



(51) International Patent Classification:

C12N 15/09 (2006.01) C12Q 1/68 (2006.01)
C12N 15/63 (2006.01) C12Q 1/02 (2006.01)

(21) International Application Number:

PCT/US2010/058341

(22) International Filing Date:

30 November 2010 (30.11.2010)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/265,219 30 November 2009 (30.11.2009) US

(71) Applicants (for all designated States except US):

TRUSTEES OF BOSTON UNIVERSITY [US/US];
One Silber Way, Boston, Massachusetts 02215 (US).
MASSACHUSETTS INSTITUTE OF TECHNOLOGY [US/US]; 77 Massachusetts Avenue, Cambridge,
Massachusetts 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): COLLINS, James,
J. [US/US]; 118 Glen Avenue, Newton, Massachusetts
02459 (US). LU, Timothy Kuan-Ta [US/US]; 303 Third
Street, Apt. 210, Cambridge, Massachusetts 02142 (US).

(74) Agents: EISENSTEIN, Ronald, I. et al.; Nixon Peabody
LLP, 100 Summer Street, Boston, Massachusetts 02110
(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, 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, IS, JP, KE, KG, KM, KN, KP,
KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD,
ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,
NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, 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.

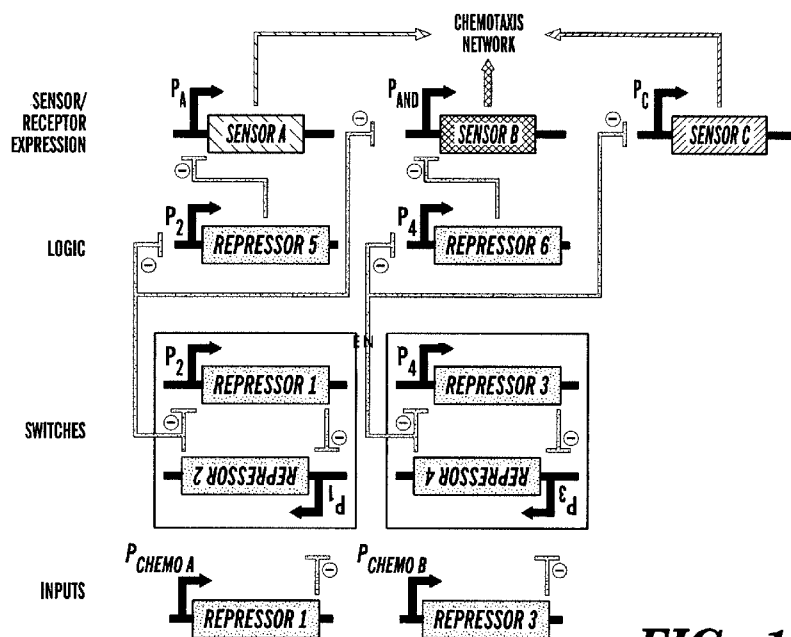
(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG,
ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, 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, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

[Continued on next page]

(54) Title: BIOLOGICAL CIRCUIT CHEMOTACTIC CONVERTERS



(57) Abstract: Described herein are novel biological circuit chemotactic converter that utilize modular components, such as genetic toggle switches and single invertase memory modules (SIMMs), for detecting and converting external inputs, such as chemoattractants, into outputs that allow for autonomous chemotaxis in cellular systems. Flexibility in these biological circuit chemotactic converter is provided by combining individual modular components, i.e., SIMMs and genetic toggle switches, together. These biological converter switches can be combined in a variety of network topologies to create network systems that regulate chemotactic responses based on the combination and nature of input signals received.

FIG. 1B



-
- | | |
|--|--|
| <ul style="list-style-type: none">— <i>as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))</i>— <i>as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))</i> | Published: <ul style="list-style-type: none">— <i>without international search report and to be republished upon receipt of that report (Rule 48.2(g))</i>— <i>with sequence listing part of description (Rule 5.2(a))</i> |
|--|--|

BIOLOGICAL CIRCUIT CHEMOTACTIC CONVERTERS

SEQUENCE LISTING

[0001] The instant application contains a Sequence Listing which has been submitted via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on November 23, 2010, is named 70158651.txt, and is 254000 bytes in size.

FIELD OF THE INVENTION

[0002] The present invention relates to biological chemotactic converters and methods of use thereof.

CROSS REFERENCE TO RELATED APPLICATIONS

[0003] This application claims the benefit under 35 U.S.C. §119(e) of U.S. Provisional Patent Application Serial No: 61/265,219 filed on November 30, 2009, the contents of which are incorporated herein in their entirety by reference.

BACKGROUND OF THE INVENTION

[0004] Circuits and circuit designs are typically based on electrical and electronic components and properties and are useful for a variety of functions. An electrical circuit is an interconnection of electrical elements, such as resistors, inductors, capacitors, transmission lines, voltage sources, current sources, and switches, and when it also contains active electronic components is known as an electronic circuit. Electronic circuits can usually be categorized as analog, digital or mixed-signal (a combination of analog and digital) electronic circuits. The basic units of analog circuits are passive (resistors, capacitors, inductors, and memristors) and active (independent power sources and dependent power sources). Components such as transistors may be represented by a model containing passive components and dependent sources. In digital electronic circuits, electric signals take on discrete values, which are not dependent upon time, to represent logical and numeric values. These values represent the information that is being processed. The transistor is one of the primary components used in discrete circuits, and combinations of these can be used to create logic gates. These logic gates can then be used in combination to create a desired output from an input.

[0005] In contrast, while some biological circuits have been developed, the utility of these circuits has been minimal, and it has been difficult to replicate the versatility and flexibility of standard electronic circuits. Many challenges remain in advancing synthetic biology from low-level gene circuitry to higher-order networks. Controlling the state of cells is a difficult but important task in biotechnology. For example, controlling transcriptional activity in cells currently relies on either constitutive promoters which are hardwired to have distinct activities or by using transcriptional activators or repressors which can be tuned by the application of inducer molecules. However, inducer

molecules are expensive, can be difficult to control in complex environments used in biotechnology, and can be toxic. Constitutive promoters cannot be easily shut off and therefore can be difficult to use when there are toxic products that are being expressed. Furthermore, constitutive promoters severely limit the flexibility of biological systems to adapt to different conditions.

SUMMARY OF THE INVENTION

[0007] We have created and describe herein novel biological chemotactic converters that provide modular systems for converting external inputs, such as chemoattractants, into outputs that allow for autonomous chemotaxis of biological systems, such as cells. To provide such functionalities, the biological circuit chemotactic converters of the invention comprise modular components that facilitate feedback loops and are capable of responding to external inputs by expressing sensors or receptors useful for detecting and responding to external signals, such as chemoattractants. For example, a biological circuit chemotactic converter can be used to convert an input chemoattractant signal into upregulation of a specific sensor receptor that permits a chemotactic response to the chemoattractant.

[0008] Provided herein are biological chemotactic converters for use in biological systems, and systems and methods for responding to specific chemotactic inputs. These biological chemotactic converters are extensible, highly modular and can function with a variety of combinations of various component parts, such as toggle switches and recombinases. Depending on the combinations of modules and component parts used in the biological chemotactic converters described herein, a chemotactic converter can respond to one or to multiple inputs.

[0006] Thus, described herein, in some aspects, are biological circuit chemotactic converters comprising at least one input module $(IM)_n$, at least one genetic toggle switch $(TS)_n$, at least one logic module $(LM)_n$, and at least two sensor modules $(SM_A \text{ and } SM_B)_{n+1}$, where $n \geq 1$.

[0007] In some aspects, the input module comprises an inducible promoter sequence (iP_A) operably linked to a repressor sequence encoding a repressor protein (R_A) . In some embodiments of these aspects, the input module further comprises at least one ribosome binding sequence. In some embodiments of these aspects, the input module further comprises at least one terminator sequence. In some embodiments of these aspects, the input module further comprises at least one degradation tag sequence. In some embodiments of these aspects, the input module further comprises at least one sequence encoding an iRNA molecule specific for a protein encoded by another module. In some embodiments of these aspects, the repressor sequence encodes for an engineered zinc-finger protein.

[0008] In other embodiments of these aspects, the input module consists essentially of an inducible promoter sequence (iP_A) operably linked to a repressor sequence (R_A) . In some embodiments of these aspects, the input module consists of an inducible promoter sequence (iP_A) operably linked to a repressor sequence (R_A) .

[0009] In some aspects, the genetic toggle switch comprises a first repressible promoter sequence (rP_1) that drives expression of a second repressor sequence (R_2), and a second repressible promoter sequence (rP_2) that drives expression of a first repressor sequence (R_1), i.e., (rP_1 - R_2 and rP_2 - R_1). In other embodiments of these aspect, the genetic toggle switch consists essentially of a first repressible promoter sequence (rP_1) that drives expression of a second repressor sequence (R_2), and a second repressible promoter sequence (rP_2) that drives expression of a first repressor sequence (R_1), i.e., (rP_1 - R_2 and rP_2 - R_1). In some embodiments of the aspect, the genetic toggle switch consists of a first repressible promoter sequence (rP_1) that drives expression of a second repressor sequence (R_2), and a second repressible promoter sequence (rP_2) that drives expression of a first repressor sequence (R_1), i.e., (rP_1 - R_2 and rP_2 - R_1).

[0010] In some such embodiments, the first and second repressor sequences encode first and second repressor molecules, such that the first repressible promoter sequence is inhibited by the first repressor molecule and the second repressible promoter sequence is inhibited by the second repressor molecule. In some embodiments of these aspects, at least one repressor molecule is a protein. In some embodiments of these aspects, at least one repressor molecule is an RNA molecule.

[0011] In other such embodiments, the first and second repressor sequences encode first and second repressor proteins, such that the first repressible promoter sequence is inhibited by the first repressor protein and the second repressible promoter sequence is inhibited by the second repressor protein. In other embodiments, the first and second repressor sequences encode first and second repressor molecules, such that the first repressible promoter sequence is inhibited by the first repressor molecule and the second repressible promoter sequence is inhibited by the second repressor molecule.

[0012] In some embodiments of these aspects, the genetic toggle switch further comprises at least one ribosome binding sequence. In some embodiments of these aspects, the genetic toggle switch further comprises at least one terminator sequence. In some embodiments of these aspects, the genetic toggle switch further comprises at least one degradation tag sequence. In some embodiments of these aspects, the genetic toggle switch further comprises at least one sequence encoding an iRNA molecule specific for a protein encoded by another module. In some embodiments of these aspects, at least one repressor sequence, R_1 , R_2 , or both, encodes for an engineered zinc-finger protein.

[0013] In some aspects, the logic module comprises a repressible promoter sequence (rP_B) operably linked to a repressor sequence encoding a repressor protein (R_B). In some embodiments of these aspects, the logic module further comprises at least one ribosome binding sequence. In some embodiments of these aspects, the logic module further comprises at least one terminator sequence. In some embodiments of these aspects, the logic module further comprises at least one degradation tag sequence. In some embodiments of these aspects, the logic module further comprises at least one sequence encoding an iRNA molecule specific for a protein encoded by another module. In some embodiments of these aspects, the repressor sequence of the logic module encodes for a protein or an RNA molecule. In some embodiments, the protein is an engineered zinc-finger protein. In other

embodiments, the logic module consists essentially of a repressible promoter sequence (rP_B) operably linked to a repressor sequence encoding a repressor protein (R_B). In other embodiments, the logic module consists of a repressible promoter sequence (rP_B) operably linked to a repressor sequence encoding a repressor protein (R_B).

[0014] In some aspects, each sensor module comprises a repressible promoter sequence (rP_C and rP_D) operably linked to a nucleic acid sequence encoding a sensor molecule (i.e., sensor A and sensor B), i.e., rP_C - sensor A, and rP_D - sensor B. In some embodiments of these aspects, at least one sensor module further comprises at least one ribosome binding sequence. In some embodiments of these aspects, at least one sensor module further comprises at least one terminator sequence. In some embodiments of these aspects, at least one sensor module further comprises at least one degradation tag sequence. In some embodiments of these aspects, at least one sensor module further comprises at least one sequence encoding an iRNA molecule specific for a protein encoded by another module. In other embodiments, each sensor module consists essentially of a repressible promoter sequence (rP_C and rP_D) operably linked to a nucleic acid sequence encoding a sensor (sensor A and sensor B), i.e., rP_C - sensor A and rP_D - sensor B. In other embodiments, each sensor module consists of a repressible promoter sequence (rP_C and rP_D) operably linked to a nucleic acid sequence encoding a sensor (sensor A and sensor B), i.e., rP_C - sensor A and rP_D - sensor B. In some embodiments of these aspects and all such aspects described herein, the at least two sensor modules encode for different sensor molecules.

[0015] In some embodiments of these aspects and all such aspects described herein, the inducible promoter sequence of the input module is induced by a biological agent, a chemical agent, a metal ion, a toxin, or a pollutant.

[0016] In some embodiments of these aspects and all such aspects described herein, the repressor encoded by the first repressor sequence of the toggle switch (R_1) and the repressor encoded by the input module (R_A) are the same repressor.

[0017] In some embodiments of these aspects and all such aspects described herein, the repressors encoded by the logic module and the input module are different repressors.

[0018] In some embodiments of these aspects and all such aspects described herein, the repressors encoded by the logic module and the genetic toggle switch are different repressors.

[0019] In some embodiments of these aspects and all such aspects described herein, the second repressible promoter sequence of the genetic toggle switch (rP_2), the repressible promoter sequence of the logic module (rP_B), and the repressible promoter sequence of one sensor module (rP_D) are repressed by the same repressor.

[0020] In some embodiments of these aspects and all such aspects described herein, the second repressible promoter sequence encoded by the genetic toggle switch (rP_2), the repressible promoter sequence encoded by the logic module (rP_B), and the repressible promoter sequence encoded by one sensor module (rP_D) comprise the same repressible promoter sequence.

[0021] In some embodiments of these aspects and all such aspects described herein, the repressible promoter sequence of one sensor module (rP_C) is repressed by the repressor encoded by the logic module (R_B).

[0022] In some embodiments of these aspects and all such aspects described herein, n is an integer value between and including 1 and 100. In some embodiments of these aspects and all such aspects described herein, n is an integer value between and including 1 and 50. In some embodiments of these aspects and all such aspects described herein, n is an integer value between and including 1 and 20. In some embodiments of these aspects and all such aspects described herein, n is an integer value between and including 1 and 15. In some embodiments of these aspects and all such aspects described herein, n is an integer value selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10.

[0023] In other aspects, described herein are methods of modulating chemotaxis in a biological system using a biological circuit chemotactic converter comprising at least one input module (IM) _{n} , at least one genetic toggle switch (TS) _{n} , at least one logic module (LM) _{n} , and at least two sensor modules (SMA and SMB) _{$n+1$} , where $n \geq 1$.

[0024] In some embodiments of these aspects, the biological system is a cell. In some embodiments of these aspects, the cell is a eukaryotic cell, a prokaryotic cell, or a synthetic cell. In other embodiments, the cell is a bacterial cell.

[0025] In other aspects, the invention provides a system that comprises one or more biological circuit chemotactic converters for converting a chemoattractant signal into a chemotactic response. In some embodiments of these aspects and all such aspects described herein, the system further comprises a cell, such as a eukaryotic cell, a prokaryotic cell, or a synthetic cell.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] **Figures 1A-1C** depicts exemplary modules for autonomous chemotactic converters. **Figure 1A** shows a chemotactic environment made up of three chemoattractant gradients (A, B, C). **Figure 1B** depicts an exemplary biological circuit chemotactic converter, whereby toggle switches control the sequential expression of three chemotaxis sensor receptors, for autonomously navigating bacteria down three chemoattractant gradients. Inputs into promoters and logic operations are shown explicitly except when the promoter name is italicized, which represents an inducible promoter. **Figure 1C** depicts the Boolean on/off values for the network genes that illustrate the sequential order of operations in the exemplary biological circuit chemotactic converter.

