 13) The bacteria of claim 10, wherein said TCS is of the OmpR-PhoB family and said REC is separated from its wild type DBD at a crossover site at amino acid 122, 137, 138 or 139, said amino acids numbered according to alignment with wild type OmpR; or wherein said TCS is of the NarL-FixJ family, and said REC is separated from its wild type DBD at a crossover site at amino acid 113, 127, 130, 132, 142 or 154, said amino acids numbered according to alignment with wild type NarL.
- 14) A method of identifying an input that activates a sensor kinase, comprising:
- a) applying a test input to the bacteria of claim 1 or 10;
 - b) determining whether said input changes expression of said reporter gene; and,
 - c) repeating steps a and b until an input that changes said reporter gene expression is identified.
- 15) A method of making a modified RR, comprising:
- a) obtaining a gene for a RR from a of a NarL-FixJ family of two component sensor kinases, cutting said gene at a site corresponding to a crossover site between amino acids 110 and 155, preferably at amino acid 113, 127, 130, 132, 142 or 154, said amino acids numbered according to alignment with wild type NarL, and operably coupling an REC domain 5' of said crossover site to a non-cognate DBD of known functionality; or,
 - b) obtaining a gene for a RR from a OmpR-PhoB family of two component sensor kinases, cutting said gene at a site corresponding to a crossover site between amino acids 110 and 151, preferably at amino acid 122, 137, 138 or 139, said amino acids numbered according to alignment with wild type OmpR, and directly coupling in frame a REC domain 5' of said crossover site to a non-cognate DBD of known functionality cut at or near the same crossover site.

16) The method of claim 16, wherein said non-cognate DBD gene is cut at or near said crossover site, preferably within 1-3 amino acids of said crossover site.

17) The method of claims 15 or 16, wherein no exogenous linker peptide is used between said cognate REC and said non-cognate DBD.

5 18) A method of making a biosensor, said method comprising engineering a bacteria to have:

a) a reporter gene under the control of a promoter;

b) a two component system (TCS) comprising a sensor kinase (SK) and a cognate response regulator (RR), said TCS comprising:

10 i) an operable SK having a known input signal;

ii) an operable rewired RR having a cognate REC domain for said SK operably fused to a non-cognate DBD that changes expression of said promoter; and,

c) wherein presence of said known input signal in an environment in which said bacteria resides is detected by a change in expression of said reporter gene.

15 19) A method of making a biosensor, comprising:

a) applying a test input signal to the bacteria of claim 10;

b) determining whether said test input signal changes expression of said reporter gene;

20 c) repeating steps a and b until an input signal that changes said reporter gene expression is identified;

d) confirming that said identified input signal is the input signal for said TCS;

e) culturing said bacteria in an environment; and,

25 f) monitoring expression of said reporter gene, wherein a change in said reporter gene expression indicates that said confirmed input signal is present in said environment.

20) A method of identifying an input signal that activates a sensor kinase, comprising:

a) applying a test input signal to a genetically engineered bacteria, comprising:

- i) a modified two component sensor system (TCS), said TCS being a member of a OmpR-PhoB or a NarL-FixJ family of two component sensor kinases, said TCS comprising:
- 5 (1) a wild type sensor kinase (SK) comprising a ligand binding domain having an unknown input signal operably coupled to a kinase domain;
- (2) a modified response regulator (RR) that is cognate to said SK, said RR comprising a cognate receiver domain (REC) operably coupled to a non-cognate DNA binding domain (DBD) of known functionality;
- 10 ii) a reporter gene under the control of an operator site that binds said DBD, such that said reporter gene is activated or repressed when said SK signals to said modified RR and said DBD binds to said operator site;
- 15 iii) wherein said REC and said non-cognate DBD are fused at a crossover site between amino acids 110-155, said amino acids numbered according to alignment with either wild type OmpR or wild type NarL, depending whether the TCS belongs to the OmpR-PhoB or the NarL-FixJ family, respectively; and
- iv) wherein no exogenous linker peptide is used between said REC and said non-cognate DBD;
- b) determining whether said input signal changes expression of said reporter gene; and,
- c) repeating steps a and b until an input signal that changes said reporter gene expression
- 20 is identified, thereby identifying a cognate input signal for said TCS.
- 21) A method of making a modified RR, comprising:
- a) obtaining a gene for a RR from a of a NarL-FixJ family of two component sensor kinases, cutting said gene at a site corresponding to a crossover site between amino acids 110 and 155, preferably at amino acid 113, 127, 130, 132, 142 or 154, said
- 25 amino acids numbered according to alignment with wild type NarL, and operably coupling an REC domain 5' of said crossover site to a non-cognate DBD gene of known functionality that is cut at or near the same crossover site; or
- b) obtaining a gene for a RR from a OmpR-PhoB family of two component sensor kinases, cutting said gene at a site corresponding to a crossover site between amino

- acids 110 and 151, preferably at amino acid 122, 137, 138 or 139, said amino acids numbered according to alignment with wild type OmpR, and directly coupling in frame a REC domain 5' of said crossover site to a non-cognate DBD gene of known functionality that is cut at or near the same crossover site; and,
- 5 c) wherein no exogenous linker peptide is used between said cognate REC and said non-cognate DBD.

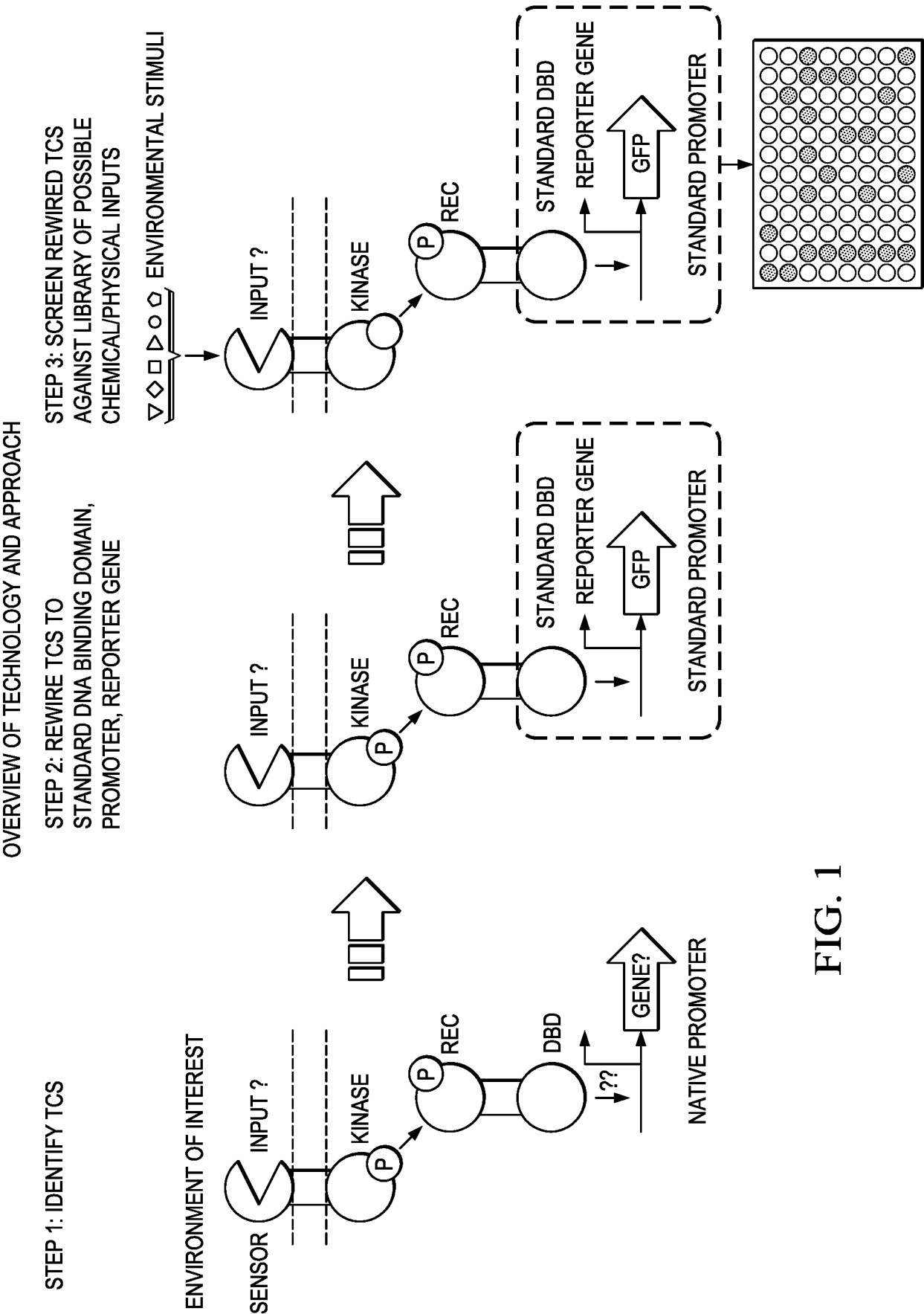
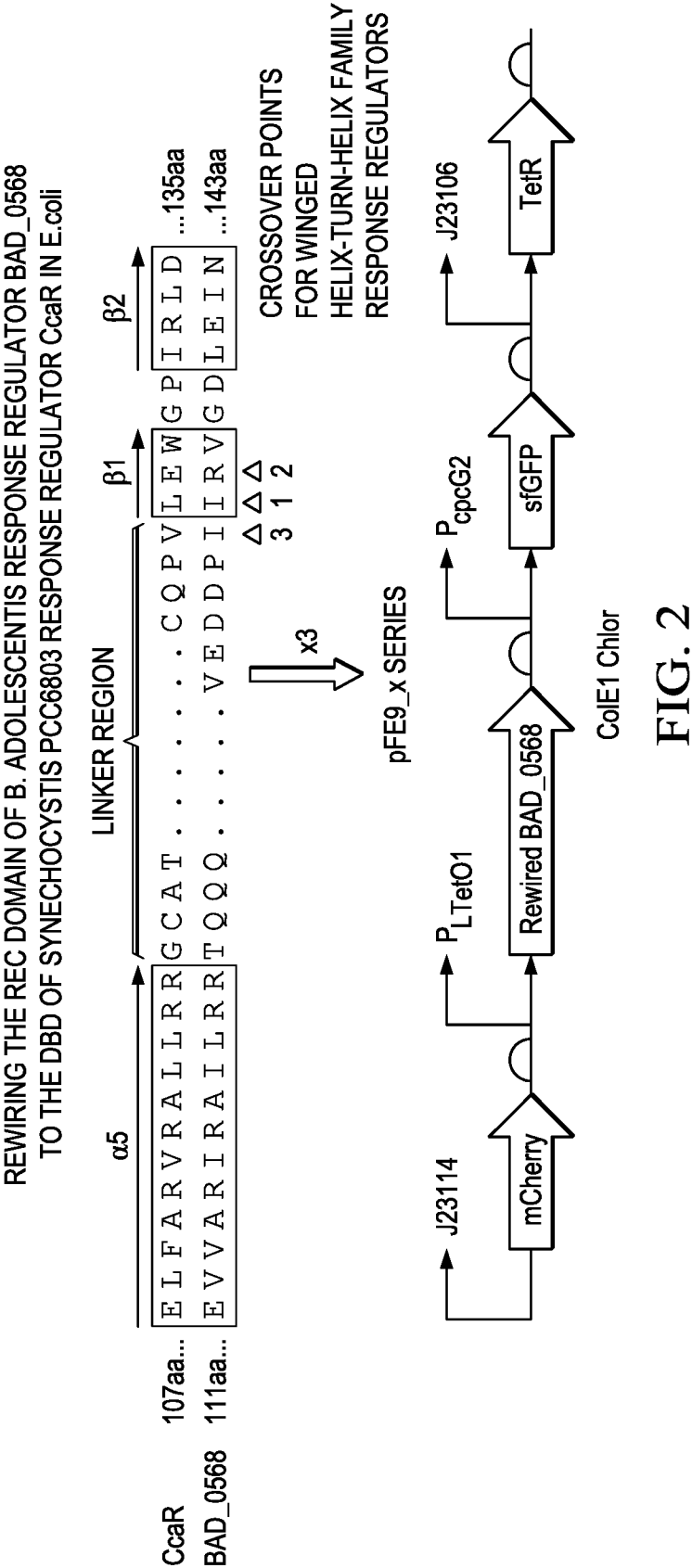