[0027] **Figures 2A-2C** depict an exemplary switch-module (SIMM) for use in the converters described herein and data from the exemplary switch module (SIMM). **Figure 2A** depicts an exemplary SIMM comprising a first inducible promoter sequence (*P inducible,1*) linked to a forward recombinase recognition site (R1RS), an inverted sequence of a second inducible promoter (*P inducible,2*), a recombinase sequence (Recombinase 1), and a reverse recombinase recognition site

(R1RS). Further upstream of the first inducible promoter sequence of the SIMM is an inverted sequence of a constitutively active inducer agent that is specific for the second inducible promoter. Downstream of the reverse recombinase recognition site of the SIMM is an inverted sequence of a third inducible promoter (*P inducible,3*) and a sequence encoding a dominant negative version of the inducer agent specific for the second inducible promoter. A further component of the exemplary circuit is a sequence of the second inducible promoter operably linked to an output product. In the absence of any inducing agent, constitutive expression of the inducer agent that is specific for the second inducible promoter drives expression from the inducible promoter operably linked to an output product. Addition of the inducing agent specific for the first inducible promoter sequence (*P inducible,1*) of the SIMM causes expression of the recombinase sequence and subsequent inversion of the recombinase sequence and the inverted sequence of the second inducible promoter (*P inducible,2*) of the SIMM, such that the second inducible promoter can now drive expression of the dominant negative inducer agent that is specific for the second inducible promoter sequence. Expression of the dominant negative inducer agent inhibits transcription of the output product. Addition of the inducing agent specific for the third inducible promoter (*P inducible,3*) causes reexpression of the inverted recombinase sequence and subsequent restoration to the initial state, and consequent expression of the output product. **Figure 2B** depicts an exemplary SIMM, as shown in Figure 2A, where the first inducible promoter sequence (*P inducible,1*) is P_{LtetO} , the forward recombinase recognition site (R1RS) is loxP, the second inducible promoter (*P inducible,2*) is P_{BAD} , the recombinase sequence encodes Cre recombinase, the reverse recombinase recognition site (R1RS) is LoxP, the constitutively active inducer agent is arabinose C, the third inducible promoter (*P inducible,3*) is P_{LlacO} , the dominant negative version of the inducer agent is a dominant negative arabinose C, and the output product is GFP. **Figure 2C** shows data obtained from the exemplary circuit of Figure 2B, where in the initial state, in the absence of any inducing agent, GFP is expressed. Addition of anhydrous tetracycline to the circuit inhibits GFP expression by permitting expression of the dominant negative arabinose C via Cre recombinase activity. Addition of IPTG then restores the circuit to its initial state whereby GFP is expressed.

[0028] Figures 3A-3B depict representative biological converter switches. **Figure 3A** shows an exemplary analog-to-digital biological converter switch that enables the discretization of analog inputs. The analog-to-digital biological converter switch is composed of a bank of modules, i.e., genetic toggle switches that have increasing response thresholds so that sequential toggling is achieved as input levels increase. The design shown enables different natural or synthetic pathways to be activated depending on distinct input ranges, which is useful in cell-based biosensing applications. Inputs into promoters and logic operations are shown explicitly except when the promoter name is italicized, which represents an inducible promoter. **Figure 3B** depicts an exemplary digital-to-analog biological converter switch that enables the programming of defined promoter activity based on combinatorial inputs. The digital-to-analog biological converter switch comprises a bank of

recombinase-based switches, known as single-invertase memory modules (SIMMs) (A.E. Friedland et al., *Science* (2009) 324:1199-1202). Each SIMM comprises an inverted promoter and a recombinase gene located between its cognate recognition site, as indicated by the arrows. Upon the combinatorial addition of inducers that activate specific P promoters, different SIMMs are flipped, enabling promoters of varying strength to drive GFP expression. This allows combinatorial programming of different levels of promoter activity.

[0029] **Figure 4** depicts a tunable genetic filter. Filter characteristics can be adjusted by tuning the degradation of RNA and protein effectors in negative feedback loops. Examples of RNA effectors include small interfering RNAs, riboregulators, and ribozymes. Examples of protein effectors include transcriptional activators and repressors.

[0030] **Figure 5A-5B** show exemplary adaptive learning networks. **Figure 5A** depicts an associative memory circuit that enables association between two simultaneous inputs (Activator A and Activator B), so that the subsequent presence of only a single input can drive its own pathway and the pathway of the other input. Associations between inputs are recorded by a promoter P_{AND} that is activated in the presence of Activator A and Activator B to toggle the memory switch. Inputs into promoters and logic operations are shown explicitly except when the promoter name is italicized, which represents an inducible promoter. **Figure 5B** depicts a winner-take-all circuit, which allows only one input out of many inputs to be recorded. This effect is achieved by a global repressor protein that gates all inputs and prevents them from being recorded if there has already been an input recorded in memory.

[0031] **Figures 6A-6C** shows an amyloid-based memory. **Figure 6A** demonstrates that amyloid-based memory can be implemented by fusing a prion-determining region (PD) to an effector gene, such as a transcriptional activator. **Figure 6B** shows that overexpressing the prion-determining region via promoter P_{OFF} causes aggregation of the fusion protein, rendering the effector inactive. **Figure 6C** shows that subsequent overexpression of chaperone proteins (HSP104), which act to disaggregate amyloids, via promoter P_{ON} releases the effector from the amyloid state and enables it to fulfill its function. Inputs into promoters and logic operations are shown explicitly except when the promoter name is italicized, which represents an inducible promoter.

[0032] **Figure 7** depicts a cell-cycle counter for biological containment. Cell-cycle counting is accomplished with a cascade of single recombinase-based memory units (e.g., SIMMs), each of which is driven by a cell-cycle-dependent promoter. After N cell-cycle events are counted, the gene circuit unlocks the expression of a toxic protein triggering cell death.

[0033] **Figure 8** depicts an exemplary SIMM, as depicted in Figures 2A-2C, that further comprises an antisense RNA (asRNA) to enhance the stability of the system. The asRNA is downstream of and in the inverted orientation with respect to the recombinase gene. The asRNA is targeted against the mRNA of the recombinase gene. Thus, when the upstream promoter is expressed, the recombinase is produced and flips the SIMM. Then, further expression from the upstream

promoter will produce asRNA targeted against recombinase mRNA and degrade it, thus preventing further recombinase from being produced and preventing the SIMM from being flipped back inadvertently.

DETAILED DESCRIPTION

[0034] Described herein are novel biological circuit chemotactic converters that provide modular designs for the conversion of changes in external signals, such as concentrations of chemoattractants, to biological responses, such as expression of sensors necessary for chemotactic responses in biological systems, such as cells. Flexibility in the biological circuit chemotactic converters described herein is provided by incorporating and utilizing various combinations of individual modular components. The biological circuit chemotactic converters are easily extendable to detect large ranges and diversity of input signals, as well as to express a variety of sensors that can translate the input signals into coordinated movements in response to chemoattractants. In other embodiments, the biological circuit chemotactic converters can be used in quorum sensing.

[0035] “Chemotaxis” refers to a biological phenomenon in which bodily cells, bacteria, and other single-cell or multicellular organisms direct their movements according to certain chemicals, termed herein as “chemoattractants” in their environment. For bacteria, such chemotactic responses are important for finding food, for example, glucose, such that the bacteria swim towards the highest concentration of food molecules. Bacteria also use such chemotactic responses to flee from poisons, such as phenol. In multicellular organisms, chemotaxis is critical to both early and subsequent phases of development, for example, movement of sperm towards the egg during fertilization; and migration of neurons or lymphocytes, respectively, as well as in normal function. In addition, it has been recognized that mechanisms that allow chemotaxis in animals can be subverted during cancer metastasis. Chemotaxis or a chemotactic response is called positive if the movement is in the direction of a higher concentration of the chemical in question, and negative if the direction is opposite. Accordingly, in some embodiments of the aspects described herein, biological circuit chemotactic converters can be designed to undergo positive chemotaxis, and in other embodiments, biological circuit chemotactic converters can be designed to undergo negative chemotaxis.

[0036] The ability of some cells to move or migrate in response chemoattractants is a useful functionality to incorporate into biological systems, such as cells, in order to provide directed and specific movements. Accordingly, provided herein are novel “biological circuit chemotactic converters,” which are engineered biological circuits comprised of modular components, that comprise an input module that can detect chemoattractants as inputs, and convert these input information into the upregulation of appropriate sensor molecules, via a series of engineered modular components, including genetic toggle switches, logic modules, and sensor modules. The upregulation of an appropriate sensor allows the biological circuit chemotactic converters to direct the chemotactic movement of a biological system, such as a cell.

[0037] In some embodiments of the aspects described herein, the biological circuit chemotactic converters can be designed to incorporate quorum sensing responses. As defined herein, “quorum sensing” refers to a type of decision-making process used by decentralized groups to coordinate behavior. Many species of bacteria use quorum sensing to coordinate their gene expression according to the local density of their population. Similarly, some social insects use quorum sensing to make collective decisions about where to nest.

[0038] Some of the best-known examples of quorum sensing come from studies of bacteria. Bacteria use quorum sensing to coordinate certain behaviors based on the local density of the bacterial population. Quorum sensing can occur within a single bacterial species as well as between diverse species, and can regulate a host of different processes, essentially serving as a simple communication network. A variety of different molecules can be used as signals, defined herein as “quorum sensing signals.” Common classes of signaling molecules are oligopeptides in Gram-positive bacteria, N-Acyl Homoserine Lactones (AHL) in Gram-negative bacteria and a family of autoinducers known as AI-2 in both Gram-negative and Gram-positive bacteria. In some embodiments of the aspects described herein, the biological circuit chemotactic converters use quorum sensing signals as input signals.

[0039] Bacteria that use quorum sensing constantly produce and secrete certain signaling molecules (called autoinducers or pheromones). These bacteria also have a receptor that can specifically detect the signaling molecule (inducer). When the inducer binds the receptor, it activates transcription of certain genes, including those for inducer synthesis. As in natural systems, there is a low likelihood of a bacterium detecting its own secreted inducer, in order for gene transcription to be activated, the cell must encounter signaling molecules secreted by other cells in its environment. When only a few other bacteria of the same kind are in the vicinity, diffusion reduces the concentration of the inducer in the surrounding medium to almost zero, so the bacteria produce little inducer. However, as the population grows the concentration of the inducer passes a threshold, causing more inducer to be synthesized. This forms a positive feedback loop, and the receptor becomes fully activated. Activation of the receptor induces the up regulation of other specific genes, causing all of the cells to begin transcription at approximately the same time. This coordinated behavior of bacterial cells can be useful in a variety of situations. Quorum sensing was first observed in *Vibrio fischeri*, a bioluminescent bacterium that lives as a mutualistic symbiont in the photophore (or light-producing organ) of the Hawaiian bobtail squid. When *V. fischeri* cells are free-living (or planktonic), the autoinducer is at low concentration and thus cells do not luminesce. However, when they are highly concentrated in the photophore (about 10^{11} cells/ml) transcription of luciferase is induced, leading to bioluminescence. The bioluminescent luciferase produced by *V. fischeri* would not be visible if it was produced by a single cell. By using quorum sensing to limit the production of luciferase to situations when cell populations are large, *V. fischeri* cells are able to avoid wasting energy on the production of useless product. Accordingly, in some embodiments of the aspects

described herein, one or more of the sensor genes of a biological circuit chemotactic converter encodes for a receptor that can detect an autoinducer.

[0040] In the Gram-negative bacteria *Escherichia coli*, cell division may be partially regulated by AI-2-mediated quorum sensing. This species uses AI-2, which is produced and processed by the *lsr* operon. Part of it encodes an ABC transporter which imports AI-2 into the cells during the early stationary (latent) phase of growth. AI-2 is then phosphorylated by the LsrK kinase, and the newly produced phospho-AI-2 can either be internalized or used to suppress LsrR, a repressor of the *lsr* operon (thereby activating the operon). Transcription of the *lsr* operon is also thought to be inhibited by dihydroxyacetone phosphate (DHAP) through its competitive binding to LsrR. Glyceraldehyde 3-phosphate has also been shown to inhibit the *lsr* operon through cAMP-CAPK-mediated inhibition. When grown normally, AI-2 presence is transient.

[0041] *E.coli* and *Salmonella enterica* do not produce AHL signals commonly found in other gram negative bacteria. However, they have a receptor that detects AHLs from other bacteria and change their gene expression in accordance with the presence of other 'quorate' populations of gram-negative bacteria.

[0042] The opportunistic bacteria *Pseudomonas aeruginosa* use quorum sensing to coordinate the formation of biofilms, swarming motility, exopolysaccharide production, and cell aggregation. These bacteria can grow within a host without harming it, until they reach a certain concentration. Then they become aggressive, their numbers sufficient to overcome the host's immune system and form a biofilm, leading to disease. In this species, AI-2 was found to increase expression of *sdiA*, a transcriptional regulator of promoters which promote *ftsQ*, part of the *ftsQAZ* operon essential for cell division. Another form of gene regulation which allows the bacteria to rapidly adapt to surrounding changes is through environmental signaling. Recent studies have discovered that anaerobiosis can significantly impact the major regulatory circuit of quorum sensing. This important link between quorum sensing and anaerobiosis has a significant impact on production of virulence factors of this organism. Garlic reportedly blocks quorum sensing in *Pseudomonas aeruginosa*. As defined herein, disrupting the quorum sensing signalling process in this way is called "quorum quenching."

Genetic Toggle Switches

[0043] Provided herein are biological modules, such as genetic toggle switches, comprising different nucleic acid and protein components, such as promoters, transcriptional activators, transcriptional repressors, recombinases, and output products, to be used in the biological circuit chemotactic converters described herein. The ability to manipulate and combine different components and modules provides flexibility in input and output responses of the biological circuit chemotactic converters described herein.

[0044] In some aspects, genetic toggle switches are provided for use as a biological module in the biological circuit chemotactic converters described herein. A "genetic toggle switch," as

defined herein, refers to a synthetic, addressable cellular memory unit or module that can be constructed from any two repressible promoters arranged in a mutually inhibitory network. A genetic toggle switch exhibits robust bistable behaviour. By “robust bistable behaviour” is meant that the toggle switch exhibits bistability over a wide range of parameter values and that the two states are tolerant of fluctuations inherent in gene expression, *i.e.*, a genetic toggle switch does not flip randomly between states. Bistability of a genetic toggle switch is possible with any set of promoters and repressors as long as a minimum set of conditions are fulfilled, as described, for example, in T.S. Gardner et al., *Nature* (2000) 403: 339-342.

[0045] Bistability of a genetic toggle switch, as described herein, arises from a mutually inhibitory arrangement of at least two repressor sequences. The product of each repressor sequence, *i.e.*, the repressor, can inhibit, at a transcriptional level, a translational level, or a combination thereof, the expression of a product encoded by the other repressor sequence. Thus, in the absence of an appropriate input or inducing agent, such as a transcriptional activating agent, two stable states are possible: a first state in which a first repressor is expressed and inhibits expression of a second repressor sequence, and a second state in which the second repressor is expressed and inhibits expression of the first repressor sequence. For example, in some aspects, repressors act at the transcriptional level, whereby a first promoter sequence drives expression of a first repressor sequence that encodes for a repressor specific for a second promoter sequence. The second promoter sequence, in turn, drives expression of a second repressor sequence that encodes for a repressor specific for a second promoter sequence. In such an aspect, switching between the two states (*i.e.*, expression of the first or second repressor) is mediated by the presence of an exogenous or endogenous input agent, such as an agent that prevents repressor binding to the currently inactive promoter. In such an embodiment, the agent permits the opposing repressor to be maximally transcribed until it stably represses the originally active promoter. In other embodiments of the aspects described herein, repressors in a genetic toggle switch can act at the translational level, whereby a first repressor encodes a product, such as an inhibitory RNA molecule, that inhibits or prevents translation of the second repressor, or causes degradation of the second repressor mRNA. In other embodiments of the aspects described herein, different repressors in a genetic toggle switch can use different mechanisms of repression, *i.e.*, transcriptional, translational, or combinations thereof.

[0046] In one embodiment of this aspect and all such aspects described herein, a genetic toggle switch comprises two different repressible promoter sequences driving expression of two sequences encoding different repressors, such that each promoter can be inhibited by the repressor transcribed by the other promoter. In such an embodiment, the genetic toggle switch comprises a first repressible promoter sequence (rP_1) that drives the transcription of a second repressor sequence (R_2), which encodes a repressor specific for the second repressible promoter sequence, and a second repressible promoter sequence (rP_2) that drives the transcription of a first repressor sequence (R_1), which encodes a repressor specific for the first repressible promoter sequence.

[0047] In some embodiments, the genetic toggle switches are implemented on plasmids, such as plasmids derived from *E. coli*. In some embodiments, the nucleic acid sequences of the promoters and repressors of the genetic toggle switch are contained or present on a single plasmid. In other embodiments, the nucleic acid sequences of the promoters and repressors of the genetic toggle switch are contained or present on multiple plasmids.

[0048] In one embodiment of this aspect and all such aspects described herein, the genetic toggle switch comprises a P_{trc-2} promoter that drives the expression of a temperature-sensitive λ repressor (cIts), and a P_{Ls1con} promoter that drives the expression of a Lac repressor. In such an embodiment, the genetic toggle is switched between states by pulses of isopropyl-b-D-thiogalactopyranoside (IPTG) and thermal pulses. For example, a pulse of IPTG permits expression of cIts driven by the P_{trc-2} promoter, as the IPTG prevents the Lac repressor from binding to the P_{trc-2} promoter. Expression of cIts maintains the state of transcription from the P_{trc-2} promoter by binding and repressing the P_{Ls1con} promoter, thus preventing Lac repressor expression and inhibition of the P_{trc-2} promoter. In contrast, a thermal pulse inhibits the cIts repressor, thus preventing cIts binding to the P_{Ls1con} promoter, and permitting expression of the Lac repressor. Expression of the Lac repressor further maintains the state of transcription from the P_{Ls1con} promoter by binding to and repressing the P_{trc-2} promoter, thus preventing cIts repressor expression and inhibition of the P_{Ls1con} promoter.

[0049] In another embodiment of this aspect and all such aspects described herein, the genetic toggle switch comprises a P_{trc-2} promoter that drives the expression of a Tet repressor (Tet), and a P_{LtetO-1} promoter that drives the expression of a Lac repressor. In such an embodiment, the genetic toggle switch is switched between states by a pulse of IPTG or a pulse of anhydrotetracycline (aTc). For example, a pulse of IPTG permits expression of Tet driven by the P_{trc-2} promoter, as the IPTG will prevent the Lac repressor from binding to the P_{trc-2} promoter. Expression of Tet maintains the state of transcription from the P_{trc-2} promoter by binding and repressing the P_{LtetO-1} promoter, thus preventing Lac repressor expression and inhibition of the P_{trc-2} promoter. In contrast, a pulse of anhydrotetracycline inhibits the Tet repressor, thus preventing Tet binding to the P_{LtetO-1} promoter, and permitting expression of the Lac repressor. Expression of the Lac repressor further maintains the state of transcription from the P_{LtetO-1} promoter by binding to and repressing the P_{trc-2} promoter, thus preventing Tet repressor expression and inhibition of the P_{LtetO-1} promoter.

Libraries of Toggle Switches

[0050] For use in the genetic toggle switches described herein, it is possible to use any set of promoters and repressors as long as they fulfill a minimum set of conditions, as described, for example, in T.S. Gardner *et al.*, Nature (2000) 403: 339-342. In some embodiments of the invention, the promoters useful in the genetic toggle switches are presented under the section entitled *Promoters* and provided in SEQ ID NOs: 1-7, SEQ ID NOs: 41-843, and SEQ ID NOs: 1005-1010.

[0051] In order to further enhance and expand the range and sensitivity of genetic toggle switches for use in the biological circuit chemotactic converters described herein, it is useful to create

libraries of genetic toggle switches with multiple interoperable repressors, such as transcriptional repressors. Thus, in some embodiments of the aspects described herein, a library of transcriptional repressors and activators can be targeted towards unique promoters with minimum crossover, using engineered zinc-finger proteins fused to transcriptional activation and repression domains.

[0052] To create such libraries, unique promoters containing sequence sites known to bind to engineered zinc-finger proteins can be synthesized. These sites are made up of three sequences, each of which is at least 3 DNA base pairs long. Each 3 base pair sequence binds to a single zinc-finger domain. Thus, in some embodiments, each complete engineered zinc-finger transcription factor contains three zinc-finger domains to target a total 9 base pair region of DNA. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 1. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 2. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 3. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 4. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 5. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 6. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 7. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 8. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 9. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 10. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 11. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 12. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 13. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 14. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 15. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 15. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 17. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 18. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 19. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is 20. In some embodiments, the number of zinc-finger domains used in a complete engineered zinc-finger transcription factor is at least 25, at least 50, at least 100, or more.

[0053] Representative examples of zinc-finger pools created for the shaded 3 base pair sequences are shown below (ML Maeder et al., *Molecular Cell* 2008: 31, 294–301):

F1				F2				F3			
GAA	GCA	GGA	GTA	GAA	GCA	GGA	GTA	GAA	GCA	GGA	GTA
GAC	GCC	GGC	GTC	GAC	GCC	GGC	GTC	GAC	GCC	GGC	GTC
GAG	GCG	GGG	GTG	GAG	GCG	GGG	GTG	GAG	GCG	GGG	GTG
GAT	GCT	GGT	GTT	GAT	GCT	GGT	GTT	GAT	GCT	GGT	GTT
TAA	TCA	TGA	TTA	TAA	TCA	TGA	TTA	TAA	TCA	TGA	TTA
TAC	TCC	TGC	TTG	TAC	TCC	TGC	TTG	TAC	TCC	TGC	TTG
TAG	TCG	TGG	TTG	TAG	TCG	TGG	TTG	TAG	TCG	TGG	TTG
TAT	TCG	TGT	TTT	TAT	TCG	TGT	TTT	TAT	TCG	TGT	TTT

Using such pools, complete engineered zinc-finger proteins containing at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, or more zinc-fingers that can target synthetic promoters can be selected.

[0054] In some embodiments, the engineered zinc-finger proteins are fused to transcriptional activation domains, for *e.g.*, VP16, VP64, p65, Gal4, α -subunit of RNA polymerase, Wild-type CRP (amino acid residues 1–209), CRP D1(residues 1–180), CRP D2 (residues 137–190), CRP D3 (residues 137–180) and CRP D4 (residues 151–168). In other embodiments, the engineered zinc-finger proteins are fused to transcriptional repressor domains *e.g.*, SKD, KRAB (Margolin et al., 1994), SNAG, Kid, Ume6, CRP, SID (Ayer et al., 1996). Thus, an engineered zinc-finger protein can be used a transcriptional activator or transcriptional repressor, depending on the requirements of the various embodiments described herein, by fusing an engineered zinc-finger protein with an appropriate transcriptional activator or transcriptional repressor domain. Non-limiting examples of methods of engineering zinc-finger proteins and transcriptional activation domains for fusion are discussed, for example, at Kwang-Hee B. et al, *Nature Biotechnology* 2003: 21, p. 275-280; R-J Kwon et al., *Biotechnology Letters* (2006) 28: 9–15; P. Blancafort et al., *PNAS*, 2005, 102: 33, p.11716–11721; J. T. Stege et al., *The Plant Journal* (2002) 32, 1077–1086; J.Y. Lee et al., *Nucleic Acids Research*, 2008, 36:16; K-S Park et al., *Nature Biotechnology*, 2003, 21:10, p.1208-1214; R.R. Beerli et al., *PNAS*, 2000, 97:4, p. 1495-1500; P. Blancafort et al., *Nature Biotechnology* 2003: 21, p. 269-274; D-k Lee, et al., *Genome Res.*, 2003, 13: 2708-2716. Interoperability of such fusion engineered zinc-finger proteins can be assessed by combinatorial addition of the different engineered zinc-finger transcription factors to determine how promoter activity is affected.

[0055] To enhance cooperativity of engineered zinc-finger-based transcription factors, in some embodiments, engineered zinc-finger-based transcription factors can be further engineered to dimerize, using dimerization domains such as leucine zipper domains. In some embodiments, the affinity of monomeric engineered-zinc finger proteins can be increased or decreased by site-directed mutagenesis of amino acids known to contact the DNA backbone and/or bases. Non-limiting examples of methods to achieve such affinity modification are discussed, for example, at J. L. Pomerantz, et al., *Biochemistry*, 1998, 37: 4, p. 965-970, and S.A. Wolfe et al., *Structure*, 2000, 8:7, p. 739-750.

[0056] Pairwise combinations of the engineered zinc-finger-based transcriptional repressors can be conducted to identify mutually-repressing transcription factors and test for bistability, for use in the genetic toggle switches and other modules described herein. In some embodiments, the ability to flip genetic toggle switches can be assessed by overexpressing transcriptional repressors one by one. Thresholds for switching between repressors in such genetic toggle switches can be modulated by changing the promoters in the toggle switch to affect, for example, binding efficiency and repression efficiency.

[0057] In some embodiments, the rates of protein synthesis of transcriptional repressors for use in the genetic toggle switches described herein can be modified by adding or modifying sequences for a ribosome binding site (RBS). In some embodiments, an RBS is placed downstream of a promoter sequence and upstream of a sequence encoding a transcriptional repressor being transcribed from that promoter.

[0058] In some embodiments, a genetic toggle switch can further comprise an output product sequence (OP) that encodes an output product, such as a protein or an RNA molecule, expression of which reflects or is indicative of the state of the genetic toggle switch. In such embodiments, the genetic toggle switch comprises a first repressible promoter sequence (rP_1) that drives the transcription of a second repressor sequence (R_2) that encodes a repressor specific for the second repressible promoter sequence, and a second repressible promoter sequence (rP_2) that drives the transcription of a first repressor sequence (R_1), which encodes a repressor specific for the first repressible promoter sequence, as well as the transcription of an output sequence (OP), *i.e.*, the genetic toggle switch comprises rP_1 - R_2 and rP_2 - R_1 -OP. In such embodiments, when the second repressible promoter is active and transcribing the first repressor and the output sequence, the toggle switch is considered in the “on” state. In such embodiments, the expression of the output product can be thought of as a digital output 00000001, in the binary system. In some embodiments of various aspects described herein, multiple genetic toggle switches are combined, each having a different output product that represents the state of that particular genetic toggle switch, such that the digital output of that combination of toggle switches is dependent on how many of the genetic toggle switches are in the “ON” state.

[0059] As used herein, a “digital output” refers to an output that can be represented in a binary format. The binary numeral system, or base-2 number system represents numeric values using two symbols, 0 and 1. More specifically, the usual base-2 system is a positional notation with a radix of 2. Owing to its implementation in digital electronic circuitry using logic gates, the binary system is used internally by all modern computers. A “bit,” as defined herein, is a binary digit. The numeric value represented by a combination of modules of the invention, for example, genetic toggle switches, is dependent upon the value assigned to each module when it is in the “on” state. For example, in those embodiments described herein where a combination of three genetic toggle switches is used, the digital output is represented as 00000000, when no genetic toggle switch is “on”. When 1 genetic

toggle switch is “on,” *i.e.*, transcribing the output product, the state is 00000001. When 2 genetic toggle switches are “ON,” the state is 00000010. When all 3 genetic toggle switches are “on,” the state is 00000011. A “byte” represents a collection of eight bits. A byte can hence be defined as a collection of 8 bits, such that 256 values or states can be represented, ranging from 0 to 255, *i.e.*, 00000000 to 11111111.

[0060] Any such output product as described herein can be utilized. In some embodiments, where the expression of an output product is driven by only one of the two promoters in a toggle switch, the output product can encode a reporter protein or a reporter RNA molecule. In some embodiments, the reporter protein is a fluorescent reporter protein, *e.g.*, green fluorescent protein. In some embodiments, where multiple genetic toggle switches are combined, each output sequence encodes for a different output product. For example, when a combination of three genetic toggle switches are combined, the output product sequences can encode for green fluorescent protein, yellow fluorescent protein, and red fluorescent protein, such that expression of all three fluorescent proteins represents a digital output of 00000011 for that combination of switches. Detection of the output products of the modules described herein can be performed using any method known to one of skill in the art, including, but not limited to, fluorescent detectors, such as microscopes and flow cytometers, luminescent detectors, quantitative PCR, Western blot analysis, etc., based on the nature of the output product being detected.

[0061] In other embodiments, each promoter in a genetic toggle switch can drive expression of an output product, such that the expression of one output product represents one digital output and expression of an output product driven by the opposing promoter represents another digital output, *i.e.*, the “on” and “off” states of a single genetic toggle switch are represented by the expression of a different output product, which can be assigned in an arbitrary manner by a skilled artisan or user, or based on the design of the circuit in which the genetic toggle switch is a component. In such embodiments, the designation of which output product expression corresponds to which state, *i.e.*, “on” or “off,” can be determined by the skilled artisan.

Single Invertase Memory Modules

[0062] Provided herein are biological modules such as single invertase memory modules, comprising different nucleic acid and protein components, such as promoters, transcriptional activators, transcriptional repressors, recombinases, and output products, to be used in the biological circuit chemotactic converters described herein. The ability to manipulate and combine different components and modules provides flexibility in input and output responses of the biological circuit chemotactic converters described herein.

[0063] In some aspects, a “single invertase memory module” is provided as a biological module for use in the biological circuit chemotactic converters described herein. A “single invertase memory module (SIMM),” as defined herein, refers to a stable, switchable bit of memory that uses recombinases, such as Cre and *flp_e*, which can invert DNA between two oppositely oriented cognate

recombinase recognition sites. A unique feature and advantage of SIMMs, of relevance to their use in the biological circuit chemotactic converters described herein, is the ability to design such SIMMs to lack both “leakiness” and mixtures of inverted and non-inverted states that can be caused by expressing recombinases independently from their cognate recognition sites. Thus, the use of SIMMs in the biological circuit chemotactic converters described herein allows for the maintenance of memory, and provides the ability to control and maintain discrete states by expressing recombinases between their cognate recognition sites.

[0064] At a minimum, a SIMM is a nucleic acid-based module comprising a recombinase sequence located between its cognate recombinase recognition sites, *i.e.*, $RRS_{for} - RC - RRS_{rev}$, where RRS_{for} is a forward recombinase recognition site; RC is a recombinase sequence encoding a recombinase that recognizes RRS_{for} and RRS_{rev} ; and RRS_{rev} is a reverse recombinase recognition site. Upon recombinase expression following activation of an upstream promoter, the recombinase causes a single inversion of the nucleic acid between the cognate recognition sites, *i.e.*, the recombinase nucleic acid sequence or $RRS_{for} - RC_{inv} - RRS_{rev}$. Any further transcription from the upstream promoter yields antisense RNA of the recombinase gene rather than sense RNA, and therefore no further recombinase protein is produced. Thus, the inversion event is discrete and stable, and does not result in a mixture of inverted and non-inverted states. The upstream promoter driving expression of the SIMM can be a promoter sequence within an upstream SIMM, another modular component, a component of the same SIMM, or be an isolated promoter sequence.

[0065] In some aspects, the SIMM further comprises an upstream promoter sequence, *i.e.*, $P - RRS_{for} - RC - RRS_{rev}$, where P is a promoter sequence. In other aspects, a SIMM comprises the recombinase sequence located between its cognate recombinase recognition sites, and further comprises an inverted inducible promoter sequence upstream of the recombinase sequence, *i.e.*, $RRS_{for} - iP_{inv} - RC - RRS_{rev}$, where RRS_{for} is a forward recombinase recognition site, iP_{inv} is an inverted promoter sequence, RC is a recombinase sequence, and RRS_{rev} is a reverse recombinase recognition site. Upon recombinase expression following activation of an upstream promoter, the recombinase causes a single inversion of the DNA between the cognate recognition sites, including the nucleic acid sequence encoding itself, *i.e.*, the recombinase nucleic acid sequence. Any further transcription from the upstream promoter yields antisense RNA of the recombinase gene rather than sense RNA, and therefore no further recombinase protein is produced. Further, the inverted promoter is now in the proper orientation to drive transcription of components of any downstream modules, for example, another SIMM. In some embodiments of the aspects described herein, the promoter is a constitutive promoter. In other embodiments of the aspects described herein, the promoter is a inducible promoter. In some embodiments, the inducible promoter is a repressible promoter. In some embodiments, the inducible promoter is activated by an activating agent.