FIG. 1



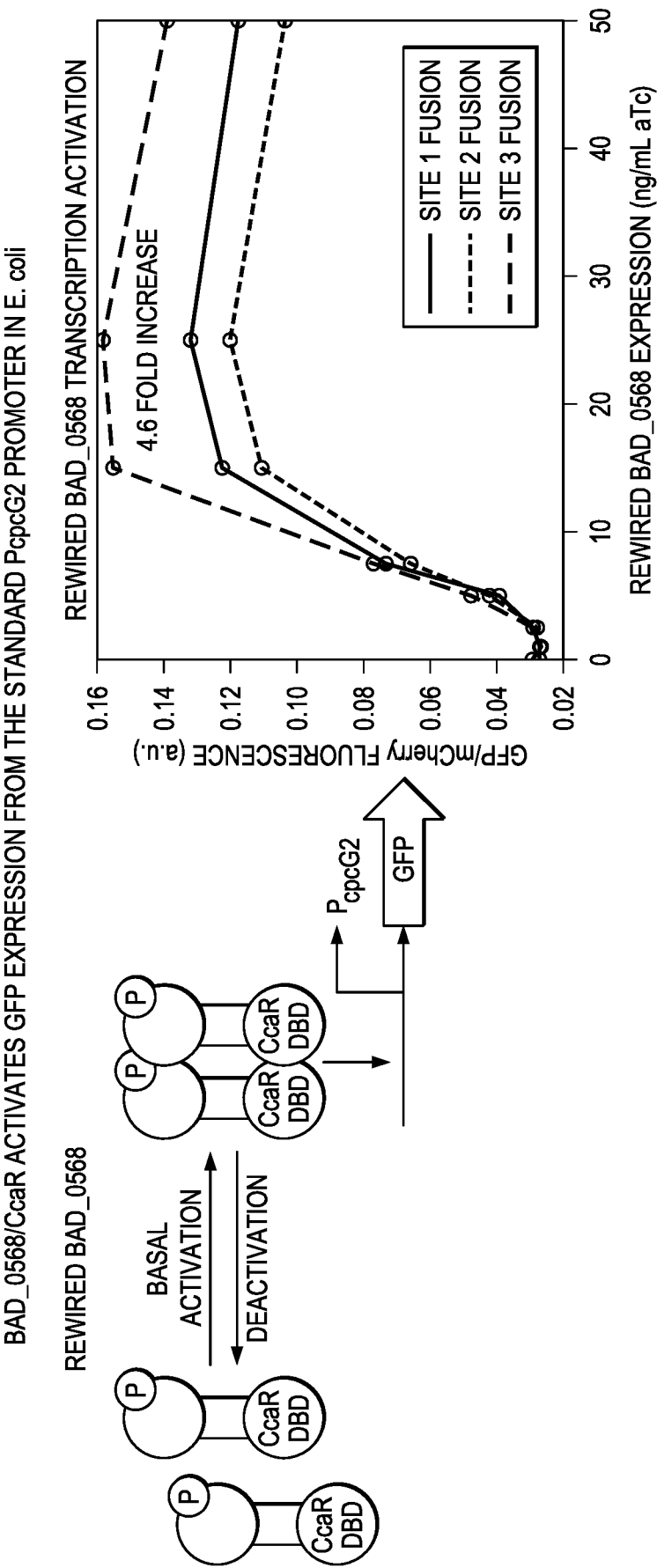


FIG. 3

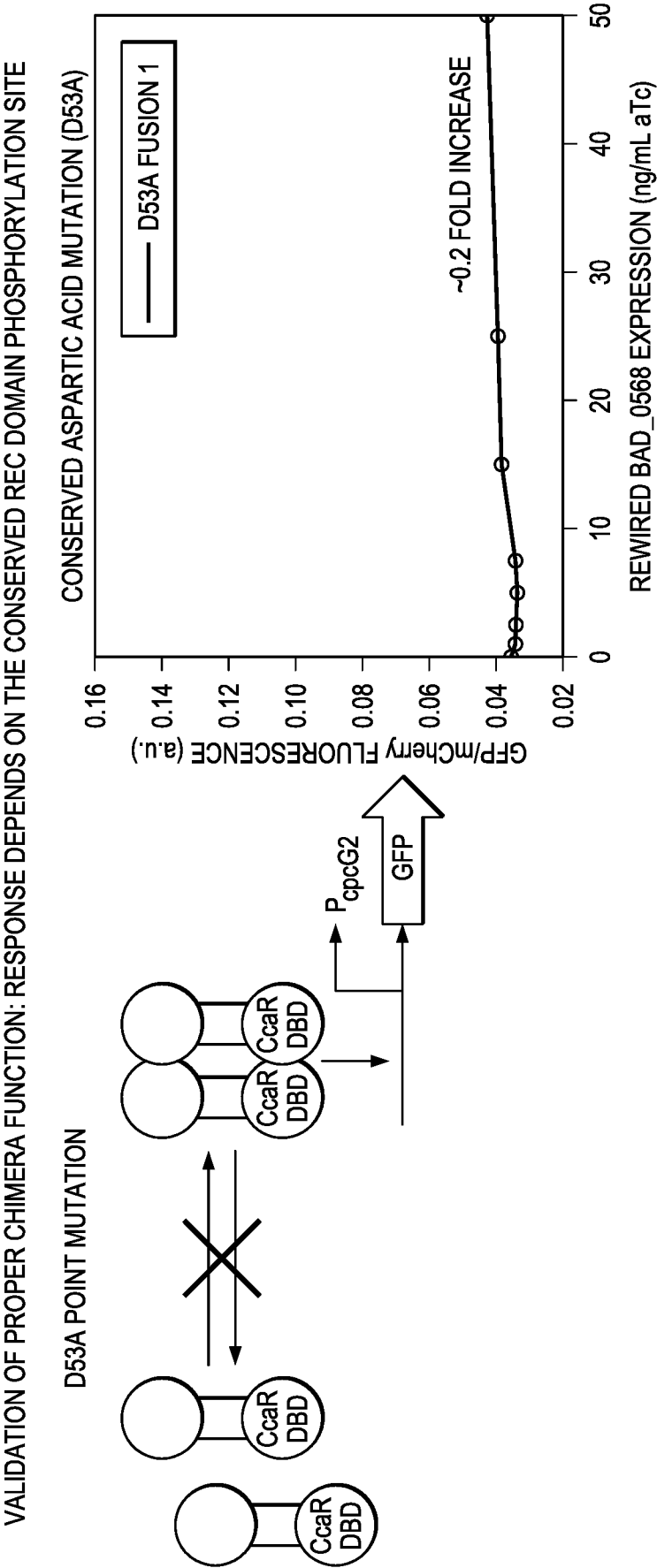
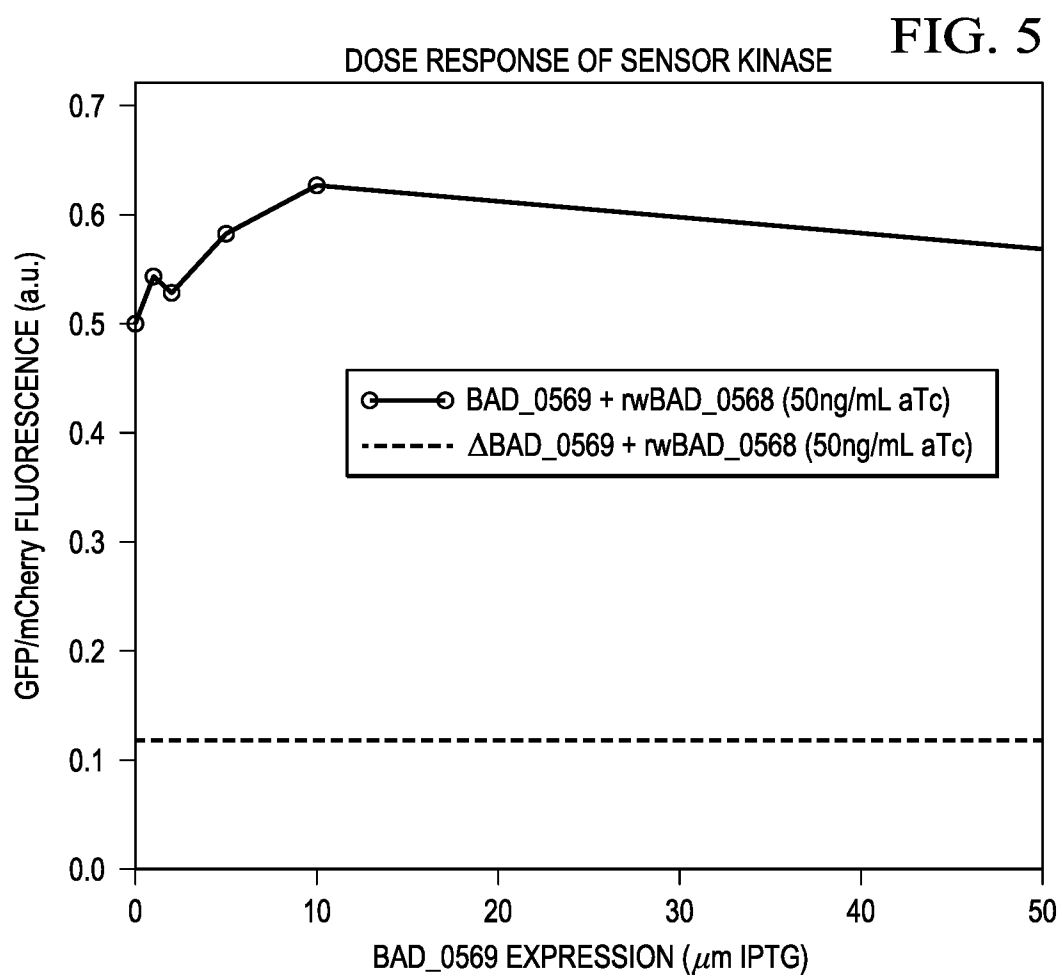


FIG. 4



VALIDATION OF PATHWAY ACTIVITY. MUTATION OF CATALYTIC HISTIDINE TO NON-CATALYTIC ASPARAGINE (N) RESIDUE

EnvZ: CONSERVED HISTIDINE AT POSITION 243 WITHIN THE HisKA REGION

HisKA REGION: ~234-289 (55aa):

...RTLIMAGVSHDLRTPLTRIRLATEMMSEQDGYLAESINKDIEECNAIIIEQFIDYLR...
 MRRLRFSPRSSFARTLLLIIVTLFASLVTTYLVVINFAILPSLQQFNKVLAYEVRMLMTDKLQLEDGTQLVVPAPFRREI
 YRELGISLYSNEAAEEAGLRWAQHYEFLSHQMAQQLGPTFVRVEVNKSSPVWLKTWLSPNIWVRVPLTEIHQGDSP
 FRYTLAIMLLAIGGAWLFIRIQNRPLVDLEHAALQVGKGIIPPLREYGASEVRSVTRAFNHMAAGVKQLADDRTLIMAG
 VSHDLRTPLTRIRLATEMMSEQDGYLAESINKDIEECNAIIIEQFIDYLRGTQEMPMMADLNAVIGEVIAAESGYEREIE
 TALYPGSEIEVKMHPLSIKRAVANMVVNAARYGNWIKVSSGTEPNRAWFQVEDDGGPGIAPEQRKHLFQPFVVRGDSARTIS
 GTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVPVTRAQGTKEG*