[0066] In some embodiments of the aspects described herein, a SIMM can use any recombinase for encoding memory, rather than only unidirectional recombinases. In some

embodiments, the recombinase is encoded between its cognate recombinase recognition sequences. In other embodiments, the recombinase is encoded outside of its cognate recombinase recognition sequences. In those embodiments where the recombinase is encoded outside of its cognate recombinase recognition sequences, the SIMM can be used as, for example, a waveform generator, such that the input or inputs that lead to recombinase expression results in constant inversion between the recombinase recognition sequences and is used to generate pulses of outputs. Such outputs can be any of the output products described herein. In some embodiments, the output is a fluorescent protein.

[0067] The recombinases and recombination recognition sequences for use in the SIMMs described herein can be selected from any known or variant, *i.e.*, engineered, recombinase or recombinase recognition sequences, as determined by a skilled artisan. In some embodiments of the various aspects described herein, the recombinase is a Cre recombinase and the recombinase recognition sites are LoxP sites or variants thereof. Alternative site-specific recombinases include, but are not limited to, 1) the Flp recombinase of the 2pi plasmid of *Saccharomyces cerevisiae* (Cox (1983) Proc. Natl. Acad. Sci. USA 80:4223) which recognize FRT sites and variants thereof; 2) the integrase of *Streptomyces phage* .PHI.C31 that carries out efficient recombination between the *attP* site of the phage genome and the *attB* site of the host chromosome (Groth et al., 2000 Proc. Natl. Acad. Sci. USA, 97: 5995); 3) the Int recombinase of bacteriophage lambda (lambda-int/*attP*) (with or without Xis) which recognizes *att* sites (Weisberg et al. In: Lambda II, supra, pp. 211-250); 4) the xerC and xerD recombinases of *E. coli* which together form a recombinase that recognizes the 28 bp dif site (Leslie and Sherratt (1995) EMBO J. 14:1561); 5) the Int protein from the conjugative transposon Tn916 (Lu and Churchward (1994) EMBO J. 13:1541); 6) TpnI and the β -lactamase transposons (Levesque (1990) J. Bacteriol. 172:3745); 7) the Tn3 resolvase (Flanagan et al. (1989) J. Mol. Biol. 206:295 and Stark et al. (1989) Cell 58:779); 8) the SpoIVC recombinase of *Bacillus subtilis* (Sato et al. J. Bacteriol. 172:1092); 9) the Hin recombinase (Galsgow et al. (1989) J. Biol. Chem. 264:10072); 10) the Cin recombinase (Hafter et al. (1988) EMBO J. 7:3991); 11) the immunoglobulin recombinases (Malynn et al. Cell (1988) 54:453); and 12) the FIMB and FIME recombinases (Blomfield et al., 1997 Mol. Microbiol. 23:705). Additional non-limiting examples of recombinases and their cognate recombinase recognition sequences that are useful for the SIMMs and biological circuit chemotactic converters described herein are provided in SEQ ID NOs: 8-25, SEQ ID NOs: 844-857, and SEQ ID NOs: 1002-1004.

[0068] The inverted promoter sequence in a SIMM can be used to drive transcription of downstream components of that SIMM or other biological modules upon recombinase activation and inversion of the promoter to the forward direction. Accordingly, an inverted promoter sequence for use in the SIMMs described herein can be a constitutive or inducible promoter, depending upon the requirements of the biological circuit chemotactic converters. Non-limiting examples of such promoter sequences for use in the SIMMs described herein are provided in SEQ ID NOs: 1-7, SEQ ID NOs: 41-843, and SEQ ID NOs: 1005 -1010.

[0069] A SIMM can further comprise one or more components, including, but not limited to, ribosome binding sequences, degradation tag sequences, translational terminator sequences, and anti-sense sequences, that are added to, for example, enhance translation of mRNA sequences for protein synthesis, prevent further transcription downstream of the recombinase, or enhance degradation of the recombinase mRNA sequence or protein sequence once the recombinase sequence has been expressed. Such additional 'parts' or components, by enhancing the fidelity and accuracy of the biological modules, such as SIMMs, permit, for example, increased numbers and combinations of biological modules and improve the capabilities of the biological circuit chemotactic converters described herein.

[0070] In other embodiments of the aspects described herein, a SIMM can further comprise one or more ribosome binding site sequences (RBSs) to promote efficient and accurate translation of the mRNA sequences for protein synthesis. RBSs are useful components for modulating the efficiency and rates of synthesis of the proteins or other outputs encoded by the biological converter switches described herein. Non-limiting examples of such RBS sequences for use in the SIMMs described herein are provided in SEQ ID NOs: 26-33 and SEQ ID NOs: 858-1001. Accordingly, in some embodiments of these aspects, a SIMM further comprises a ribosome binding site upstream of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - RBS - RC - RRS_{rev}$, where RBS is a ribosome binding site. In other aspects, a SIMM further comprises both an inverted promoter sequence and a ribosome binding site upstream of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - iP_{inv} - RBS - RC - RRS_{rev}$, where iP_{inv} is an inverted promoter sequence and RBS is a ribosome binding site sequence.

[0071] In other embodiments of the aspects described herein, one or more terminator sequences can be added to a SIMM to prevent activation of downstream genes or modules by an upstream promote sequence. Terminator sequences can be added to the end of, for example, the sequence encoding a recombinase in a SIMM, to prevent further transcription downstream of the recombinase. Thus, terminator sequences are useful in the biological converter switches described herein to prevent unwanted transcription driven by activation of the various modules. Non-limiting examples of such terminators for use in the SIMMs described herein are provided in Tables 1-5. Accordingly, in some embodiments of these aspects, a SIMM further comprises a transcriptional terminator sequence downstream of the recombinase sequence, *i.e.*, $RRS_{for} - RC - T - RRS_{rev}$, where T is a terminator sequence. In other embodiments, a SIMM further comprises both an inverted promoter sequence and a terminator sequence upstream and downstream respectively of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - iP_{inv} - RC - T - RRS_{rev}$, where iP_{inv} is an inverted promoter sequence and T is a terminator sequence.

[0072] Degradation tag sequences are also provided for use in the SIMMs and biological converter switches described herein to enhance degradation of a protein expressing the tag. The ability to add degradation tags to the proteins encoded by the SIMMs and biological converter switches

described herein provides an additional layer of regulation and control of the modules. Non-limiting examples of such degradation tag sequences for use in the SIMMs described herein are provided in SEQ ID NOs: 34-40. Accordingly, in some embodiments of the aspects described herein, a SIMM further comprises a protein degradation tag sequence downstream of the recombinase sequence, *i.e.*, the SIMM comprises *i.e.*, the SIMM comprises $RRS_{for} - RC - D - RRS_{rev}$, where D is a degradation tag sequence. In other embodiments, a SIMM further comprises both an inverted promoter sequence and a degradation tag sequence upstream and downstream respectively of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - iP_{inv} - RC - D - RRS_{rev}$, where iP_{inv} is an inverted promoter sequence and D is a degradation tag sequence.

[0073] In further embodiments of these aspects, a SIMM comprises both a ribosome binding site upstream of the recombinase sequence and a protein degradation tag sequence downstream of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - RBS - RC - D - RRS_{rev}$. In some embodiments, a SIMM further comprises an inverted promoter sequence and ribosome binding site upstream of the recombinase sequence, and a protein degradation tag sequence downstream of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - iP_{inv} - RBS - RC - D - RRS_{rev}$. In some embodiments of these aspects, a SIMM comprises both a protein degradation tag sequence and a transcriptional terminator sequence downstream of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - iP_{inv} - RC - D - T - RRS_{rev}$. In some embodiments, a SIMM further comprises an inverted promoter sequence upstream of the recombinase sequence, and a protein degradation tag sequence and a terminator sequence downstream of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - iP_{inv} - RC - D - T - RRS_{rev}$. In some embodiments of these aspects, a SIMM further comprises a ribosome binding site upstream of the recombinase sequence and a terminator sequence downstream of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - iP_{inv} - RBS - RC - T - RRS_{rev}$. In some embodiments, a SIMM further comprises an inverted promoter sequence and ribosome binding site upstream of the recombinase sequence, and a terminator sequence downstream of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - iP_{inv} - RBS - RC - T - RRS_{rev}$.

[0074] In some particular embodiments of these aspects, a SIMM can further comprise a ribosome binding site upstream of the recombinase sequence, and protein degradation tag and transcriptional terminator sequences downstream of the recombinase sequence, *i.e.*, $RRS_{for} - iP_{inv} - RBS - RC - D - T - RRS_{rev}$. In other such embodiments, a SIMM further comprises an inverted promoter sequence and ribosome binding site upstream of the recombinase sequence, and protein degradation tag and transcriptional terminator sequences downstream of the recombinase sequence, *i.e.*, the SIMM comprises $RRS_{for} - iP_{inv} - RBS - RC - D - T - RRS_{rev}$. In such embodiments, the combined addition of an RBS, a transcriptional terminator sequence, and a degradation tag to the SIMM provides an enhanced ability to regulate and control expression of the recombinase encoded by the SIMM.

[0075] In some embodiments of the aspects described herein, a SIMM can further comprise or be designed to include an antisense RNA sequence downstream of and in an inverted orientation in respect to the sequence encoding the recombinase, which is specific for the recombinase mRNA, *i.e.*, $RRS_{for} - RC - asRNA_{inv} - RRS_{rev}$, where $asRNA_{inv}$ is an inverted antisense RNA sequence. In such embodiments, upon expression of the recombinase protein, in response, for example, to activation of an upstream promoter, the recombinase flips the sequences in the SIMM flanked by the recombinase recognition sites, such that the recombinase sequence is in the inverted orientation and the sequence encoding the antisense RNA is in the forward direction, *i.e.*, $RRS_{for} - asRNA - RC_{inv} - RRS_{rev}$. The inversion event prevents further transcription and translation of the recombinase sequence, while transcription of the sequence encoding the antisense RNA specific for the recombinase enhances degradation of any transcribed recombinase mRNA sequence remaining. In further embodiments, the SIMM can further comprise a ribosome binding site upstream of the recombinase sequence, protein degradation tag or transcriptional terminator sequences downstream of the recombinase sequence, or any combination thereof, *i.e.*, $RRS_{for} - RBS - RC - asRNA_{inv} - D - T - RRS_{rev}$, such that expression of the recombinase is regulated by a combination of elements to ensure accuracy and fidelity of the SIMM.

[0076] In some embodiments of these aspects and all such aspects described herein, a SIMM can be designed so that it can be reset by placing an additional promoter sequence in an inverted orientation downstream of the reverse recombinase recognition site, *i.e.*, $RRS_{for} - RBS - RC - RRS_{rev} - iP_{inv}$; where iP_{inv} is an inverted inducible promoter sequence. Upon activation of the promoter, the state of such a SIMM is flipped from its inverted state back to its original state, when the recombinase sequence is in the inverted orientation. In some embodiments, the same reverse inducible promoter can be used throughout an entire set of SIMMs within an a biological converter switch, such that a single inducer can be used to perform a global reset of all the SIMMs in the biological converter switch.

[0077] In some embodiments of these aspects and all such aspects described herein, a SIMM comprises a forward recombinase recognition site sequence, an inverted promoter sequence, a ribosome binding site sequence, a recombinase gene sequence, and a reverse recombinase recognition site sequence, and a sequence encoding an output product, *i.e.*, $RRS_{for} - P_{inv} - RC - RRS_{rev} - OP$; where RRS_{for} is a forward recombinase recognition sequence; P_{inv} is an inverted promoter sequence; RC is a recombinase gene sequence encoding a recombinase that is specific for RRS_{for} and RRS_{rev} ; RRS_{rev} is a reverse recombinase recognition sequence, and OP is an output product sequence. In such embodiments, upon expression of the recombinase, the sequence between the two recombinase recognition sites is inverted, resulting in termination of recombinase expression and allowing for the inverted promoter sequence to be in the appropriate direction to drive expression of the output gene sequence and any downstream modules. In some embodiments, the promoter is an inducible promoter. In some embodiments, the inducible promoter is a repressible promoter, or a promoter that can be

activated by an activating agent. In such embodiments, the SIMM can further comprise degradation tags, ribosome binding sites, transcriptional terminator sequences, and antisense RNA sequences, as described herein, to add further regulatory capacities to the SIMM.

[0078] In other embodiments of these aspects and all such aspects described herein, a SIMM can comprise an inducible promoter sequence (iP), a forward recombinase recognition site sequence (RRS_{for}), an inverted sequence of a constitutive promoter (P_{inv}), a recombinase gene sequence (RC), a reverse recombinase recognition site sequence (RRS_{rev}), and an output product sequence (OP), *i.e.*, iP-RRS_{for} - iP_{inv} - RC - RRS_{rev} - OP. In such embodiments, activation of the inducible promoter drives expression of the recombinase, which inverts the sequence between the two recombinase recognition sites, resulting in termination of recombinase expression, and allowing for the inverted, constitutive promoter sequence to be in the appropriate direction to drive expression of the output product sequence and any downstream modules, for example, one or more additional SIMMs. In some such embodiments, the SIMM can further comprise an additional inducible promoter in inverted orientation between the reverse recombinase recognition site sequence and the output product sequence, *i.e.*, (iP-RRS_{for} - iP_{inv} - RC - RRS_{rev} - iP_{2,inv}-OP), such that the upon activation of the reverse promoter, the state of the system is flipped from its inverted state back to its original state. In some such embodiments, the first and second inducible promoters are induced by different agents. In all such embodiments, the SIMM can further comprise one or more components such as degradation tags, ribosome binding sites, transcriptional terminator sequences, and antisense RNA sequences to further regulate the activated and steady-states of the SIMM.

[0079] In some embodiments of these aspects and all such aspects described herein, one or more additional components can be added to the SIMMs to increase the utility or functionality of the SIMM for use in the biological circuit chemotactic converters provided herein. In some embodiments, a SIMM comprises a forward promoter sequence, a forward recombinase recognition sequence, a ribosome-binding site sequence, a recombinase gene sequence, a transcriptional terminator sequence, an inverted output product sequence, an inverted ribosome-binding site sequence, and a reverse recombinase recognition site sequence, *i.e.*, P_{for} - RRS_{for} - RBS - RC - T - OP_{inv} - RBS_{inv} - RRS_{rev}; where P_{for} is a forward promoter sequence; RRS_{for} is a forward recombinase recognition sequence; RBS is a ribosome-binding site sequence; RC is a recombinase gene sequence encoding a recombinase that recognizes RRS_{for} and RRS_{rev}; T is a transcriptional terminator sequence; OP_{inv} is the inverted sequence of any gene that can be used as an output; RBS_{inv} is an inverted ribosome-binding site sequence; and RRS_{rev} is a reverse recombinase recognition sequence. In such embodiments, upon activation of the forward promoter (P_{for}), the recombinase gene (RC) is expressed, causing inversion of the sequence between the two recombinase recognition sequences (RRS_{for} and RRS_{rev}), thus allowing for expression of the output product sequence that is no longer in the inverted direction. In some embodiments, the output product sequence encodes a transcriptional repressor or activator. In some embodiments, the output product sequence encodes a reporter gene.

[0080] In other embodiments of these aspects and all such aspects described herein, a SIMM is provided that comprises a forward promoter sequence, a forward recombinase recognition sequence, a ribosome-binding site sequence, a recombinase gene sequence, a ribosome-binding site sequence, an output product sequence, an inverted ribosome-binding site sequence, and a reverse recombinase recognition site sequence, *i.e.*, $P_{\text{for}} - \text{RRS}_{\text{for}} - \text{RBS} - \text{RC} - \text{RBS} - \text{OP} - \text{RBS}_{\text{inv}} - \text{RRS}_{\text{rev}}$, where P_{for} is a forward promoter; RRS_{for} is a forward recombinase recognition sequence; RBS are ribosome-binding site sequences; RC is a recombinase gene sequence encoding a recombinase that recognizes RRS_{for} and RRS_{rev} ; OP is the sequence of any output product, such as a protein or RNA molecule; RBS_{inv} is an inverted ribosome-binding site sequence; and RRS_{rev} is a reverse recombinase recognition sequence. In such embodiments, activation of the forward promoter sequence results in expression of both the recombinase and the output product, and then upon inversion of the sequence due to the activity of the recombinase, the expression of the output product is shut off. Thus, in such embodiments, the SIMM creates a single pulse of expression of an output gene product. In some embodiments, the output product sequence encodes a transcriptional repressor or transcriptional activator. In other embodiments, the output product sequence encodes an RNA molecule, such as an iRNA molecule, an antisense RNA molecule, or a microRNA molecule. Other non-limiting examples of output products for use in the SIMMs described herein are reporter proteins (*e.g.*, green fluorescent protein), transcription factors, transcriptional repressors, or RNA molecules, such as riboswitches in prokaryotic and mammalian cells, as well as short-hairpin RNAs, antisense RNA molecules, and microRNA molecules in mammalian cells (F. J. Isaacs, *Nat Biotechnol* 22, 841 (2004)). Further non-limiting examples of output products for use in the SIMMs described herein, are provided in the sections entitled “*Output Products*” and “*RNA Molecule Components and Output Products*.”

Biological Circuit Chemotactic Converters

[0081] Described herein are novel biological circuit chemotactic converters that provide the ability to convert input signals received into expression of specific “sensory receptors” or “sensors.” Expression of such sensors enable a biological system, such as a natural or synthetic (*e.g.*, artificial) cell, to activate the necessary molecular components to move, or chemotaxis, in response to a chemotactic signal. The modular nature of the biological circuit chemotactic converters described herein permits flexibility and expansion of the converters to vary the range and sensitivity of input signals to which the biological circuit chemotactic converters can respond, and increases the numbers and combinations of sensors that are expressed, depending on the specific input signals received. Such signals that can act as input signals to the biological circuit chemotactic converters described herein include, but are not limited to, concentrations of inducing agents, which may include biological agents such as pheromones, hormones, growth factors, metabolites, and the like; concentrations of chemicals, environmental byproducts, metal ions, and other such molecules or agents; light levels; temperature; mechanical stress; or electrical signals, such as currents and voltages.

[0082] The biological circuit chemotactic converters provided herein comprise at least five modules- an input module, a genetic toggle switch, a logic module, and at least two sensor modules. The input module comprises an inducible promoter sequence operably linked to a repressor sequence, *i.e.*, iP_A-R_A . In some embodiments of the aspects described herein, the input module further comprises an enhancer, a ribosome binding site, a degradation tag, an antisense RNA or microRNA sequence, and/or a terminator sequence. In some embodiments of the aspects described herein, the inducible promoter is an engineered promoter. In some embodiments of the aspects described herein, the inducing agent for the inducible promoter is an exogenous chemical, metal ion, toxin, or pollutant.

[0083] The genetic toggle switches provided herein are bistable switches that comprise a mutually inhibitory arrangement of repressor sequences, the products of which are specific for the sequence of the promoter sequence the other repressor sequence is operably linked to. In some embodiments of the aspects described herein, the genetic toggle switch comprises a first repressible promoter sequence (rP_1) that drives the transcription of a second repressor sequence (R_2) that encodes a repressor specific for the second repressible promoter sequence, and a second repressible promoter sequence (rP_2) that drives the transcription of a first repressor sequence (R_1), that represses transcription from rP_1 , *i.e.*, the genetic toggle switch (TS) comprises rP_1-R_2 and rP_2-R_1 . In some embodiments of the aspects described herein, the repressor encoded by the input module R_A acts as a repressor for rP_1 , such that when the input module receives an input and transcribes R_A , then transcription driven by rP_1 of the genetic toggle switch is repressed, allowing transcription to proceed from rP_2 , resulting in transcription of R_1 of the genetic toggle switch. R_1 serves as a further repressor of transcription driven by rP_1 of the genetic toggle switch. In some embodiments of the aspects described herein, the repressor of the input module (R_A) and the first repressor of the toggle switch (R_1) are the same repressor, *i.e.*, $R_A=R_1$.

[0084] The logic module of the biological circuit chemotactic converters described herein comprises a repressible promoter operably linked to a repressor sequence, *i.e.*, rP_B-R_B . In some embodiments, the logic module further comprises an enhancer, a ribosome binding site, a degradation tag, an antisense RNA or microRNA sequence, and/or a terminator sequence. In some embodiments of these aspects, the repressible promoter of the logic module (rP_B) is repressed by the repressor encoded by the second repressor of the genetic toggle switch, such that rP_B is repressed by R_2 . In some embodiments of these aspects, the inducing agent for the inducible promoter is a biological agent, such as a pheromone, hormone, growth factor, or metabolite; or an exogenous chemical, metal ion, toxin, or pollutant. In some embodiments of these aspects, $rP_B = rP_2$ of the genetic toggle switch. Thus, in such embodiments, in the absence of an inducing agent received by the inducible promoter of the input module (iP_A), the genetic toggle switch transcribes the second repressor (R_2), which represses both rP_2 of the genetic toggle switch and the promoter (rP_B) of the logic module.

[0085] The sensor modules of the biological circuit chemotactic converter described herein each comprise a repressible promoter operably linked to a sequence encoding a sensor, *i.e.*, rP_C -sensor

A and rP_D -sensor B. In some embodiments, the two sensor sequences encode for a different sensor receptor. In some embodiments, the repressible promoter of one sensor module is repressed by the repressor encoded by the logic module, and the repressible promoter of the other sensor module is repressed by the repressor encoded by the second repressor (R_2) of the genetic toggle switch. For example, in such embodiments, rP_C is repressed by R_B of the logic module, which prevents transcription of sensor A, and rP_C is repressed by R_2 of the genetic toggle switch, which prevents transcription of sensor B. In some embodiments, $rP_D = rP_2$. In some embodiments, $rP_D = rP_B$. In some embodiments, $rP_D = rP_B = rP_2$. Thus, in such biological circuit chemotactic converters, when no input is received by the input module, then the second repressor of the genetic toggle switch (R_2) is expressed and inhibits expression of the sensor having a promoter repressible by R_2 , as well as the first repressor of the genetic toggle switch, and repressor B of the logic module. In the example described herein, in the absence of an input, only sensor A would be expressed.

[0086] Accordingly, in such embodiments, upon the input module receiving an appropriate input signal, repressor A of the input module is transcribed, which represses transcription from rP_1 of the genetic toggle switch, allowing transcription of the first repressor from rP_2 of the genetic toggle switch, and further inhibition of transcription from rP_1 . In the absence of the second repressor of the genetic toggle switch, repression of the logic module and the R_2 repressible sensor module is removed, (*i.e.*, sensor B), thus allowing transcription of repressor B of the logic module, and sensor B. Expression of repressor B of the logic module represses transcription from the sensor module having a promoter that can be repressed by repressor B, *i.e.*, in this example, sensor A.

[0087] In some embodiments of the aspects described herein, the biological circuit chemotactic converters comprise multiple such modular components to increase the sensitivity and range of chemotactic responses in a biological system, such as a cell or organism. For example, in some embodiments, the biological circuit chemotactic converter comprises at least 2 input modules, at least 2 genetic toggle switches, at least 2 logic modules, and at least 3 sensor modules, as depicted in Figure 1B. In such an embodiment, upon the input modules receiving appropriate input signals, different sensors are expressed, allowing different chemotactic responses to be achieved. For example, in the embodiment described in Figure 1B, when neither input module receives an input signal, only sensor A is expressed, and the biological circuit chemotactic converter can be responsive, *i.e.*, undergo chemotaxis, to the chemoattractant recognized by sensor A. Similarly, in the example shown in Figure 1B, if the first input module, which encodes for repressor 1, receives an appropriate input signal, then the specific combination of repressors expressed results in expression of only sensor B, and the biological circuit chemotactic converter can be responsive to the chemoattractant recognized by sensor B. If the second input module, in the example shown in Figure 1B, receives an appropriate input signal to begin transcription of repressor 3, then only sensor C is expressed and the biological circuit chemotactic converter can undergo chemotaxis to the chemoattractant recognized by sensor C.

[0088] In other embodiments of this aspect and all such aspects described herein, the biological circuit chemotactic converter comprises: at least 3 input modules, at least 3 genetic toggle switches, at least 3 logic modules, and at least 4 sensor modules; at least 4 input modules, at least 4 genetic toggle switches, at least 4 logic modules, and at least 5 sensor modules; at least 5 input modules, at least 5 genetic toggle switches, at least 5 logic modules, and at least 6 sensor modules; at least 6 input modules, at least 6 genetic toggle switches, at least 6 logic modules, and at least 7 sensor modules; at least 7 input modules, at least 7 genetic toggle switches, at least 7 logic modules, and at least 8 sensor modules; at least 8 input modules, at least 8 genetic toggle switches, at least 8 logic modules, and at least 9 sensor modules; at least 9 input modules, at least 9 genetic toggle switches, at least 9 logic modules, and at least 10 sensor modules; at least 10 input modules, at least 10 genetic toggle switches, at least 10 logic modules, and at least 11 sensor modules; at least 10 input modules, at least 10 genetic toggle switches, at least 10 logic modules, and at least 11 sensor modules; at least 11 input modules, at least 11 genetic toggle switches, at least 11 logic modules, and at least 12 sensor modules; at least 12 input modules, at least 12 genetic toggle switches, at least 12 logic modules, and at least 13 sensor modules; at least 13 input modules, at least 13 genetic toggle switches, at least 13 logic modules, and at least 14 sensor modules; at least 14 input modules, at least 14 genetic toggle switches, at least 14 logic modules, and at least 15 sensor modules; at least 15 input modules, at least 15 genetic toggle switches, at least 15 logic modules, and at least 16 sensor modules; at least 16 input modules, at least 16 genetic toggle switches, at least 16 logic modules, and at least 17 sensor modules; at least 17 input modules, at least 17 genetic toggle switches, at least 17 logic modules, and at least 18 sensor modules; at least 18 input modules, at least 18 genetic toggle switches, at least 18 logic modules, and at least 19 sensor modules; at least 19 input modules, at least 19 genetic toggle switches, at least 19 logic modules, and at least 20 sensor modules; at least 20 input modules, at least 20 genetic toggle switches, at least 20 logic modules, and at least 21 sensor modules; at least 50 input modules, at least 50 genetic toggle switches, at least 50 logic modules, and at least 51 sensor modules; at least 100 input modules, at least 100 genetic toggle switches, at least 100 logic modules, and at least 101 sensor modules; or at least n input modules, at least n genetic toggle switches, at least n logic modules, and at least $n+1$ sensor modules, where n is any integer value greater than or equal to 1.

[0089] In all embodiments of the aspects described herein, the modules of the biological circuit chemotactic converters can further comprise additional components such as one or more RBSs, degradation tag sequences, termination sequences, antisense RNA or microRNA sequences, and/or enhancer sequences. In some embodiments, the input module further comprises a ribosome binding sequence, a degradation tag sequence, a termination sequence, an antisense RNA, a microRNA sequence, an enhancer sequence, or any combination thereof. In some embodiments, the genetic toggle switch further comprises a degradation tag sequence, a termination sequence, an antisense RNA, a microRNA sequence, an enhancer sequence, or any combination thereof. In some embodiments, the logic module further comprises a ribosome binding sequence, a degradation tag

sequence, a termination sequence, an antisense RNA, a microRNA sequence, an enhancer sequence, or any combination thereof. In some embodiments, the sensor module further comprises a ribosome binding sequence, a degradation tag sequence, a termination sequence, an antisense RNA, a microRNA sequence, an enhancer sequence, or any combination thereof.

Applications of Biological Circuit Chemotactic Converters

[0090] Provided herein are biological circuit chemotactic converters for use in engineering complex behavioral phenotypes in cellular systems, such as prokaryotic, eukaryotic, or synthetic cells. The novel biological circuit chemotactic converters described herein combine the power of protein engineering methods with systems biology approaches to elicit targeted responses in cellular systems.

Chemotactic Responses in Cellular Organisms

[0091] Chemotaxis in bacteria has been studied as a model system for receptor signaling and intracellular signal transduction. To actively seek energy sources or find optimal physicochemical environments, bacteria such as *E. coli* employ a system of alternating tumbling and swimming phases. Bacteria such as *E. coli* are unable to choose the direction in which they swim, and are unable to swim in a straight line for more than a few seconds due to rotational diffusion. In the presence of a chemical gradient bacteria will chemotax, or direct their overall motion based on the gradient. If the bacterium senses that it is moving in the correct direction, for example, toward attractant/away from repellent, it keeps swimming in a straight line for a longer time before tumbling. If it is moving in the wrong direction, it tumbles sooner and try a new direction at random. Thus, bacteria like *E. coli* use temporal sensing, to find, for example, the location with the highest concentration of attractant. Even under very high concentrations, it can still distinguish very small differences in concentration. Fleeing from a repellent works with the same efficiency.

[0092] The helical nature of the individual flagellar filament is critical for this movement to occur. Some bacteria, such as *E. coli*, have several flagella per cell (4–10 typically), that can rotate in two ways. Counter-clockwise rotation aligns the flagella into a single rotating bundle, causing the bacterium to swim in a straight line. Clockwise rotation breaks the flagella bundle apart such that each flagellum points in a different direction, causing the bacterium to tumble in place. The protein that makes up the flagellar filament, flagellin, is quite similar among all flagellated bacteria.

[0093] Some bacteria, such as *Vibrio*, are monoflagellated and have a single flagellum at one pole of the cell, and used a different method of chemotaxis. Others possess a single flagellum that is kept inside the cell wall. These bacteria move by spinning the whole cell, which is shaped like a corkscrew.

[0094] Chemical gradients are sensed through multiple transmembrane receptors, called methyl accepting chemotaxis proteins (MCPs), which vary in the molecules that they detect. In some aspects of the invention, sequences encoding these receptors, or fragments thereof, are used in the sensor modules of the biological circuit chemotactic converters. These receptors can bind attractants or repellents directly or indirectly through interaction with proteins of periplasmic space. The

signals from these receptors are transmitted across the plasma membrane into the cytosol, where Che proteins are activated. The Che proteins alter the tumbling frequency, and alter the receptors.

[0095] The proteins CheW and CheA bind to the receptor. The activation of the multiple transmembrane receptor by an external stimulus causes autophosphorylation in the histidine kinase, CheA, at a single highly conserved histidine residue. CheA in turn transfers phosphoryl groups to conserved aspartate residues in the response regulators CheB and CheY. This mechanism of signal transduction is called a “two-component system” and is a common form of signal transduction in bacteria. CheY induces tumbling by interacting with the flagellar switch protein FliM, inducing a change from counter-clockwise to clockwise rotation of the flagellum. Change in the rotation state of a single flagellum can disrupt the entire flagella bundle and cause a tumble.

[0096] CheB, when activated by CheA, acts as a methylesterase, removing methyl groups from glutamate residues on the cytosolic side of the receptor. It works antagonistically with CheR, a methyltransferase, which adds methyl residues to the same glutamate residues. The more methyl residues are attached to the receptor, the less sensitive the receptor. As the signal from the receptor induces demethylation of the receptor in a feedback loop, the system is continuously adjusted to environmental chemical levels, remaining sensitive for small changes even under extreme chemical concentrations. This regulation allows the bacterium to 'remember' chemical concentrations from the recent past, a few seconds, and compare them to those it is currently experiencing, thus 'know' whether it is traveling up or down a gradient. Additional regulatory mechanisms such as receptor clustering and receptor-receptor interactions also modulate the signalling pathway.

[0097] The mechanism by which eukaryotic cells undergo chemotaxis differs from that in bacteria, however, sensing of chemical gradients is still a crucial step in the process. In contrast to prokaryotes, such as *E.coli*, the size of eukaryotic cells allows for the possibility of detecting gradients via dynamic and polarized distribution of receptors. Induction of these receptors by chemoattractants or chemorepellents results in migration towards or away from the chemotactic substance.

[0098] Levels of receptors, intracellular signalling pathways and the effector mechanisms all represent diverse, eukaryotic type components that are contemplated for use in the modules and biological circuit chemotactic converters described herein. In eukaryotic unicellular cells, amoeboid movement and cilium or the eukaryotic flagellum are the main effectors (e.g. Amoeba or Tetrahymena). Some eukaryotic cells of higher vertebrate origin, such as immune cells also move to where they need to be by sensing chemotactic gradients. Besides immune competent cells (granulocyte, monocyte, lymphocyte) a large group of cells - considered previously to be fixed into tissues - are also motile in special physiological (e.g. mast cell, fibroblast, endothelial cells) or pathological conditions (e.g. metastases). Chemotaxis has high significance in the early phases of embryogenesis as development of germ layers is guided by gradients of biological signal molecules.