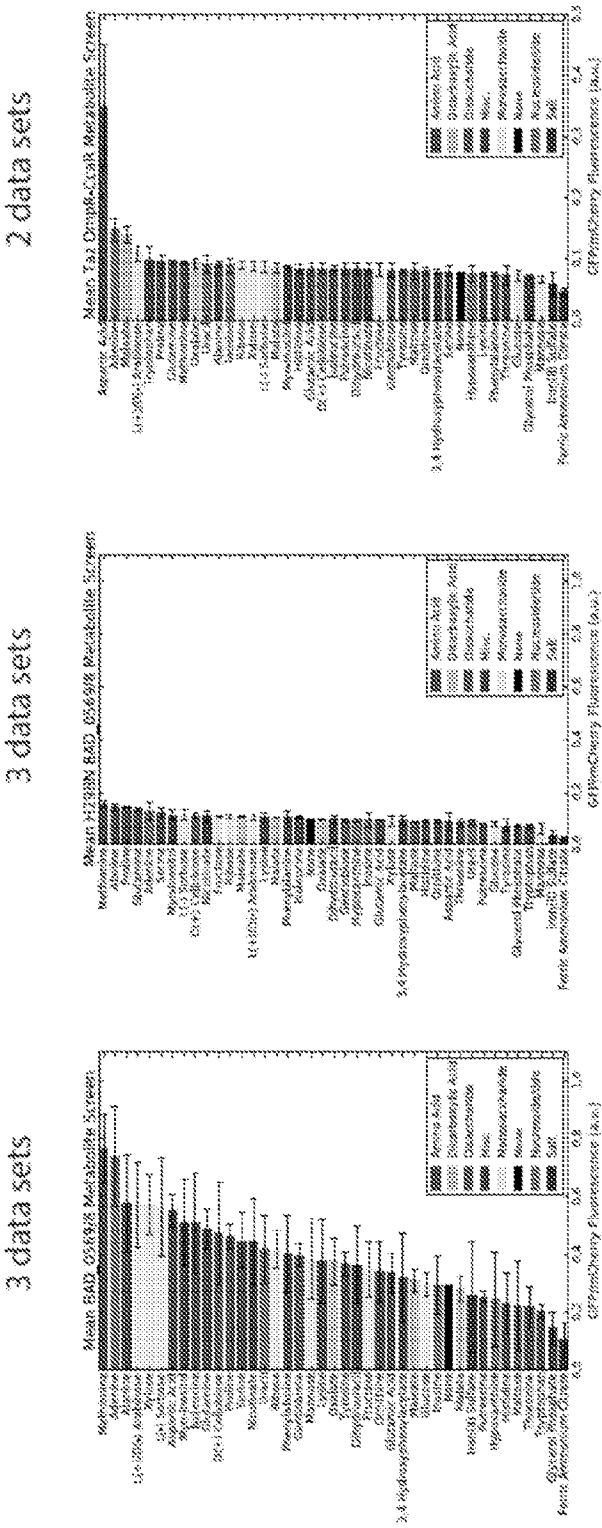
BAD_0569: CONSERVED HISTIDINE IN THE HisKA REGION

HisKA REGION: 287-353 (66aa):

...EKMKRFVSDASHELRTPLAAIHGYAELYKMQRDMPGALERADESIEHIEASSARMTVLVEDLLSLAR...
 MQPPRSLPKQNKVWSRFTRRIQAIPLSTKLVTCTIVLLTIGTIGISFIRTLVGNYLLOKTDQLINQAQLIFNSMDSLS
 SDTGDDGRSLMNTYYVEVRDSEYKSTGAGSVPMRLRDGVVSEPSLPADGSDGVTLGQPFTRAVVHIITTSRTPDHSIMQA
 AQSPWRVVALPWNEKTRTGQVKDVGFFIGLSLSDQIDTANTLTRFCAMVGIAVVLIGAILGLTLVQSTLAPLKRIEKTA
 AKIAAGDLSQRPVDPPESTEVGSLMSLNTMLTRIEESFHAQEEETTEKMKRFVSDASHELRTPLAAIHGYAELYKMQRDM
 PGALERADESIEHIEASSARMTVLVEDLLSLARLDEGRGIDITQQVKLTSVVRDAADDLHALDPDRGISCGQVVLQPGTD
 MDHPAQAFQNGQMPQIELKGDASRLRQVVTNIVGNIHRYTPADSPVEISMGVLPASISPELSRMPSPNEQSLRHLVEAI
 EVGQSMQVGMNYAIVRFSDHGPVPPPEARSKIIFERFYTADPSRARQKGGTGLGMAIAQSVVKAHHGFCASGSEGTGLIL
 TVVLPPIAPVEPKPQPI TASENRKNEKNRKS*

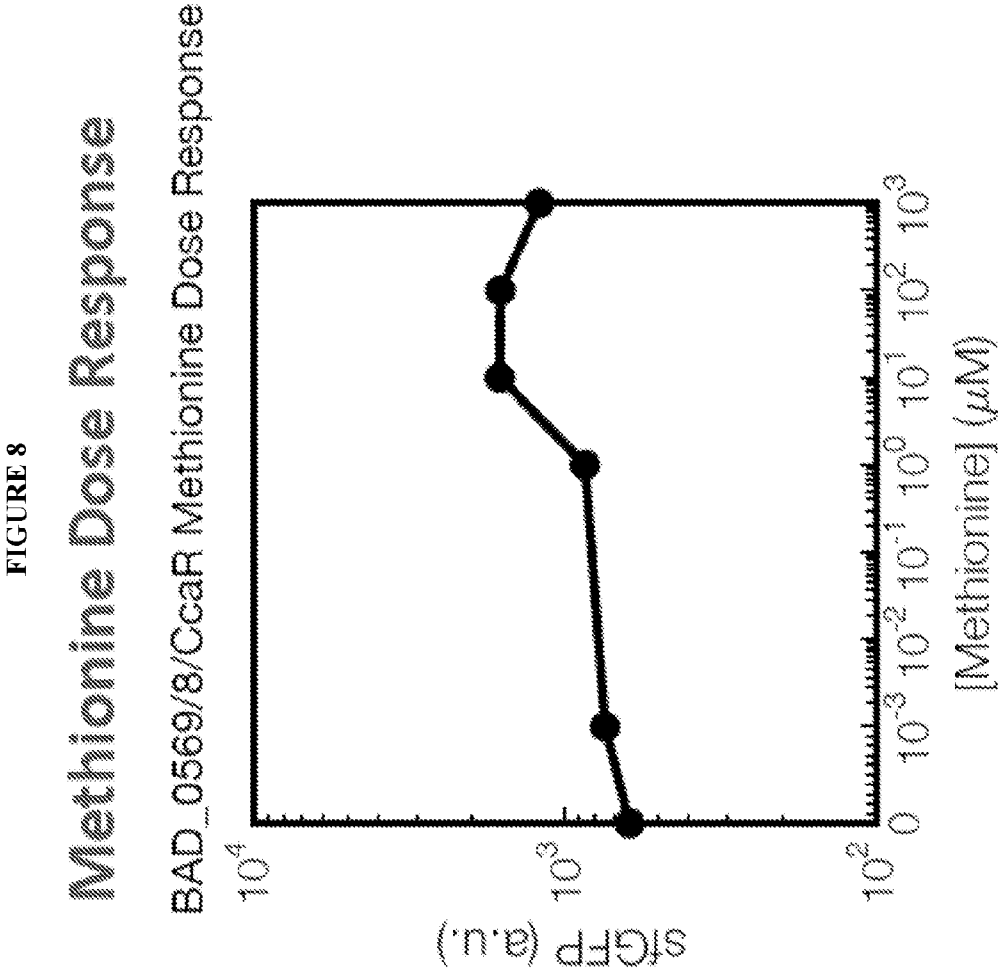
FIG. 6

FIGURE 7
Response of rewired BAD_0569/8/CcaR pathway to panel of
mammalian gut metabolites in E. coli



Note: The "None" conditions were only run once
*Malate for the third run was removed for H298N BAD_0569/8
since the mCherry value was questionably low (600 a.u. versus a normal 3000 a.u.)
*Cysteine, Valine, and Leucine were removed for all conditions since no growth was observed