[0099] In eukaryotic cells, the mechanisms regulating chemotaxis include mechanisms by which an external chemotactic gradient is sensed and turned into an intracellular PIP3 gradient, which

results in a gradient and the activation of a signaling pathway, culminating in the polymerisation of actin filaments. The growing distal end of actin filaments develops connections with the internal surface of the plasma membrane via different sets of peptides and results in the formation of pseudopods. Cilia of eukaryotic cells can also produce chemotaxis; in this case it is mainly a Ca^{2+} -dependent induction of the microtubular system of the basal body and the beat of the 9+2 microtubules within cilia. The orchestrated beating of hundreds of cilia is synchronized by a submembranous system built between basal bodies.

[00100] For the most part, eukaryotic cells sense the presence of chemotactic stimuli through the use of 7-transmembrane (or serpentine) heterotrimeric G-protein coupled receptors. Accordingly, in some embodiments, heterotrimeric G-protein coupled receptors are contemplated for use in the modules and biological circuit chemotactic converters described herein. Other professional chemotaxis receptors in eukaryotes that can be used in the modules and biological circuit chemotactic converters described herein are those triggered by formyl peptides, i.e., formyl peptide receptors (FPR); chemokines, i.e., chemokine receptors (CCR or CXCR); and leukotriene, i.e., leukotriene receptors (BLT), as described herein. In addition, induction of a wide set of membrane receptors (e.g. amino acids, insulin, vasoactive peptides) can also elicit migration of the cell and be used in the modules and biological circuit chemotactic converters described herein.

[00101] While some chemotaxis receptors are expressed in the surface membrane with long-term characteristics as they are determined genetically, others have short-term dynamics as they are assembled ad hoc in the presence of the ligand. The diverse features of the chemotaxis receptors and ligands allows for the possibility of selecting chemotactic responder cells with a simple chemotaxis assay. The main groups of primary ligands in eukaryotic cells include:

[00102] Formyl peptides, which are di-, tri-, tetrapeptides of bacterial origin, released from bacteria in vivo or after decomposition of the cell. A typical member of this group is the N-formylmethionyl-leucyl-phenylalanine (fMLF or fMLP). The bacterial origin fMLF as a key component of inflammation has characteristic chemoattractant effects in neutrophil granulocytes and monocytes.

[00103] Complement 3a (C3a) and complement 5a (C5a) are intermediate products of the complement cascade. Their synthesis is joined to the three alternative pathways (classical, lectin dependent and alternative) of complement activation by a convertase enzyme. The main target cells of these derivatives are neutrophil granulocytes and monocytes as well.

[00104] Chemokines belong to a special class of cytokines. Their groups (C, CC, CXC, CX3C chemokines) represent not only structurally related molecules with a special arrangement of disulfide bridges, but their target cell specificity is also diverse: CC chemokines act on monocytes (e.g. RANTES), CXC chemokines are neutrophil granulocyte specific (e.g. IL-8).

[00105] Leukotrienes belong to the group eicosanoids. They are significant lipid mediators of the arachidonic acid cascade converted by 5-lipoxygenase. Their predominant member is leukotriene

B4 (LTB4) which elicits adhesion, chemotaxis and aggregation of leukocytes. The characteristic chemoattractant effect of LTB4 is induced via G-protein linked seven-transmembrane spanning leukotriene receptors which are highly expressed in inflammation and allergy.

Engineering Chemotactic Responses in Cellular Organisms

[00106] By integrating various modular components to provide specific responses chemotactic pathways, it is possible to engineer an artificial chemotactic system, as described herein. In some aspects, biological circuit chemotactic converters are provided for use in organisms, such as bacteria, to modulate responses to chemotactic gradients. A variety of processing elements can be integrated at the genetic level to obtain gradient responses in such a converter, such as stable biphasic switches (for example, the genetic toggle switches described herein). Integration with another input can be achieved through biological logic modules, such as AND gates. The usage of such biological circuit chemotactic converters to modulate a system response can be modeled to characterize the dynamics of the genetic circuitry and can also be fine-tuned during empirical testing of the system.

[00107] In some aspects, biological circuit chemotactic converters are provided for use in engineering bacteria, such as *E.coli*, to provide directed chemotactic responses. In some embodiments of these aspects, the chemotactic response is further coupled to the production of an output product specific for the chemoattractant response. For example, upon sensing an environmental toxin, an organism engineered with a biological circuit chemotactic converter described herein, can undergo chemotaxis towards the source of the environmental toxin, resulting in an output “elimination response,” such as the production of an enzyme that degrades the toxin. In other such embodiments, the biological circuit chemotactic converter can be used to mobilize an engineered organism towards the source of the environmental toxin to allow for visual identification of a region where there is high concentration of a toxin or pollutant, by, for example, coupling the biological circuit chemotactic converter to a reporter gene, such as GFP. In some embodiments, the biological circuit chemotactic converter is used in a response to eliminate a target, such as another bacteria.

[00108] In some embodiments of these aspects, fusion or chimeric polypeptides are encoded by the sensor modules of the biological circuit chemotactic converters described herein. Such fusion polypeptides can comprise a sensor domain, a transmembrane domain, and a cytoplasmic domain. In some embodiments, the domains are selected from chemotactic receptors of different organisms, *e.g.*, different bacteria, or *e.g.*, a eukaryotic sensor domain coupled to a prokaryotic cytoplasmic domain.

[00109] The methods and uses of the biological circuit chemotactic converters described herein can involve *in vivo*, *ex vivo*, or *in vitro* systems. The term “*in vivo*” refers to assays or processes that occur in or within an organism, such as a multicellular animal. In some of the aspects described herein, a method or use can be said to occur “*in vivo*” when a unicellular organism, such as a bacteria, is used. The term “*ex vivo*” refers to methods and uses that are performed using a living cell with an intact membrane that is outside of the body of a multicellular animal or plant, *e.g.*, explants, cultured

cells, including primary cells and cell lines, transformed cell lines, and extracted tissue or cells, including blood cells, among others. The term "*in vitro*" refers to assays and methods that do not require the presence of a cell with an intact membrane, such as cellular extracts, and can refer to the introducing of a biological circuit chemotactic converters in a non-cellular system, such as a media not comprising cells or cellular systems, such as cellular extracts.

[00110] A cell or cellular system to be engineered for use with the biological circuit chemotactic converters described herein can be any cell. As defined herein, a "cell" is the basic structural and functional unit of all known living organisms. It is the smallest unit of life that is classified as a living thing, and is often called the building block of life. Some organisms, such as most bacteria, are unicellular (consist of a single cell). Other organisms, such as humans, are multicellular.

[00111] A "natural cell," as defined herein, refers to any prokaryotic or eukaryotic cell that occurs in nature. A "prokaryotic cell" can comprise a cell envelope and a cytoplasmic region that contains the cell genome (DNA) and ribosomes and various sorts of inclusions. In some embodiments, the cell is a eukaryotic cell. As used herein, a "eukaryotic cell" comprises membrane-bound compartments in which specific metabolic activities take place, such as for example, a nucleus. In some embodiments, the cell is an artificial or synthetic cell. As defined herein, an "artificial cell" or a "synthetic cell" is a minimal cell formed from artificial parts, that can do many things a natural cell can do, such as transcribe and translate proteins and generate ATP.

[00112] Host cells of use in the aspects described herein for transformation or transfection with the biological circuit chemotactic converters described herein include any host cell that is capable of supporting the activation and expression of the biological circuit chemotactic converters. In some embodiments of the aspects described herein, the cells are bacterial cells. The term "bacteria" as used herein is intended to encompass all variants of bacteria, for example, prokaryotic organisms and cyanobacteria. Bacteria are small (typical linear dimensions of around 1 μ m), non-compartmentalized, with circular DNA and ribosomes of 70S. The term bacteria also includes bacteria subdivisions of Eubacteria and Archaeobacteria. Eubacteria can be further subdivided on the basis of their staining using Gram stain, and both gram-positive and gram-negative eubacteria, which depends upon a difference in cell wall structure are also included, as well as classified based on gross morphology alone (into cocci, bacilli, etc.).

[00113] In some embodiments, the bacterial cells are gram-negative cells and in alternative embodiments, the bacterial cells are gram-positive cells. Non-limiting examples of species of bacterial cells useful for engineering with the biological circuit chemotactic converters of the invention include, without limitation, cell from *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species of *Pseudomonas*, *Streptomyces*, and *Staphylococcus*. Other examples of bacterial cells that can be genetically engineered for use with the biological circuit chemotactic converters of the invention include, but are not limited to, cells from *Yersinia* spp., *Escherichia* spp., *Klebsiella* spp.,

Bordetella spp., *Neisseria* spp., *Aeromonas* spp., *Francisella* spp., *Corynebacterium* spp., *Citrobacter* spp., *Chlamydia* spp., *Hemophilus* spp., *Brucella* spp., *Mycobacterium* spp., *Legionella* spp., *Rhodococcus* spp., *Pseudomonas* spp., *Helicobacter* spp., *Salmonella* spp., *Vibrio* spp., *Bacillus* spp., and *Erysipelothrix* spp. In some embodiments, the bacterial cells are *E.coli* cells. Other examples of organisms from which cells can be transformed or transfected with the biological circuit chemotactic converters described herein include, but are not limited to the following: *Staphylococcus aureus*, *Bacillus subtilis*, *Clostridium butyricum*, *Brevibacterium lactofermentum*, *Streptococcus agalactiae*, *Lactococcus lactis*, *Leuconostoc lactis*, *Streptomyces*, *Actinobacillus actinobycetemcomitans*, *Bacteroides*, *cyanobacteria*, *Escherichia coli*, *Helobacter pylori*, *Selenomonas ruminantium*, *Shigella sonnei*, *Zymomonas mobilis*, *Mycoplasma mycoides*, or *Treponema denticola*, *Bacillus thuringiensis*, *Staphylococcus lugdunensis*, *Leuconostoc oenos*, *Corynebacterium xerosis*, *Lactobacillus plantarum*, *Streptococcus faecalis*, *Bacillus coagulans*, *Bacillus ceretus*, *Bacillus popilliae*, *Synechocystis* strain PCC6803, *Bacillus liquefaciens*, *Pyrococcus abyssi*, *Selenomonas nominantium*, *Lactobacillus hilgardii*, *Streptococcus ferus*, *Lactobacillus pentosus*, *Bacteroides fragilis*, *Staphylococcus epidermidis*, *Streptococcus epidermidis*, *Zymomonas mobilis*, *Streptomyces phaeochromogenes*, *Streptomyces ghanaensis*, *Halobacterium* strain GRB, and *Halobaferax* sp. strain Aa2.2.

[00114] Cells of use in the various aspects described herein upon transformation or transfection with the biological circuit chemotactic converters described herein include any cell that is capable of supporting the activation and expression of the biological circuit chemotactic converters. In some embodiments of the aspects described herein, a cell can be from any multicellular or eukaryotic organism. Examples of eukaryotic cells that can be useful in aspects described herein include eukaryotic cells selected from, *e.g.*, mammalian, insect, yeast, or plant cells. The biological circuit chemotactic converters described herein can be introduced into a variety of cells including, *e.g.*, fungal, plant, or animal (nematode, insect, plant, bird, reptile, or mammal (*e.g.*, a mouse, rat, rabbit, hamster, gerbil, dog, cat, goat, pig, cow, horse, whale, monkey, or human)). The cells can be primary cells, immortalized cells, stem cells, or transformed cells. In some embodiments, the cells comprise stem cells. Expression vectors for the components of the biological circuit chemotactic converters generally comprise a promoter and/or an enhancer suitable for expression in a particular host cell of interest. Other cell types also contemplated for use with the biological circuit chemotactic converters described herein, include, but are not limited to, reproductive cells including sperm, ova and embryonic cells, and non-reproductive cells, such as kidney, lung, spleen, lymphoid, cardiac, gastric, intestinal, pancreatic, muscle, bone, neural, brain, and epithelial cells. Also contemplated for use with the biological circuit chemotactic converters described herein are stem cells, including human embryonic stem cells, pluripotent stem cells, multipotent stem cells, and induced pluripotent stem cells (iPSCs), as those terms are understood by one of skill in the art.

[00115] As used herein, the term "stem cells" is used in a broad sense and includes traditional stem cells, progenitor cells, preprogenitor cells, reserve cells, and the like. The term "stem cell" or "progenitor cell" are used interchangeably herein, and refer to an undifferentiated cell which is capable of proliferation and giving rise to more progenitor cells having the ability to generate a large number of mother cells that can in turn give rise to differentiated, or differentiable daughter cells. Stem cells for use with the biological circuit chemotactic converters and the methods described herein can be obtained from endogenous sources such as cord blood, or can be generated using *in vitro* or *ex vivo* techniques as known to one of skill in the art. For example, a stem cell can be an induced pluripotent stem cell (iPS cell).

[00116] Exemplary stem cells include, but are not limited to, embryonic stem cells, adult stem cells, pluripotent stem cells, induced pluripotent stem cells (iPS cells), neural stem cells, liver stem cells, muscle stem cells, muscle precursor stem cells, endothelial progenitor cells, bone marrow stem cells, chondrogenic stem cells, lymphoid stem cells, mesenchymal stem cells, hematopoietic stem cells, central nervous system stem cells, peripheral nervous system stem cells, and the like. Descriptions of stem cells, including method for isolating and culturing them, can be found in, among other places, Embryonic Stem Cells, Methods and Protocols, Turksen, ed., Humana Press, 2002; Weisman *et al.*, Annu. Rev. Cell. Dev. Biol. 17:387 403; Pittinger *et al.*, Science, 284:143 47, 1999; Animal Cell Culture, Masters, ed., Oxford University Press, 2000; Jackson *et al.*, PNAS 96(25):14482 86, 1999; Zuk *et al.*, Tissue Engineering, 7:211 228, 2001 ("Zuk *et al.*"); Atala *et al.*, particularly Chapters 33 41; and U.S. Pat. Nos. 5,559,022, 5,672,346 and 5,827,735. Descriptions of stromal cells, including methods for isolating them, can be found in, among other places, Prockop, Science, 276:71 74, 1997; Theise *et al.*, Hepatology, 31:235 40, 2000; Current Protocols in Cell Biology, Bonifacino *et al.*, eds., John Wiley & Sons, 2000 (including updates through March, 2002); and U.S. Pat. No. 4,963,489; Phillips BW and Crook JM, Pluripotent human stem cells: A novel tool in drug discovery. BioDrugs. 2010 Apr 1;24(2):99-108; Mari Ohnuki *et al.*, Generation and Characterization of Human Induced Pluripotent Stem Cells, Current Protocols in Stem Cell Biology Unit Number: UNIT 4A., September, 2009.

[00117] As indicated above, there are different levels or classes of cells falling under the general definition of a "stem cell." These are "totipotent," "pluripotent" and "multipotent" stem cells. The term "totipotency" or "totipotent" refers to a cell with the degree of differentiation describing a capacity to make all of the cells in the adult body as well as the extra-embryonic tissues including the placenta. The fertilized egg (zygote) is totipotent as are the early cleaved cells (blastomeres)

[00118] The term "pluripotent" or a "pluripotent state" as used herein refers to a cell with the capacity, under different conditions, to differentiate to cell types characteristic of all three germ cell layers: endoderm (gut tissue), mesoderm (including blood, muscle, and vessels), and ectoderm (such as skin and nerve). Pluripotent cells are characterized primarily by their ability to differentiate to all three germ layers, using, for example, a nude mouse teratoma formation assay. Pluripotency is also

evidenced by the expression of embryonic stem (ES) cell markers, although the preferred test for pluripotency is the demonstration of the capacity to differentiate into cells of each of the three germ layers. In some embodiments, a pluripotent cell is an undifferentiated cell.

[00119] The term "multipotent" when used in reference to a "multipotent cell" refers to a cell that is able to differentiate into some but not all of the cells derived from all three germ layers. Thus, a multipotent cell is a partially differentiated cell. Multipotent cells are well known in the art, and examples of multipotent cells include adult stem cells, such as for example, hematopoietic stem cells and neural stem cells. Multipotent means a stem cell can form many types of cells in a given lineage, but not cells of other lineages. For example, a multipotent blood stem cell such as a "hematopoietic stem cells" refers to all stem cells or progenitor cells found *inter alia* in bone marrow and peripheral blood that are capable of differentiating into any of the specific types of hematopoietic or blood cells, such as erythrocytes, lymphocytes, macrophages and megakaryocytes. The term "multipotency" refers to a cell with the degree of developmental versatility that is less than totipotent and pluripotent.

[00120] In the context of cell ontogeny, the adjectives "differentiated", or "differentiating" are relative terms. The term "differentiation" in the present context means the formation of cells expressing markers known to be associated with cells that are more specialized and closer to becoming terminally differentiated cells incapable of further differentiation. The pathway along which cells progress from a less committed cell, to a cell that is increasingly committed to a particular cell type, and eventually to a terminally differentiated cell is referred to as progressive differentiation or progressive commitment. Cells which are more specialized (*e.g.*, have begun to progress along a path of progressive differentiation) but not yet terminally differentiated are referred to as partially differentiated. Differentiation is a developmental process whereby cells assume a specialized phenotype, *e.g.*, acquire one or more characteristics or functions distinct from other cell types. In some cases, the differentiated phenotype refers to a cell phenotype that is at the mature endpoint in some developmental pathway (a so called terminally differentiated cell). In many, but not all tissues, the process of differentiation is coupled with exit from the cell cycle. In these cases, the terminally differentiated cells lose or greatly restrict their capacity to proliferate. However, we note that in the context of this specification, the terms "differentiation" or "differentiated" refer to cells that are more specialized in their fate or function than at a previous point in their development, and includes both cells that are terminally differentiated and cells that, although not terminally differentiated, are more specialized than at a previous point in their development. The development of a cell from an uncommitted cell (for example, a stem cell), to a cell with an increasing degree of commitment to a particular differentiated cell type, and finally to a terminally differentiated cell is known as progressive differentiation or progressive commitment. A cell that is "differentiated" relative to a progenitor cell has one or more phenotypic differences relative to that progenitor cell. Phenotypic differences include, but are not limited to morphologic differences and differences in gene expression and biological activity, including not only the presence or absence of an expressed marker, but also

differences in the amount of a marker and differences in the co-expression patterns of a set of markers.

[00121] In other embodiments of the aspects described herein, biological circuit chemotactic converters can be introduced into a non-cellular system such as a virus or phage, by direct integration of the nucleic acid sequences encoding the various components and modules of the biological converter switch, for example, into the viral genome. A virus for use with the biological converter switches described herein can be a dsDNA virus (*e.g.* *Adenoviruses*, *Herpesviruses*, *Poxviruses*), a ssDNA viruses ((+)sense DNA) (*e.g.* *Parvoviruses*); a dsRNA virus (*e.g.* *Reoviruses*); a (+)ssRNA viruses ((+)sense RNA) (*e.g.* *Picornaviruses*, *Togaviruses*); (–)ssRNA virus (–)sense RNA) (*e.g.* *Orthomyxoviruses*, *Rhabdoviruses*); a ssRNA-Reverse Transcriptase viruses ((+)sense RNA with DNA intermediate in life-cycle) (*e.g.* *Retroviruses*); or a dsDNA- Reverse Transcriptase virus (*e.g.* *Hepadnaviruses*).

[00122] Viruses can also include plant viruses and bacteriophages or phages. Examples of phage families that can be used with the biological circuit chemotactic converters described herein include, but are not limited to, *Myoviridae* (*T4-like viruses*; *P1-like viruses*; *P2-like viruses*; *Mu-like viruses*; *SPO1-like viruses*; *φH-like viruses*); *Siphoviridae* λ -like viruses (*T1-like viruses*; *T5-like viruses*; *c2-like viruses*; *L5-like viruses*; *ψM1-like viruses*; *φC31-like viruses*; *N15-like viruses*); *Podoviridae* (*T7-like viruses*; *φ29-like viruses*; *P22-like viruses*; *N4-like viruses*); *Tectiviridae* (*Tectivirus*); *Corticoviridae* (*Corticovirus*); *Lipothrixviridae* (*Alphalipothrixvirus*, *Betalipothrixvirus*, *Gammalipothrixvirus*, *Deltalipothrixvirus*); *Plasmaviridae* (*Plasmavirus*); *Rudoviridae* (*Rudivirus*); *Fuselloviridae* (*Fusellovirus*); *Inoviridae* (*Inovirus*, *Plectrovirus*); *Microviridae* (*Microvirus*, *Spiromicrovirus*, *Bdellomicrovirus*, *Chlamydiamicrovirus*); *Leviviridae* (*Levivirus*, *Allolevivirus*) and *Cystoviridae* (*Cystovirus*). Such phages can be naturally occurring or engineered phages.

[00123] In some embodiments of the aspects described herein, the biological circuit chemotactic converters are introduced into a cellular or non-cellular system using a vector or plasmid. As used herein, the term “vector” is used interchangeably with “plasmid” to refer to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. Vectors capable of directing the expression of genes and/or nucleic acid sequence to which they are operatively linked are referred to herein as “expression vectors.” In general, expression vectors of utility in the methods and biological circuit chemotactic converters described herein are often in the form of “plasmids,” which refer to circular double stranded DNA loops which, in their vector form are not bound to the chromosome.

[00124] Other expression vectors can be used in different embodiments, for example, but not limited to, plasmids, episomes, bacteriophages or viral vectors, and such vectors can integrate into the host’s genome or replicate autonomously in the particular cellular system used. Other forms of expression vectors known by those skilled in the art which serve the equivalent functions can also be used. Expression vectors comprise expression vectors for stable or transient expression encoding the

DNA. A vector can be either a self replicating extrachromosomal vector or a vector which integrates into a host genome. One type of vector is a genomic integrated vector, or "integrated vector", which can become integrated into the chromosomal DNA or RNA of a host cell, cellular system, or non-cellular system. In some embodiments, the nucleic acid sequence or sequences encoding the biological circuit chemotactic converters integrates into the chromosomal DNA or RNA of a host cell, cellular system, or non-cellular system along with components of the vector sequence. In other embodiments, the nucleic acid sequence encoding the biological circuit chemotactic converter directly integrates into chromosomal DNA or RNA of a host cell, cellular system, or non-cellular system, in the absence of any components of the vector by which it was introduced. In such embodiments, the nucleic acid sequence encoding the biological circuit chemotactic converter can be integrated using targeted insertions, such as knock-in technologies or homologous recombination techniques, or by non-targeted insertions, such as gene trapping techniques or non-homologous recombination. The number of copies of a biological circuit chemotactic converter that integrate into the chromosomal DNA or RNA of a cellular or non-cellular system can impact the fidelity of the system, and thus it is preferred that only one copy is integrated per cellular system. Accordingly, in some embodiments of the aspects described herein, only one copy of a biological circuit chemotactic converter is integrated in the chromosomal DNA or RNA of a cellular or non-cellular system. In some embodiments, the number of copies is less than 10, less than 9, less than 8, less than 7, less than 6, less than 6, less than 4, less than 3, or less than 2.

[00125] Another type of vector is an episomal vector, *i.e.*, a nucleic acid capable of extra-chromosomal replication. Such plasmids or vectors can include plasmid sequences from bacteria, viruses or phages. Such vectors include chromosomal, episomal and virus-derived vectors *e.g.*, vectors derived from bacterial plasmids, bacteriophages, yeast episomes, yeast chromosomal elements, and viruses, vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, cosmids and phagemids. A vector can be a plasmid, bacteriophage, bacterial artificial chromosome (BAC) or yeast artificial chromosome (YAC). A vector can be a single or double-stranded DNA, RNA, or phage vector. In some embodiments, the biological circuit chemotactic converters are introduced into a cellular system using a BAC vector.

[00126] The vectors comprising the biological circuit chemotactic converters described herein can be "introduced" into cells as polynucleotides, preferably DNA, by techniques well-known in the art for introducing DNA and RNA into cells. The term "transduction" refers to any method whereby a nucleic acid sequence is introduced into a cell, *e.g.*, by transfection, lipofection, electroporation, biolistics, passive uptake, lipid:nucleic acid complexes, viral vector transduction, injection, contacting with naked DNA, gene gun, and the like. The vectors, in the case of phage and viral vectors can also be introduced into cells as packaged or encapsidated virus by well-known techniques for infection and transduction. Viral vectors can be replication competent or replication defective. In the latter case, viral propagation generally occurs only in complementing host cells. In some embodiments, the

biological circuit chemotactic converters are introduced into a cell using other mechanisms known to one of skill in the art, such as a liposome, microspheres, gene gun, fusion proteins, such as a fusion of an antibody moiety with a nucleic acid binding moiety, or other such delivery vehicle.

[00127] The biological circuit chemotactic converters or the vectors comprising the biological biological circuit chemotactic converters described herein can be introduced into a cell using any method known to one of skill in the art. The term "transformation" as used herein refers to the introduction of genetic material (*e.g.*, a vector comprising a biological circuit chemotactic converters) comprising one or more modules or biological circuit chemotactic converters described herein into a cell, tissue or organism. Transformation of a cell can be stable or transient. The term "transient transformation" or "transiently transformed" refers to the introduction of one or more transgenes into a cell in the absence of integration of the transgene into the host cell's genome. Transient transformation can be detected by, for example, enzyme linked immunosorbent assay (ELISA), which detects the presence of a polypeptide encoded by one or more of the transgenes. Alternatively, transient transformation can be detected by detecting the activity of the protein encoded by the transgene. The term "transient transformant" refers to a cell which has transiently incorporated one or more transgenes.

[00128] In contrast, the term "stable transformation" or "stably transformed" refers to the introduction and integration of one or more transgenes into the genome of a cell or cellular system, preferably resulting in chromosomal integration and stable heritability through meiosis. Stable transformation of a cell can be detected by Southern blot hybridization of genomic DNA of the cell with nucleic acid sequences, which are capable of binding to one or more of the transgenes. Alternatively, stable transformation of a cell can also be detected by the polymerase chain reaction of genomic DNA of the cell to amplify transgene sequences. The term "stable transformant" refers to a cell or cellular, which has stably integrated one or more transgenes into the genomic DNA. Thus, a stable transformant is distinguished from a transient transformant in that, whereas genomic DNA from the stable transformant contains one or more transgenes, genomic DNA from the transient transformant does not contain a transgene. Transformation also includes introduction of genetic material into plant cells in the form of plant viral vectors involving epichromosomal replication and gene expression, which may exhibit variable properties with respect to meiotic stability. Transformed cells, tissues, or plants are understood to encompass not only the end product of a transformation process, but also transgenic progeny thereof.

Component Parts

Promoters, and Promoter Inducing, Activating and Repressing Agents

[00129] Provided herein are promoters and promoter sequences are for use in the various modules and components of the biological circuit chemotactic converters described herein.

[00130] The term "promoter," as used herein, refers to any nucleic acid sequence that regulates the expression of another nucleic acid sequence by driving transcription of the nucleic acid

sequence, which can be a heterologous target gene encoding a protein or an RNA. Promoters can be constitutive, inducible, repressible, tissue-specific, or any combination thereof. A promoter is a control region of a nucleic acid sequence at which initiation and rate of transcription of the remainder of a nucleic acid sequence are controlled. A promoter can also contain genetic elements at which regulatory proteins and molecules can bind, such as RNA polymerase and other transcription factors. In some embodiments of the aspects, a promoter can drive the expression of a transcription factor that regulates the expression of the promoter itself, or that of another promoter used in another modular component of the biological circuit chemotactic converters described herein.

[00131] A promoter can be said to drive expression or drive transcription of the nucleic acid sequence that it regulates. The phrases “operably linked,” “operatively positioned,” “operatively linked,” “under control,” and “under transcriptional control” indicate that a promoter is in a correct functional location and/or orientation in relation to a nucleic acid sequence it regulates to control transcriptional initiation and/or expression of that sequence. An “inverted promoter,” as used herein, refers to a promoter in which the nucleic acid sequence is in the reverse orientation, such that what was the coding strand is now the non-coding strand, and vice versa. Inverted promoter sequences can be used in various embodiments of the invention to regulate the state of a module or a switch. In addition, in various embodiments of the invention, a promoter can be used in conjunction with an “enhancer,” which refers to a cis-acting regulatory sequence involved in the transcriptional activation of a nucleic acid sequence downstream of the promoter. The enhancer can be located at any functional location before or after the promoter, and/or the encoded nucleic acid.

[00132] A promoter can be one naturally associated with a gene or sequence, as can be obtained by isolating the 5' non-coding sequences located upstream of the coding segment and/or exon of a given gene or sequence. Such a promoter can be referred to as “endogenous.” Similarly, in some embodiments, an enhancer can be one naturally associated with a nucleic acid sequence, located either downstream or upstream of that sequence.

[00133] Alternatively, certain advantages are gained by positioning a coding nucleic acid segment under the control of a “recombinant promoter” or “heterologous promoter,” which refer to a promoter that is not normally associated with the encoded nucleic acid sequence it is operably linked to in its natural environment. A recombinant or heterologous enhancer refers to an enhancer not normally associated with a given nucleic acid sequence in its natural environment. Such promoters or enhancers can include promoters or enhancers of other genes; promoters or enhancers isolated from any other prokaryotic, viral, or eukaryotic cell; and synthetic promoters or enhancers that are not “naturally occurring,” *i.e.*, comprise different elements of different transcriptional regulatory regions, and/or mutations that alter expression through methods of genetic engineering that are known in the art. In addition to producing nucleic acid sequences of promoters and enhancers synthetically, promoter sequences can be produced using recombinant cloning and/or nucleic acid amplification technology, including PCR, in connection with the biological converter switches and modules

disclosed herein (see, *e.g.*, U.S. Pat. No. 4,683,202, U.S. Pat. No. 5,928,906, each incorporated herein by reference). Furthermore, it is contemplated that control sequences that direct transcription and/or expression of sequences within non-nuclear organelles such as mitochondria, chloroplasts, and the like, can be employed as well.

Inducible Promoters

[00134] As described herein, an “inducible promoter” is one that is characterized by initiating or enhancing transcriptional activity when in the presence of, influenced by, or contacted by an inducer or inducing agent. An “inducer” or “inducing agent,” as defined herein, can be endogenous, or a normally exogenous compound or protein that is administered in such a way as to be active in inducing transcriptional activity from the inducible promoter. In some embodiments, the inducer or inducing agent, *i.e.*, a chemical, a compound or a protein, can itself be the result of transcription or expression of a nucleic acid sequence (*i.e.*, an inducer can be a transcriptional repressor protein expressed by another component or module), which itself can be under the control of an inducible promoter. In some embodiments, an inducible promoter is induced in the absence of certain agents, such as a repressor. Examples of inducible promoters include but are not limited to, tetracycline, metallothionine, ecdysone, mammalian viruses (*e.g.*, the adenovirus late promoter; and the mouse mammary tumor virus long terminal repeat (MMTV-LTR)) and other steroid-responsive promoters, rapamycin responsive promoters and the like.

[00135] Inducible promoters useful in the biological circuit chemotactic converters, systems, and methods described herein are capable of functioning in both prokaryotic and eukaryotic host organisms. In some embodiments of the different aspects described herein, mammalian inducible promoters are included, although inducible promoters from other organisms, as well as synthetic promoters designed to function in a prokaryotic or eukaryotic host can be used. One important functional characteristic of the inducible promoters described herein is their ultimate inducibility by exposure to an externally applied inducer, such as an environmental inducer. Exemplary environmental inducers include exposure to heat (*i.e.*, thermal pulses or constant heat exposure), various steroidal compounds, divalent cations (including Cu^{2+} and Zn^{2+}), galactose, tetracycline, IPTG (isopropyl- β -D thiogalactoside), as well as other naturally occurring and synthetic inducing agents and gratuitous inducers.