8/27



9/27

FIGURE 9: Amino acid sequences of wild-type (native) and chimeric RRs used herein	
Native full length wHTH Response Regulators used: Key No Style = REC Domain, Bold = Linker, <u>Underline</u> = DBD. Wild-type RRs contain linkers between the two domains, but these are native sequences, not exogenous linkers.	
OmpR (E. coli) ACCN: NP_417864 SEQ ID NO: 1	
MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA VIAFGKFKLNLGTREMFREDEPMPLTSGEFAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQ ISRLRRMVEEDPAHPRIYQTVWGLGYVFPDGSKA	
CcaR (Synechocystis PCC6803) ACCN: WP_010874216 SEQ ID NO: 2	
MRILLVEDDLPLAETLAEALSDQLYTVDIATDASLAWDYASRLEYDLVILDVMLPELDGITLCQKWRSHS YLMPILMMTARDTINDKITGLDAGADDYVVKPVDLGELFARVRALLRRGCATCQPVLEWGPIRLDPST YEVSYDNEVLSLTRKEYSILELLLRNGRRVLSRSMIIDSIIWKLESPPEEDTVKVHVRSRLRQKLKSAGLSA DAIETVHGIGYRLANLTEKSLCQGKN	
CopR (Synechocystis PCC 6803) ACCN: WP_010873936 SEQ ID NO: 3	
MRLLLVEDEPDGLMALEKALRRRENYVVDWVQDGNLAWSYLDQGWVNYTLAIFDWMVPGLSGLELC QKLRGQRSSLPILMLTAKDQIADRVGLDAGADDYLIKPFPGMAELLARLSLQRRSPQLQVQVQ WWLDYGTFAVVTPEQARITLTAKEFQLLEYFMKHPQQILSSEIKNQLWALSASTSNVVAQVRLLR RKLEESYSHNLIETVYGLGYRFQPHPTAEQ	
ManR (Synechocystis PCC6803) ACCN: WP_010872074 SEQ ID NO: 4	
MANILLVDDENALTEPLSKALGHQGHGTIDVADQGKTGLAMAIAGQYDLLILDWMLPQVSGLEICRQIRIL GHSTPVLFLTAKDTLDDRVLGADAGDDYLIKPFELRELLARVRALLRRQSHGETITETLGA ~~VKNNLLS~~ VNNVSLDVANQVAYCQGGRIALSEKEVALLTLFLQAPGQILSHEEYSHLWPGESPSSNVLAALVRL RRKIEQPNAPRLINSVYGKGYCFEAN	
BceR (B. subtilis) ACCN: WP_004399109 SEQ ID NO: 5	
MFKLLIEDDESLEFHEIKDRLTGWSYDVYGIQDFSQVLQFEAAVNPDCVIIDVQLPKFDGFHWCRLIRS RSNVPIFLSSRDHPADMVMSMLQGADDFIQKPFHFDVLIQAMFRRVHHYNTEPSTIKTWCGAAV DAEQNLVSNDKGSVELTKNEMFILKQLEQKNKIVSREELIRSLWNDERFVSDNTLTNVNRLRKKLDA LQLGAYIETKVGQGYIAKEEDKFYD	
PsdR (B. subtilis) ACCN: WP_003244535 SEQ ID NO: 6	
MYRILLVEDDERIASLLGGHLQKYGYEVKIAEQNLNDIKLEFAEMKPDVLDDINLPFFDGFYWCRQIRTS NAPIIFISARTDELNQVMAIENGDDYITKPFHLEVVMKIKSVLRRTYGEYSPSLPQESRIVELGGLTIY PDQNEAEWNSVRILFSQKEFQLLSIFVREHKKIVSRDELLEALWDDVDVDDNTLTNVNRLRRLKEN AGLTDICISTIRGQGYQFQVNRKDEAEC	
YxdJ (B. subtilis) ACCN: WP_003243527 SEQ ID NO: 7	
MNKIMIVEDSEDIRGLLQNYLEKYGYQTVAADF ~~TAVLDVFLREKPDVVLLDINLPAYDGYWCRQIRQ~~ HSTSPIIFISARSGEMDQVMAIENGDDYIEKPFSDIVLAKIKSQIRRAYGEYAAKQGEKVVEYAGVQ LFVERFELRFQDEKSELSKESKLEVLLEGEKVTSRDRLMEKTDWTDIFIDNTLVN ~~VYITRLRKKLRE~~ LNAPVSIEAVRGEQYQLRAQS	
BAD_0569 (B. adolescentis) ACCN: WP_003808701 SEQ ID NO: 8	
MSKPIEASIVVDDEPSIRELLVASLHFAGFEVNTAASGSEAIEVIERLQPDILVDVMLPDI ~~DGFTVTRRI~~ RQEGITTPVLYLTARDDTQDKVMGLTVGGDDYVTKPFSLEEVVARIRAILRR ~~TQQQVEDDPIIRVGDLEI~~ NEDSHDVSRAQPIDLSPTYEKLLRYLMDNEGRVLSKAQILDHVWQYDWGGDAAIVESYISYLRKKVD GIVIEDENGDKHKVTPLIETKRGIGYMIRAPK	
Rewired wHTH Response Regulators demonstrated: Key No Style = N-Terminal RR (REC Domain Donor), <u>Underline</u> = C-Terminal RR (DBD Donor)	
OmpR-CcaR (122aa) SEQ ID NO: 9	
MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRGCATCQPVLEWGPI RLDPSTYEVSYDNEVLSLTRKEYSILELLLRNGRRVLSRSMIIDSIIWKLESPPEEDTVKVHVRSRLRQKLK SAGLSADAIETVHGIGYRLANLTEKSLCQGKN	
OmpR-CcaR (137aa): SEQ ID NO: 10	
MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA <u>VLEWGPIRLDPSTYEVSYDNEVLSLTRKEYSILELLLRNGRRVLSRSMIIDSIIWKLESPPEEDTVKVHVR</u> <u>SLRQKLKSAGLSADAIETVHGIGYRLANLTEKSLCQGKN</u>	
OmpR-CcaR (138aa) SEQ ID NO: 11	
MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA <u>VLEWGPIRLDPSTYEVSYDNEVLSLTRKEYSILELLLRNGRRVLSRSMIIDSIIWKLESPPEEDTVKVHVR</u> <u>SLRQKLKSAGLSADAIETVHGIGYRLANLTEKSLCQGKN</u>	

10/27

FIGURE 9: Amino acid sequences of wild-type (native) and chimeric RRs used herein	
OmpR-CcaR (139aa) SEQ ID NO: 12	<u>MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC</u> <u>RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA</u> <u>VIAWGPIRLDPSTYEVSVDNEVLSLTRKEYSILELLLRNGRRVLSRSMIIDSIIWKLESPPEEDTVKVHVR</u> <u>SLRQKLKSAGLSADAETVHGIGYRLANLTEKSLCQGKN</u>
OmpR-CcaR (140aa) SEQ ID NO: 13	<u>MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC</u> <u>RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA</u> <u>VIAFGPIRLDPSTYEVSVDNEVLSLTRKEYSILELLLRNGRRVLSRSMIIDSIIWKLESPPEEDTVKVHVRS</u> <u>LRQKLKSAGLSADAETVHGIGYRLANLTEKSLCQGKN</u>
OmpR-CcaR (142aa) SEQ ID NO: 14	<u>MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC</u> <u>RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA</u> <u>VIAFGKIRLDPSTYEVSVDNEVLSLTRKEYSILELLLRNGRRVLSRSMIIDSIIWKLESPPEEDTVKVHVRS</u> <u>LRQKLKSAGLSADAETVHGIGYRLANLTEKSLCQGKN</u>
OmpR-CcaR (150aa) SEQ ID NO: 15	<u>MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC</u> <u>RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA</u> <u>VIAFGKFKLNLGTREVSVDNEVLSLTRKEYSILELLLRNGRRVLSRSMIIDSIIWKLESPPEEDTVKVHVR</u> <u>SLRQKLKSAGLSADAETVHGIGYRLANLTEKSLCQGKN</u>
OmpR-ManR (137aa) SEQ ID NO: 16	<u>MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC</u> <u>RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA</u> <u>VLSVNNVSLDVANQVAYCQGQRIALSEKEVALLTLFLQAPGQILSHEEIIYSHLWPGESPSSNVLAALV</u> <u>RLLRKIEQPNAPRLINSVYGKGYCFEAN</u>
OmpR-PsdR (137aa) SEQ ID NO: 17	<u>MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC</u> <u>RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA</u> <u>VVELGGLTIYPDQNEAEWNSVRILFSQKEFQLLSIFVREHKKIVSRDELLEALWDDVDFVDDNTLTNVN</u> <u>NRLRRKLENAGLTDICISTIRGQGYQFQVNRKDEAEC</u>
OmpR-YxdJ (137aa) SEQ ID NO: 18	<u>MQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAQMDRLLTRESFHLMVLDLMLPGEDGLSIC</u> <u>RRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPPELLARIRAVLRRQANELPGAPSQEEA</u> <u>VVEYAGVQLFVERFELRFQDEKSELSKKEVLLERGEKVTSRDRLMEKTWDTDIFIDDNTLNVI</u> <u>TRLRKKLRELNAPVSIEAVRGEQYQLRAQS</u>
CcaR-CopR (137aa) SEQ ID NO: 19	<u>MRILLVEDDLPLAETLAEALSDQLYTVDIATDASLAWDYASRLEYDLVILDVMLPELDGITLCQKWRSHS</u> <u>YLMPILMMTARDTINDKITGLDAGADDYVVKPVDLGELFARVRALLRRGCATCQPVLSVNNVSLDVAN</u> <u>QVAYCQGQRIALSEKEVALLTLFLQAPGQILSHEEIIYSHLWPGESPSSNVLAALVRLLRKIEQPNAP</u> <u>RLINSVYGKGYCFEAN</u>
CcaR-ManR (137aa) SEQ ID NO: 20	<u>MRILLVEDDLPLAETLAEALSDQLYTVDIATDASLAWDYASRLEYDLVILDVMLPELDGITLCQKWRSHS</u> <u>YLMPILMMTARDTINDKITGLDAGADDYVVKPVDLGELFARVRALLRRGCATCQPVLSVNNVSLDVAN</u> <u>QVAYCQGQRIALSEKEVALLTLFLQAPGQILSHEEIIYSHLWPGESPSSNVLAALVRLLRKIEQPNAP</u> <u>RLINSVYGKGYCFEAN</u>
CcaR-BceR (137aa) SEQ ID NO: 21	<u>MRILLVEDDLPLAETLAEALSDQLYTVDIATDASLAWDYASRLEYDLVILDVMLPELDGITLCQKWRSHS</u> <u>YLMPILMMTARDTINDKITGLDAGADDYVVKPVDLGELFARVRALLRRGCATCQPVKTWCGAAVDAE</u> <u>QNLVSNDKGSVELTKNEMFILKQLIEQKNKIVSREELIRSLWNDERFVSDNTLTNVNRLRKKLDALQL</u> <u>GAYIETKVGQGYIAKEEDKFYD</u>
CcaR-PsdR (137aa) SEQ ID NO: 22	<u>MRILLVEDDLPLAETLAEALSDQLYTVDIATDASLAWDYASRLEYDLVILDVMLPELDGITLCQKWRSHS</u> <u>YLMPILMMTARDTINDKITGLDAGADDYVVKPVDLGELFARVRALLRRGCATCQPVVVELGGLTIYPDQ</u> <u>NEAEWNSVRILFSQKEFQLLSIFVREHKKIVSRDELLEALWDDVDFVDDNTLTNVNRLRKKLENAGLT</u> <u>DCISTIRGQGYQFQVNRKDEAEC</u>
CcaR-YxdJ (137aa): SEQ ID NO: 23	<u>MRILLVEDDLPLAETLAEALSDQLYTVDIATDASLAWDYASRLEYDLVILDVMLPELDGITLCQKWRSHS</u> <u>YLMPILMMTARDTINDKITGLDAGADDYVVKPVDLGELFARVRALLRRGCATCQPVVEYAGVQLFVER</u> <u>FELRFQDEKSELSKKEVLLERGEKVTSRDRLMEKTWDTDIFIDDNTLNVIITRLRKKLRELNAP</u> <u>VSIEAVRGEQYQLRAQS</u>

11/27

FIGURE 9: Amino acid sequences of wild-type (native) and chimeric RRs used herein	
BAD_0569-CcaR (137aa): SEQ ID NO: 24	
MSKPIEASIVVVDDEPSIRELLVASLHFAGFEVNTAASGSEAEVIERLQPDILVLDVMLPDIDGFTVTRRI RQEGITTPVLYLTARDDTQDKVMGLTVGGDDYVTKPFSLEEVVARIRAILRRTQQQVEDDPVLEWGPI RLDPSTYEVSYDNEVLSLTRKEYSILELLLRNGRRVLSRSMIIDSIIWKLESPPEEDTVKVHVRSRLRQKLK SAGLSADAIETVHGIGYRLANLTEKSLCQGKN	
Native HTH Response Regulators:	
i) Key No Style = REC Domain, Bold = Linker, <u>Bold Underline</u> = Flexible alpha 6, Underline = DBD	
NarL (E. coli) ACCN: NP_415739 SEQ ID NO: 25	
MSNQEPATILLIDHPMLRTGVKQLISMADITVVGGEASNGEQGIELAESLDPDLILLDLNMPGMNGLE TLDKLREKSLSGRIVVFSVSNHEEDVVTALKRGADGYLLKDMEPEDLLKALHQAAGEMVLSEALTPV LAASLRANRATTERDVNQLTPRERDILKIAQGLPNKMIARRLDITESTVKVHVKHMLKMKMLKSRVEA AVVWHQERIF	
UhpA (E. coli) ACCN: NP_418125 SEQ ID NO: 26	
MITVALIDDHLIVRSGFAQLLGLEPDLQVVAEFGSGREALAGLPGRGVQVCICDISMPDISGLELLSQLP KGMATIMLSVHDSALVEQALNAGARGFLSKRCSPELIAAVHTVATGGCYLTPDIAIKLASGRQDPLT KRERQVAEKLAQGMVKEIAAELGLSPKTVHVRANLMEKLGVSNDVELARRMFDGW	
YdfI (B. subtilis) ACCN: WP_003244318 SEQ ID NO: 27	
MNKVLIVDDHLVVREGLKLLIETNDQYTIIGEAENGKAVRLADELEPDILMDLYMPEMSGLEAIKQIKE KHDTPIILTTYNEDHLMIEGIELGAKGYLLKDTSSSETLFHTMDAAIRGNVLLQPDILKRLQEIQFERMKK QRNETQLTEKEVIVLKAIAGLKSIAFDLGVSERTVKSRLTSIYNKLGANSRTEAVTIAMQKGILTIDN	
LiaR (B. subtilis) ACCN: WP_003243201 SEQ ID NO: 28	
MIRVLLIDDHVMVRMGAAFLAEPDIEVIGESDGSSEGVRLAVELSPDVILMDLVMEGMDGIEATKQI CRELSDPKIIVLTSFIDDDKVYPVIEAGALSYLLKTSKAAEIAAIRAASKGEPKLESKVAGKVLRLRHS GENALPHESLTKRELEILCLIAEGKTNKEIGEELFITIKTVKTHITNLSKLDVSDRTQAAVYAHRNHLVN	
FusR (E. coli) ACCN: AAG54714 SEQ ID NO: 29	
MIRVVLVDDHVVRSGFAQLLSLEDDLEVIGQYSSAAQAWSALIRDDVNVAVIDIAMPDENGLSLLKRL RAQKPQFRAILSIYDAPTFVQSALDAGASGYLTKRCGPPEELVQAVRSVGLGGHYLCADAIARLGGG QPAQALEILTPREREVFELLVKGDSVKEIAFKLELSHKTIVHVHRANVLGKLNCHSTIELVHFALDHLLA GH	
Rewired HTH Response Regulators:	
No Style = N-Terminal RR (REC Domain Donor), Underline = C-Terminal RR (DBD Donor)	
NarL-YdfI (131aa) SEQ ID NO: 30	
MSNQEPATILLIDHPMLRTGVKQLISMADITVVGGEASNGEQGIELAESLDPDLILLDLNMPGMNGLE TLDKLREKSLSGRIVVFSVSNHEEDVVTALKRGADGYLLKDMEPEDLLKALHQAAGEMVLSPDILKRL QEIQFERMKKQRNETQLTEKEVIVLKAIAGLKSIAFDLGVSERTVKSRLTSIYNKLGANSRTEAVTIA MQKGILTIDN	
UhpA-YdfI (131aa) SEQ ID NO: 31	
MITVALIDDHLIVRSGFAQLLGLEPDLQVVAEFGSGREALAGLPGRGVQVCICDISMPDISGLELLSQLP KGMATIMLSVHDSALVEQALNAGARGFLSKRCSPELIAAVHTVATGGCYLTPDILKRLQEIQFERMK KQRNETQLTEKEVIVLKAIAGLKSIAFDLGVSERTVKSRLTSIYNKLGANSRTEAVTIAMQKGILTID N	
UhpA-LiaR (131aa) SEQ ID NO: 32	
MITVALIDDHLIVRSGFAQLLGLEPDLQVVAEFGSGREALAGLPGRGVQVCICDISMPDISGLELLSQLP KGMATIMLSVHDSALVEQALNAGARGFLSKRCSPELIAAVHTVATGGCYLTSKVAGKVLRLRHSG ENALPHESLTKRELEILCLIAEGKTNKEIGEELFITIKTVKTHITNLSKLDVSDRTQAAVYAHRNHLVN	
FusR-YdfI (129aa) SEQ ID NO: 33	
MIRVVLVDDHVVRSGFAQLLSLEDDLEVIGQYSSAAQAWSALIRDDVNVAVIDIAMPDENGLSLLKRL RAQKPQFRAILSIYDAPTFVQSALDAGASGYLTKRCGPPEELVQAVRSVGLGGHYLQPDILKRLQEIQF ERMKKQRNETQLTEKEVIVLKAIAGLKSIAFDLGVSERTVKSRLTSIYNKLGANSRTEAVTIAMQKG ILTIDN	

12/27

FIGURE 10

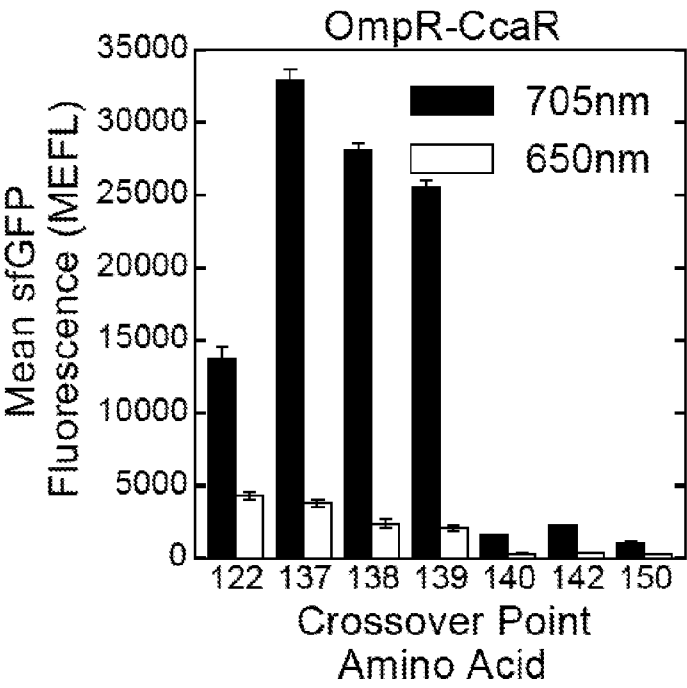
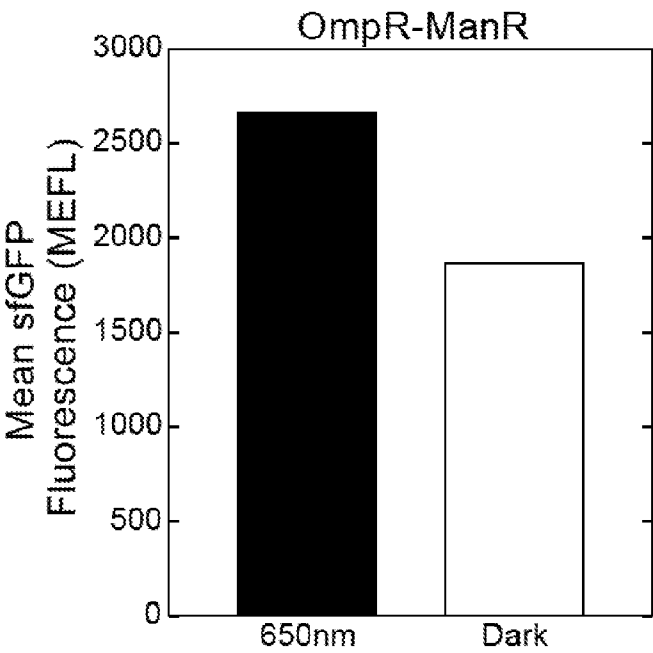


FIGURE 11



13/27

FIGURE 12

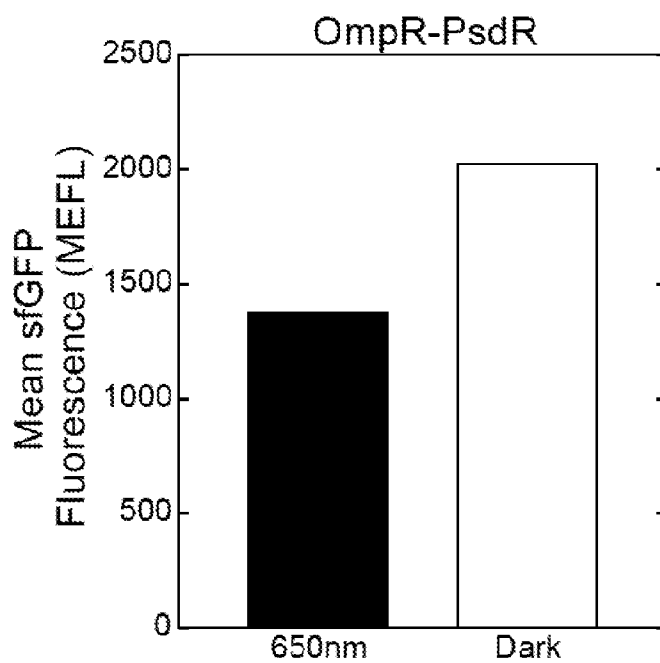
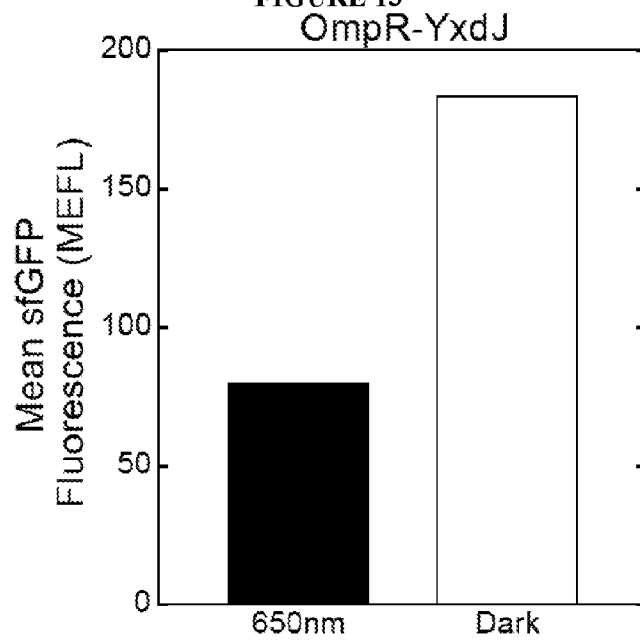


FIGURE 13



14/27

FIGURE 14

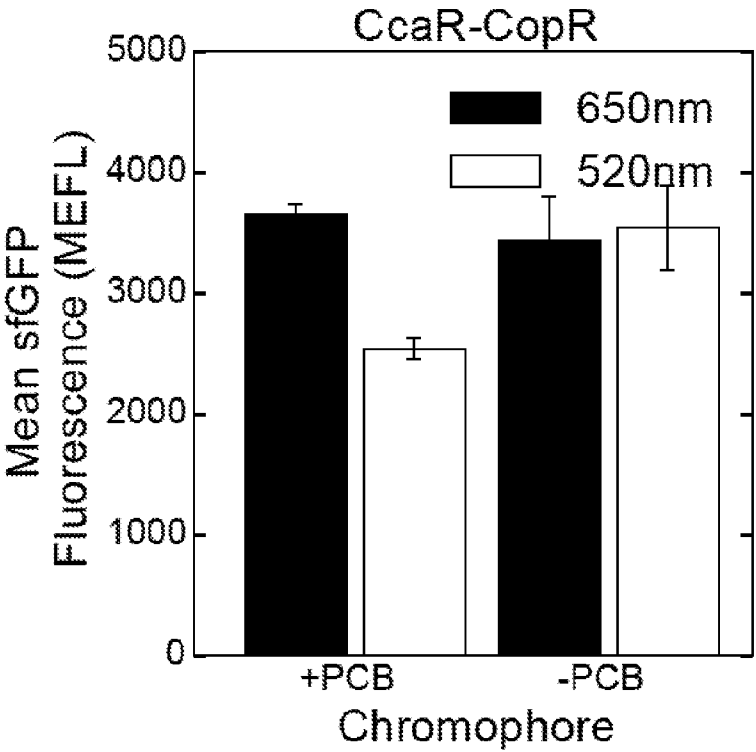
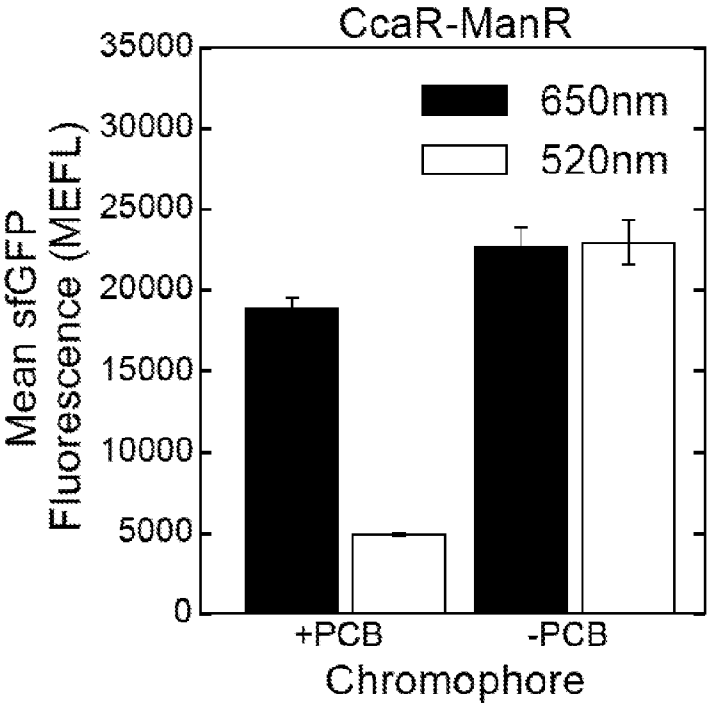


FIGURE 15



15/27

FIGURE 16
CcaR-BceR

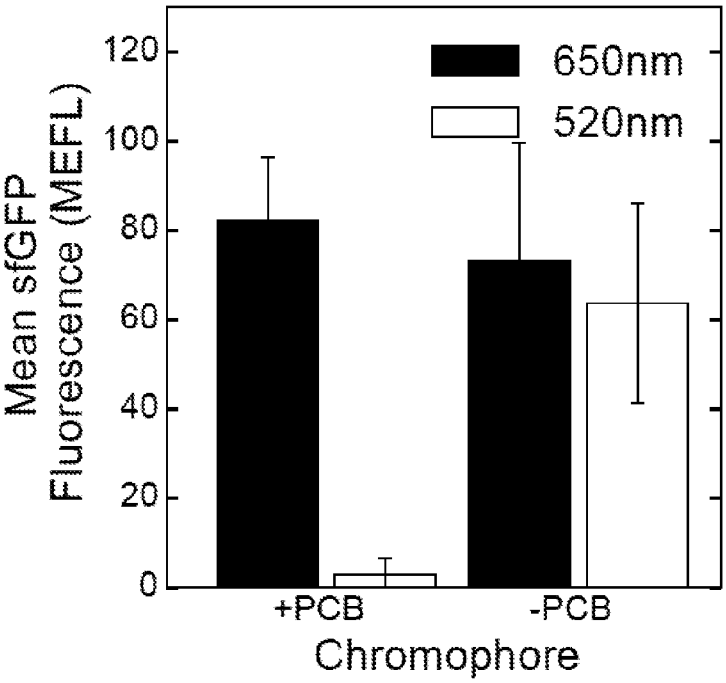
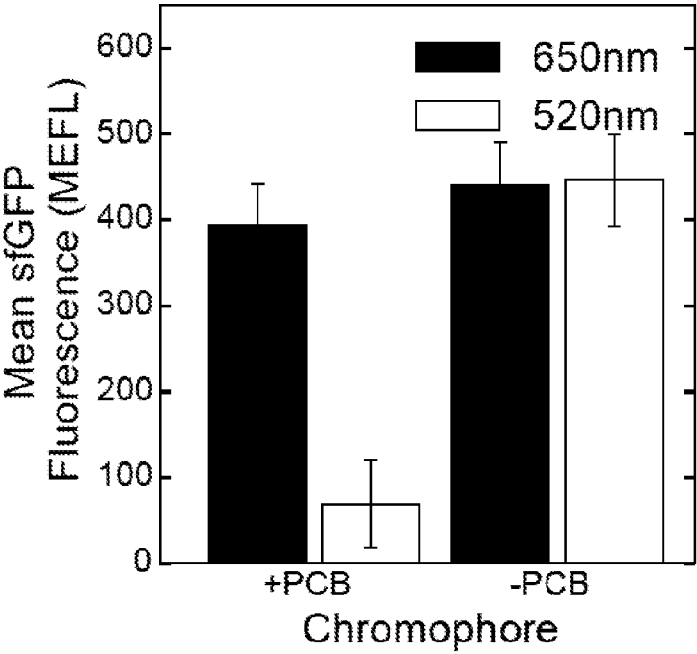


FIGURE 17
CcaR-PsdR



16/27

FIGURE 18

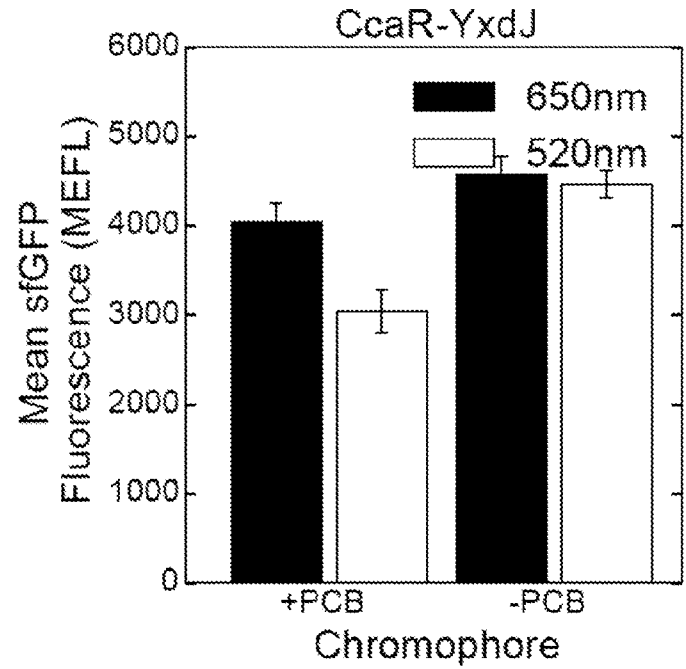
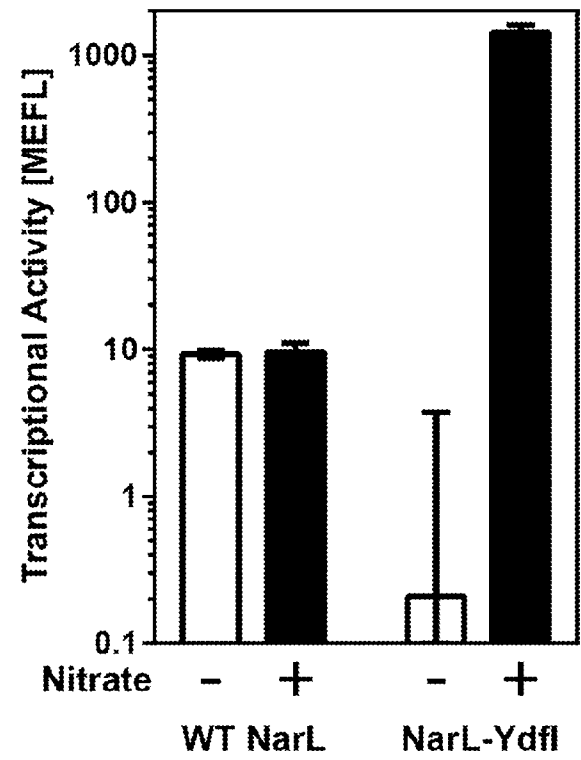


FIGURE 19



17/27

FIGURE 20

G6P Transfer Function
UhpB/UhpA-YdfI | $P_{ydfJ-115}$

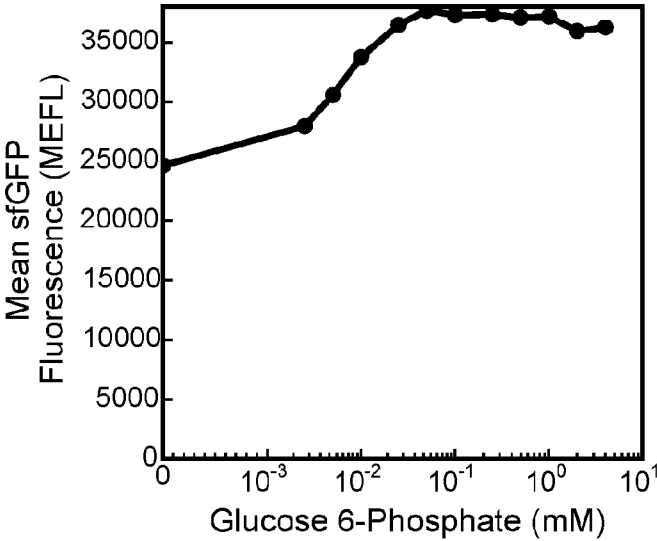


FIGURE 21

G6P Transfer Function
UhpB/UhpA-LiaR | P_{yhc-86}

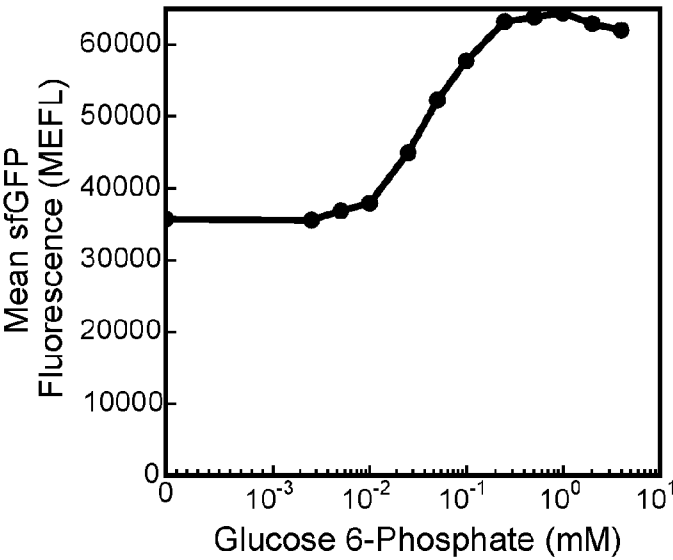


FIGURE 22

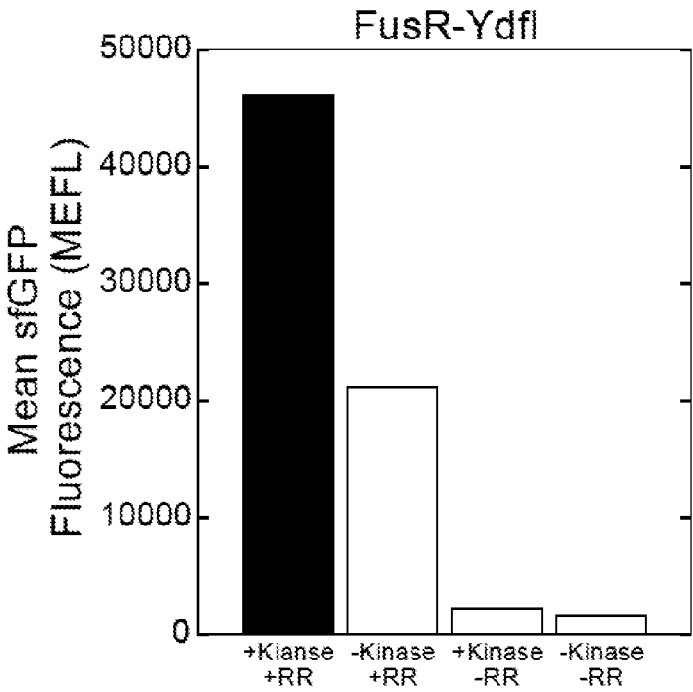
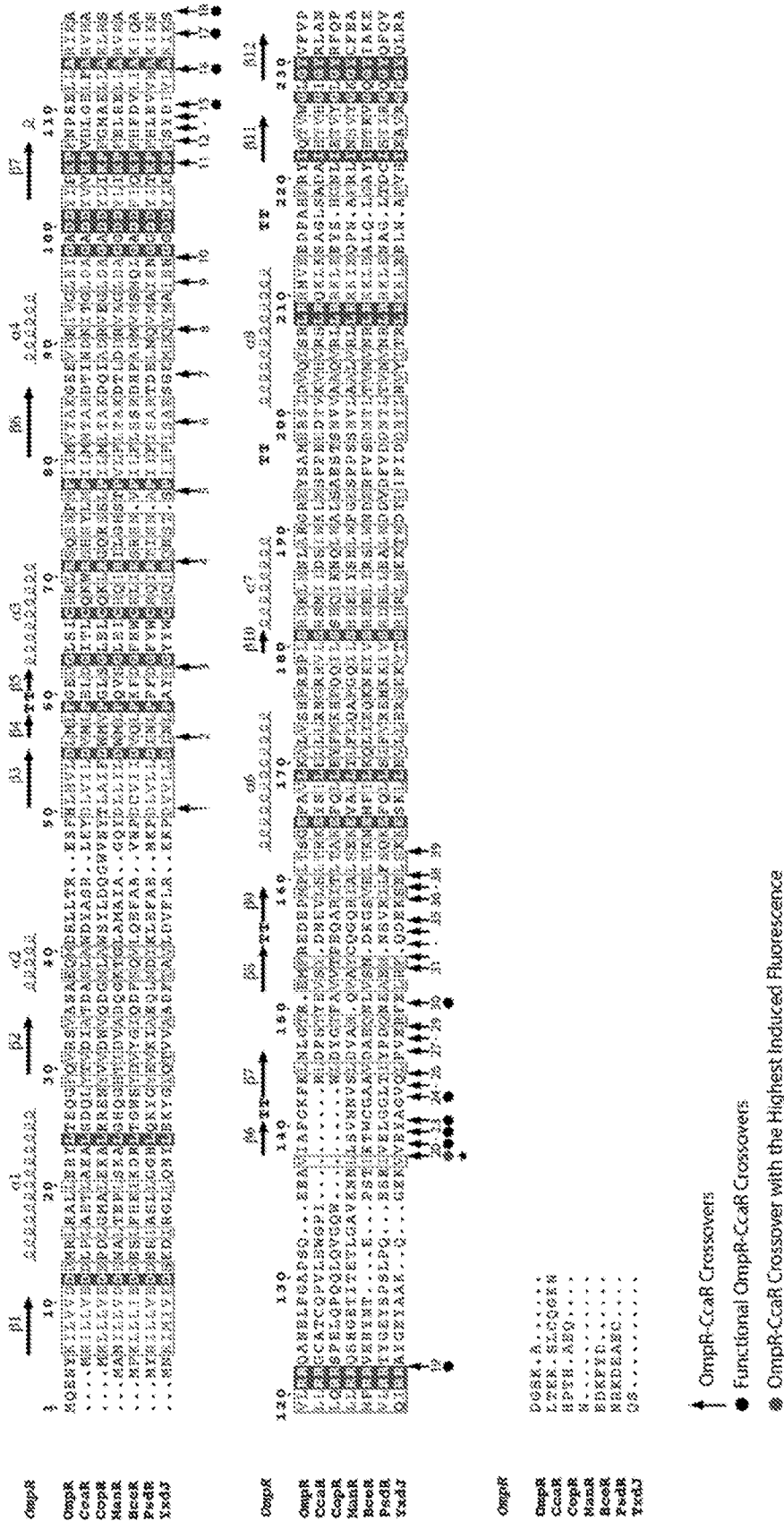


FIGURE 23



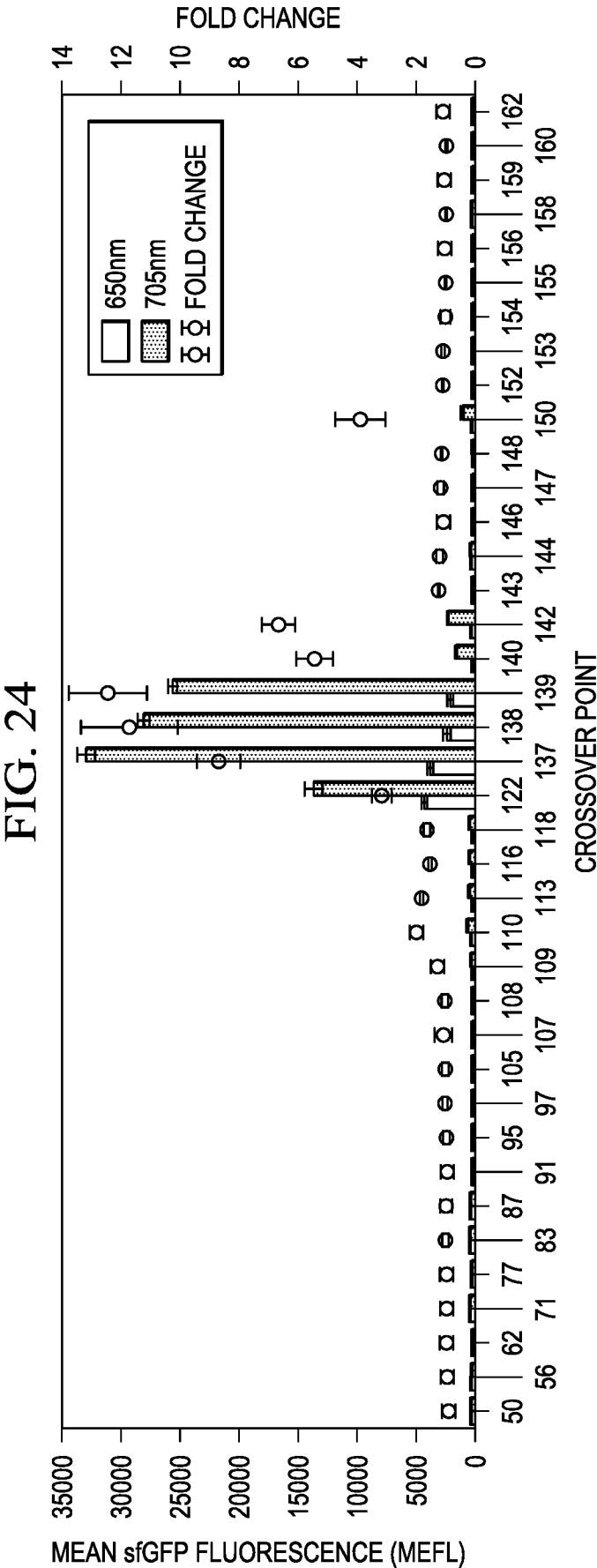


FIGURE 25

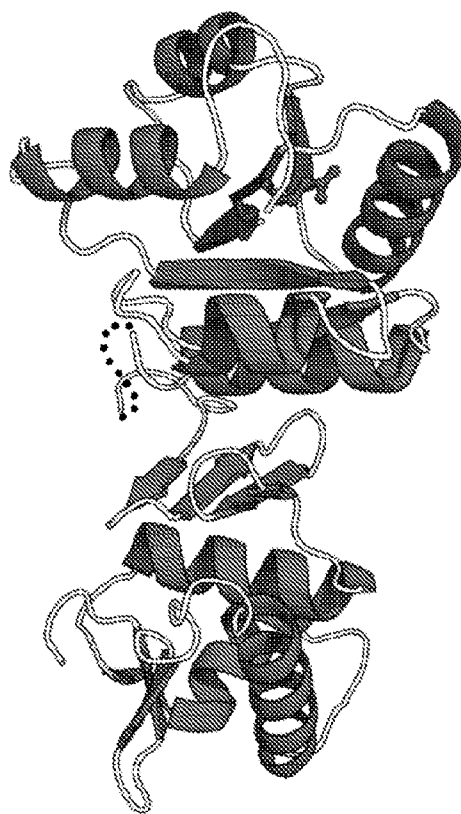
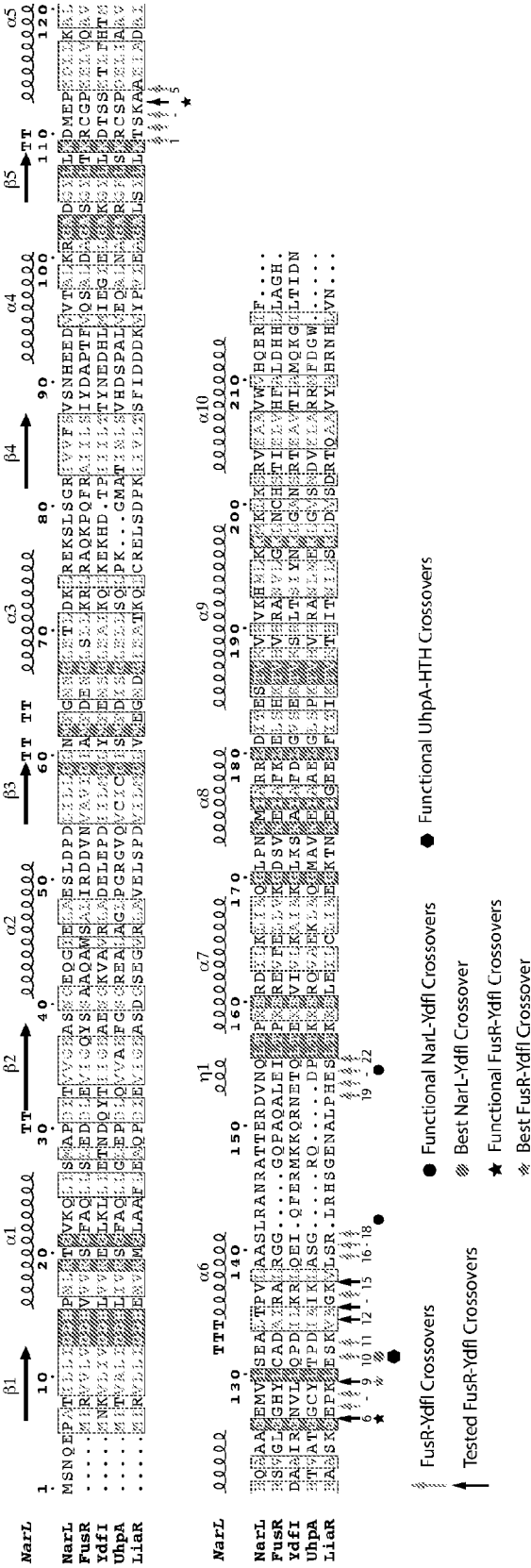


FIGURE 26



23/27

FIGURE 27

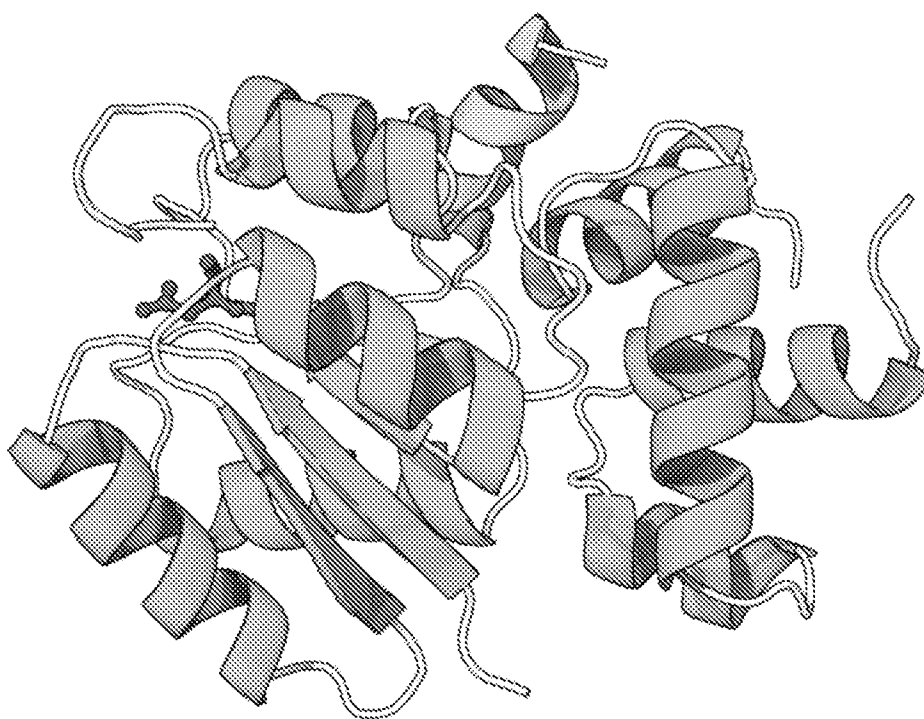


FIGURE 28

Partial listing of embodiments, any one of which can be combined with any other one or more or portions thereof.

A genetically engineered bacteria, comprising:
 a modified heterologous two component sensor system (TCS) from a different species of bacteria, said TCS being a member of a members of a OmpR-PhoB or NarL-FixJ family of two component sensor kinases, said TCS comprising:
 a sensor kinase (SK) comprising a ligand binding domain operably coupled to a kinase domain, and
 a modified response regulator (RR) that is cognate to said SK, said RR comprising a cognate receiver domain (REC) operably coupled to a heterologous DNA binding domain (DBD) of known functionality, and
 a reporter gene under the control of a DNA binding site that binds said DBD, such that said reporter gene is expressed when said SK activates said modified RR and said DBD binds to said DNA binding site.

A genetically engineered bacteria, comprising:
 a modified heterologous two component sensor system (TCS) from a different species of bacteria, said TCS comprising:
 a sensor kinase (SK) comprising a ligand binding domain operably coupled to a kinase domain, and
 a modified response regulator (RR) that is cognate to said SK, said RR comprising a cognate receiver domain (REC) operably coupled to a non-cognate DNA binding domain (DBD) of known functionality, and
 a reporter gene under the control of a DNA binding site that binds said DBD, such that said reporter gene is expressed when said SK activates said modified RR by phosphorylating said REC domain and said DBD binds to said DNA binding site.

A genetically engineered bacteria, comprising:
 a modified heterologous two component sensor system (TCS) from a different species of bacteria, said TCS being a member a OmpR-PhoB or a NarL-FixJ family of two component sensor kinases, said TCS comprising:
 a sensor kinase (SK) comprising a ligand binding domain operably coupled to a kinase domain, and
 a modified response regulator (RR) that is cognate to said SK, said RR comprising a cognate receiver domain (REC) operably coupled to a non-cognate DNA binding domain (DBD) of known functionality, and
 a reporter gene under the control of a DNA binding site that binds said DBD, such that said reporter gene is expressed when said SK activates said modified RR and said DBD binds to said DNA binding site.

A genetically engineered bacteria, comprising:
 a modified two component sensor system (TCS), said TCS comprising:
 a wild type sensor kinase (SK) comprising a ligand binding domain operably coupled to a kinase domain, and
 a modified response regulator (RR) that is cognate to said SK, said RR comprising a cognate receiver domain (REC) operably coupled to a non-cognate DNA binding domain (DBD) of known functionality, and
 a reporter gene under the control of a promoter element containing an operator site that is bound by said DBD, such that said reporter gene is activated or repressed when said SK signals to said modified RR and said DBD binds to said DNA binding site.

FIGURE 28

Partial listing of embodiments, any one of which can be combined with any other one or more or portions thereof.

A genetically engineered bacteria, comprising:
a modified two component sensor system (TCS), said TCS being a member of a OmpR-PhoB or a NarL-FixJ family of two component sensor kinases, said TCS comprising:

a sensor kinase (SK) comprising a ligand binding domain of unknown input operably coupled to a kinase domain, and
a modified response regulator (RR) that is cognate to said SK, said RR comprising a cognate receiver domain (REC) operably coupled to a non-cognate DNA binding domain (DBD) of known functionality, and

a reporter gene under the control of a DNA binding site that binds said DBD, such that said reporter gene is activated or repressed when said SK signals to said modified RR and said DBD binds to said DNA binding site;
wherein said REC is separated from its wild type DBD at a crossover site between amino acids 110-155, said amino acids numbered according to alignment with either wild type OmpR or wild type NarL, depending whether the TCS belongs to the OmpR-PhoB or the NarL-FixJ family, respectively.

A genetically engineered bacteria, said bacteria expressing a two component sensor system (TCS) comprising i) a sensor kinase gene comprising a ligand binding domain operably coupled to a kinase domain, and ii) a response regulator gene comprising a receiver domain operably coupled to an heterologous DNA binding domain (DBD), said bacteria also comprising a DNA binding site that binds said DBD that is operably coupled to a reporter gene.

A genetically engineered gram positive bacteria expressing a two component sensor system (TCS) from a gram negative bacteria, said TCS comprising a sensor kinase gene and a response regulator gene. Preferably the SK of any bacteria herein described is wild type and the non-cognate DBD is compatible with the host bacteria.

A genetically engineered gram positive bacteria expressing a two component sensor system (TCS) from a gram negative bacteria, said TCS comprising a sensor kinase (SK) gene and a response regulator (RR) gene (or vice versa). Preferably, the DBD of the RR is a non-cognate DBD operably fused to the REC domain, which remains cognate to the SK.

A bacteria as described, a single expression vector encoding both said SK and said modified RR.

A bacteria as described, said reporter gene being encoded on an expression vector.

A bacteria as described, said reporter gene being integrated into a genome of said bacteria.

A bacteria as described, said response regulator gene encoding a receiver domain operably coupled to a heterologous DBD, said bacteria also comprising a DNA targeted by said DBD that is operably coupled to a reporter gene.

A bacteria as described, where said bacteria is gram positive and said TCS is from a gram negative species, or vice versa.

A bacteria as described, where said bacteria is gram positive and said TCS is from a gram negative species, or vice versa.

A bacteria as described, wherein said SK and RR are members of a OmpR-PhoB family of TCSs or a member of a NarL-FixJ family of TCSs.

A bacteria as described, wherein said SK and RR are members of a OmpR-PhoB family of TCSs or a member of a NarL-FixJ family of TCSs.

A bacteria as herein described, where said bacteria is the same bacterium wherein which said TCS evolved, and a native SK and RR of said TCS is knocked out.

FIGURE 28

Partial listing of embodiments, any one of which can be combined with any other one or more or portions thereof.

A bacteria as herein described, wherein:

said TCS is a member of a OmpR-PhoB family and said REC is separated from its wild type DBD at a crossover site between amino acids 110 and 151, preferably 122, 137, 138 or 139, said amino acids numbered according to alignment with wild type OmpR, or

said TCS is a member of a NarL-FixJ family and said REC is separated from its wild type DBD at a crossover site between amino acids 110 and 155, preferably 113, 127, 130, 132, 142 or 154, said amino acids numbered according to alignment with wild type NarL. Preferably, wherein no exogenous linker peptide is used between said REC and said non-cognate DBD. Preferably the non-cognate DBD is also cut at the same crossover site, although sites nearby may suffice.

A biosensor for a ligand, said biosensor comprising a bacteria expressing a two component sensor system (TCS), said TCS comprising a sensor kinase gene encoding a ligand binding domain that binds said ligand operably coupled to a kinase domain, and a response regulator encoding a receiver domain that is activated by said kinase domain operably coupled to a heterologous DNA binding domain that can change expression of a reporter gene, also found in said bacteria.

A method of identifying an input signal that activates a sensor kinase, comprising applying a test input signal to a genetically engineered bacteria, comprising:

- a modified two component sensor system (TCS), said TCS being a member of a OmpR-PhoB or a NarL-FixJ family of two component sensor kinases, said TCS comprising:
 - a wild type sensor kinase (SK) comprising a ligand binding domain having an unknown input signal operably coupled to a kinase domain;
 - a modified response regulator (RR) that is cognate to said SK, said RR comprising a cognate receiver domain (REC) operably coupled to a non-cognate DNA binding domain (DBD) of known functionality;
- a reporter gene under the control of an operator site that binds said DBD, such that said reporter gene is activated or repressed when said SK signals to said modified RR and said DBD binds to said operator site;
- wherein said REC is separated from its wild type DBD at a crossover site between amino acids 110-155, said amino acids numbered according to alignment with either wild type OmpR or wild type NarL, depending whether the TCS belongs to the OmpR-PhoB or the NarL-FixJ family, respectively; and
- wherein no exogenous linker peptide is used between said REC and said non-cognate DBD;

determining whether said input signal changes expression of said reporter gene, and

repeating applying and determining steps until an input signal that changes said reporter gene expression is identified, thereby identifying a cognate input signal for said TCS.

A method of making a biosensor, comprising

- applying a test input signal to a bacteria as herein described,
- determining whether said test input signal changes expression of said reporter gene, and
- repeating applying and determining steps until an input signal that changes said reporter gene expression is identified,
- confirming that said identified input signal is the input signal for said TCS;
- culturing said bacteria in an environment, and

FIGURE 28

Partial listing of embodiments, any one of which can be combined with any other one or more or portions thereof.

monitoring expression of said reporter gene, wherein a change in said reporter gene expression indicates that said confirmed input signal is present in said environment.

A method of making a biosensor, said method comprising engineering a bacteria to have:
 a reporter gene under the control of a promoter,
 a heterologous two component system (TCS) comprising a sensor kinase (SK) and a cognate response regulator (RR), said TCS comprising:
 an operable SK having a known input ligand;
 an operable rewired RR having a cognate REC domain for said SK operably fused to a non-cognate DBD that activates said promoter,
 wherein presence of an input ligand in an environment in which said bacteria resides is detected by expression of said reporter gene.

A method of making a biosensor, said method comprising engineering a bacteria to have:
 a reporter gene under the control of a promoter,
 a heterologous two component system (TCS) comprising a sensor kinase (SK) and a cognate response regulator (RR), said TCS comprising:
 an operable SK having a known input signal;
 an operable rewired RR having a cognate REC domain for said SK operably fused to a non-cognate DBD that changes expression of said promoter,
 wherein presence of said known input signal in an environment in which said bacteria resides is detected by a change in expression of said reporter gene.

A method of making a modified RR, comprising obtaining a gene for a RR from a of a OmpR-PhoB or NarL-FixJ family of two component sensor kinases, cutting said gene at a site corresponding to a domain separation site between amino acids 110 and 151, preferably at amino acid 122, 137, 138 or 139, said amino acids numbered according to alignment with wild type OmpR, and operably coupling an REC domain 5' of said domain separation site to a non-cognate DBD of known functionality. Preferably, the non-cognate DBD is cut at or near the same crossover site.

A method of screening for a ligand that activates a sensor kinase, comprising i) applying a test ligand to the bacteria described herein, ii) determining whether said ligand activates said reporter gene, and iii) repeating steps i and ii until a ligand that activates said reporter gene expression is identified.

A method of screening for an input that activates a sensor kinase, comprising
 g) applying a test input to the bacteria as described,
 h) determining whether said input activates expression of said reporter gene, and
 i) repeating steps a and b until an input that activates said reporter gene expression is identified.