[00136] The promoters for use in the biological circuit chemotactic converters and modules described herein encompass the inducibility of a prokaryotic or eukaryotic promoter by, in part, either of two mechanisms. In some embodiments of these aspects, the biological converter switches and their component modules comprise suitable inducible promoters that can be dependent upon transcriptional activators that, in turn, are reliant upon an environmental inducer. In other embodiments, the inducible promoters can be repressed by a transcriptional repressor, which itself is rendered inactive by an environmental inducer, such as the product of a sequence driven by another promoter. Thus, unless specified otherwise, an inducible promoter can be one that is induced by an

inducing agent that positively activates a transcriptional activator, or one which is derepressed by an inducing agent that negatively regulates a transcriptional repressor. In such embodiments of the various aspects described herein where it is required to distinguish between an activating and a repressing inducing agent, explicit distinction will be made.

[00137] Inducible promoters that are useful in the biological circuit chemotactic converters and methods of use disclosed herein include those controlled by the action of latent transcriptional activators that are subject to induction by the action of environmental inducing agents. Some non-limiting examples include the copper-inducible promoters of the yeast genes CUP1, CRS5, and SOD1 that are subject to copper-dependent activation by the yeast ACE1 transcriptional activator (see *e.g.* Strain and Culotta, 1996; Hottiger et al., 1994; Lapinskas et al., 1993; and Gralla et al., 1991). Alternatively, the copper inducible promoter of the yeast gene CTT1 (encoding cytosolic catalase T), which operates independently of the ACE1 transcriptional activator (Lapinskas et al., 1993), can be utilized. The copper concentrations required for effective induction of these genes are suitably low so as to be tolerated by most cell systems, including yeast and *Drosophila* cells. Alternatively, other naturally occurring inducible promoters can be used in the present invention including: steroid inducible gene promoters (see *e.g.* Oligino et al. (1998) *Gene Ther.* 5: 491-6); galactose inducible promoters from yeast (see *e.g.* Johnston (1987) *Microbiol Rev* 51: 458-76; Ruzzi et al. (1987) *Mol Cell Biol* 7: 991-7); and various heat shock gene promoters. Many eukaryotic transcriptional activators have been shown to function in a broad range of eukaryotic host cells, and so, for example, many of the inducible promoters identified in yeast can be adapted for use in a mammalian host cell as well. For example, a unique synthetic transcriptional induction system for mammalian cells has been developed based upon a GAL4-estrogen receptor fusion protein that induces mammalian promoters containing GAL4 binding sites (Brasemann et al. (1993) *Proc Natl Acad Sci USA* 90: 1657-61). These and other inducible promoters responsive to transcriptional activators that are dependent upon specific inducers are suitable for use with the biological circuit chemotactic converters described herein.

[00138] Inducible promoters useful in the biological circuit chemotactic converters, methods of use and systems described herein also include those that are repressed by “transcriptional repressors,” which are subject to inactivation by the action of environmental, external agents, or the product of another gene. Such inducible promoters can also be termed “repressible promoters” where it is required to distinguish between other types of promoters in a given module or component of the biological switch converters described herein. Examples include prokaryotic repressors molecules that can transcriptionally repress eukaryotic promoters that have been engineered to incorporate appropriate repressor-binding operator sequences. Preferred repressors for use in the modules and methods described herein are sensitive to inactivation by physiologically benign agent. Thus, where a lac repressor protein is used to control the expression of a promoter sequence that has been engineered to contain a lacO operator sequence, treatment of the host cell with IPTG will cause the dissociation

of the lac repressor from the engineered promoter containing a lacO operator sequence and allow transcription to occur. Similarly, where a tet repressor is used to control the expression of a promoter sequence that has been engineered to contain a tetO operator sequence, treatment of the host cell with tetracycline will cause dissociation of the tet repressor from the engineered promoter and allow transcription of the sequence downstream of the engineered promoter to occur.

[00139] An inducible promoter useful in the biological circuit chemotactic converters, methods and systems as described herein can be induced by one or more physiological conditions, such as changes in pH, temperature, radiation, osmotic pressure, saline gradients, cell surface binding, and the concentration of one or more extrinsic or intrinsic inducing agents. The extrinsic inducer or inducing agent can comprise amino acids and amino acid analogs, saccharides and polysaccharides, nucleic acids, protein transcriptional activators and repressors, cytokines, toxins, petroleum-based compounds, metal containing compounds, salts, ions, enzyme substrate analogs, hormones, and combinations thereof. In specific embodiments, the inducible promoter is activated or repressed in response to a change of an environmental condition, such as the change in concentration of a chemical, metal, temperature, radiation, nutrient or change in pH. Thus, an inducible promoter useful in the methods and systems as described herein can be a phage inducible promoter, nutrient inducible promoter, temperature inducible promoter, radiation inducible promoter, metal inducible promoter, hormone inducible promoter, steroid inducible promoter, and/or hybrids and combinations thereof.

[00140] Promoters that are inducible by ionizing radiation can be used in certain embodiments, where gene expression is induced locally in a cell by exposure to ionizing radiation, such as UV or x-rays. Radiation inducible promoters include the non-limiting examples of fos promoter, c-jun promoter or at least one CArG domain of an Egr-1 promoter. Further non-limiting examples of inducible promoters include promoters from genes such as cytochrome P450 genes, inducible heat shock protein genes, metallothionein genes, hormone-inducible genes, such as the estrogen gene promoter, and such. In further embodiments, an inducible promoter useful in the methods and systems described herein can be Zn²⁺ metallothionein promoter, metallothionein-1 promoter, human metallothionein IIA promoter, lac promoter, lacO promoter, mouse mammary tumor virus early promoter, mouse mammary tumor virus LTR promoter, triose dehydrogenase promoter, herpes simplex virus thymidine kinase promoter, simian virus 40 early promoter or retroviral myeloproliferative sarcoma virus promoter. Examples of inducible promoters also include mammalian probasin promoter, lactalbumin promoter, GRP78 promoter, or the bacterial tetracycline-inducible promoter. Other examples include phorbol ester, adenovirus E1A element, interferon, and serum inducible promoters.

[00141] Inducible promoters useful in the modules and biological circuit chemotactic converters described herein for *in vivo* uses can include those responsive to biologically compatible agents, such as those that are usually encountered in defined animal tissues. An example is the human PAI-1 promoter, which is inducible by tumor necrosis factor. Further suitable examples include

cytochrome P450 gene promoters, inducible by various toxins and other agents; heat shock protein genes, inducible by various stresses; hormone-inducible genes, such as the estrogen gene promoter, and such.

[00142] The administration or removal of an inducer or repressor as described herein results in a switch between the “on” or “off” states of the transcription of the operably linked heterologous target gene. Thus, as defined herein, the “on” state of a promoter operably linked to a nucleic acid sequence, refers to the state when the promoter is actively driving transcription of the operably linked nucleic acid sequence, *i.e.*, the linked nucleic acid sequence is expressed. Several small molecule ligands have been shown to mediate regulated gene expressions, either in tissue culture cells and/or in transgenic animal models. These include the FK1012 and rapamycin immunosuppressive drugs (Spencer et al., 1993; Magari et al., 1997), the progesterone antagonist mifepristone (RU486) (Wang, 1994; Wang et al., 1997), the tetracycline antibiotic derivatives (Gossen and Bujard, 1992; Gossen et al., 1995; Kistner et al., 1996), and the insect steroid hormone ecdysone (No et al., 1996). All of these references are herein incorporated by reference. By way of further example, Yao discloses in U.S. Pat. No. 6,444,871, which is incorporated herein by reference, prokaryotic elements associated with the tetracycline resistance (*tet*) operon, a system in which the *tet* repressor protein is fused with polypeptides known to modulate transcription in mammalian cells. The fusion protein is then directed to specific sites by the positioning of the *tet* operator sequence. For example, the *tet* repressor has been fused to a transactivator (VP16) and targeted to a *tet* operator sequence positioned upstream from the promoter of a selected gene (Gussen et al., 1992; Kim et al., 1995; Hennighausen et al., 1995). The *tet* repressor portion of the fusion protein binds to the operator thereby targeting the VP16 activator to the specific site where the induction of transcription is desired. An alternative approach has been to fuse the *tet* repressor to the KRAB repressor domain and target this protein to an operator placed several hundred base pairs upstream of a gene. Using this system, it has been found that the chimeric protein, but not the *tet* repressor alone, is capable of producing a 10 to 15-fold suppression of CMV-regulated gene expression (Deuschle et al., 1995).

[00143] One example of a repressible promoter useful in the modules and biological circuit chemotactic converters as disclosed herein is the Lac repressor (*lacR*)/operator/inducer system of *E. coli* that has been used to regulate gene expression by three different approaches: (1) prevention of transcription initiation by properly placed *lac* operators at promoter sites (Hu and Davidson, 1987; Brown et al., 1987; Figge et al., 1988; Fuerst et al., 1989; Deuschle et al., 1989; (2) blockage of transcribing RNA polymerase II during elongation by a LacR/operator complex (Deuschle et al. (1990); and (3) activation of a promoter responsive to a fusion between LacR and the activation domain of herpes simplex virus (HSV) virion protein 16 (VP16) (Labow et al., 1990; Baim et al., 1991). In one version of the Lac system, expression of *lac* operator-linked sequences is constitutively activated by a LacR-VP16 fusion protein and is turned off in the presence of isopropyl- β -D-1-thiogalactopyranoside (IPTG) (Labow et al. (1990), cited *supra*). In another version of the system, a

lacR-VP16 variant is used that binds to lac operators in the presence of IPTG, which can be enhanced by increasing the temperature of the cells (Baim et al. (1991), cited supra). Thus, in some embodiments of the aspects described herein, components of the Lac system are utilized. For example, a lac operator (LacO) can be operably linked to tissue specific promoter, and control the transcription and expression of the heterologous target gene and another repressor protein, such as the TetR. Accordingly, the expression of the heterologous target gene is inversely regulated as compared to the expression or presence of Lac repressor in the system.

[00144] Components of the tetracycline (Tc) resistance system of *E. coli* that have also been found to function in eukaryotic cells and been used to regulate gene expression can also be used in the various aspects described herein. For example, the Tet repressor (TetR), which binds to tet operator (tetO) sequences in the absence of tetracycline and represses gene transcription, has been expressed in plant cells at sufficiently high concentrations to repress transcription from a promoter containing tet operator sequences (Gatz, C. et al. (1992) *Plant J.* 2:397-404). In some embodiments described herein, the Tet repressor system is similarly utilized.

[00145] A temperature- or heat-inducible gene regulatory system can also be used in the in the biological circuit chemotactic converters and methods described herein, such as the exemplary TIGR system comprising a cold-inducible transactivator in the form of a fusion protein having a heat shock responsive regulator, rheA, fused to the VP16 transactivator (Weber et al., 2003a). The promoter responsive to this fusion thermosensor comprises a rheO element operably linked to a minimal promoter, such as the minimal version of the human cytomegalovirus immediate early promoter. At the permissive temperature of 37°C, the cold-inducible transactivator transactivates the exemplary rheO-CMVmin promoter, permitting expression of the target gene. At 41°C, the cold-inducible transactivator no longer transactivates the rheO promoter. Any such heat-inducible or -regulated promoter can be used in accordance with the modules, biological circuit chemotactic converters, and methods described herein, including but not limited to a heat-responsive element in a heat shock gene (*e.g.*, hsp20-30, hsp27, hsp40, hsp60, hsp70, and hsp90). See Easton et al. (2000) *Cell Stress Chaperones* 5(4):276-290; Csermely et al. (1998) *Pharmacol Ther* 79(2): 129-1 68; Ohtsuka & Hata (2000) *Int J Hyperthermia* 16(3):231-245; and references cited therein. Sequence similarity to heat shock proteins and heat-responsive promoter elements have also been recognized in genes initially characterized with respect to other functions, and the DNA sequences that confer heat inducibility are suitable for use in the disclosed gene therapy vectors. For example, expression of glucose-responsive genes (*e.g.*, grp94, grp78, mortalin/grp75) (Merrick et al. (1997) *Cancer Lett* 119(2): 185-1 90; Kiang et al. (1998) *FASEB J* 12(14):1571-16-579), calreticulin (Szewczenko-Pawlikowski et al. (1997) *Mol Cell Biochem* 177(1 -2): 145-1 52); clusterin (Viard et al. (1999) *J Invest Dermatol* 112(3):290-296; Michel et al. (1997) *Biochem J* 328(Pt1):45-50; Clark & Griswold (1997) *J Androl* 18(3):257-263), histocompatibility class I gene (HLA-G) (Ibrahim et al. (2000) *Cell Stress Chaperones* 5(3):207-218), and the Kunitz protease isoform of amyloid precursor protein (Shepherd et al. (2000) *Neuroscience*

99(2):31 7-325) are upregulated in response to heat. In the case of clusterin, a 14 base pair element that is sufficient for heat-inducibility has been delineated (Michel et al. (1997) *Biochem J* 328(Pt1):45-50). Similarly, a two sequence unit comprising a 10- and a 14-base pair element in the calreticulin promoter region has been shown to confer heat-inducibility (Szewczenko-Pawlikowski et al. (1997) *Mol Cell Biochem* 177(1 -2): 145-1 52).

[00146] Other inducible promoters useful in the various embodiments of the aspects described herein include the erythromycin-resistance regulon from *E.coli*, having repressible (E_{off}) and inducible (E_{on}) systems responsive to macrolide antibiotics, such as erythromycin, clarithromycin, and roxithromycin (Weber et al., 2002). The E_{off} system utilizes an erythromycin-dependent transactivator, wherein providing a macrolide antibiotic represses transgene expression. In the E_{on} system, the binding of the repressor to the operator results in repression of transgene expression. Therein, in the presence of macrolides gene expression is induced.

[00147] Fussenegger et al. (2000) describe repressible and inducible systems using a Pip (pristinamycin-induced protein) repressor encoded by the streptogramin resistance operon of *Streptomyces coelicolor*, wherein the systems are responsive to streptogramin-type antibiotics (such as, for example, pristinamycin, virginiamycin, and Synercid). The Pip DNA-binding domain is fused to a VP16 transactivation domain or to the KRAB silencing domain, for example. The presence or absence of, for example, pristinamycin, regulates the Pip_{ON} and Pip_{OFF} systems in their respective manners, as described therein.

[00148] Another example of a promoter expression system useful for the modules and biological circuit chemotactic converters described herein utilizes a quorum-sensing (referring to particular prokaryotic molecule communication systems having diffusible signal molecules that prevent binding of a repressor to an operator site, resulting in derepression of a target regulon) system. For example, Weber et al. (2003b) employ a fusion protein comprising the *Streptomyces coelicolor* quorum-sensing receptor to a transactivating domain that regulates a chimeric promoter having a respective operator that the fusion protein binds. The expression is fine-tuned with non-toxic butyrolactones, such as SCB1 and MP133.

[00149] In some embodiments, multiregulated, multigene gene expression systems that are functionally compatible with one another can be utilized in the aspects described herein (see, for example, Kramer et al. (2003)). For example, in Weber et al. (2002), the macrolide-responsive erythromycin resistance regulon system is used in conjunction with a streptogramin (PIP)-regulated and tetracycline-regulated expression systems.

[00150] Other promoters responsive to non-heat stimuli can also be used. For example, the mortalin promoter is induced by low doses of ionizing radiation (Sadekova (1997) *Int J Radiat Biol* 72(6):653-660), the hsp27 promoter is activated by 17- β -estradiol and estrogen receptor agonists (Porter et al. (2001) *J Mol Endocrinol* 26(1):31-42), the HLA-G promoter is induced by arsenite, hsp promoters can be activated by photodynamic therapy (Luna et al. (2000) *Cancer Res* 60(6): 1637-1

644). A suitable promoter can incorporate factors such as tissue-specific activation. For example, hsp70 is transcriptionally impaired in stressed neuroblastoma cells (Drujan & De Maio (1999) 12(6):443-448) and the mortalin promoter is up-regulated in human brain tumors (Takano et al. (1997) Exp Cell Res 237(1):38-45). A promoter employed in methods of the present invention can show selective up-regulation in tumor cells as described, for example, for mortalin (Takano et al. (1997) Exp Cell Res 237(1):38-45), hsp27 and calreticulin (Szewczenko-Pawlikowski et al. (1997) Mol Cell Biochem 177(1-2): 145-152; Yu et al. (2000) Electrophoresis 21(14):3058-3068), grp94 and grp78 (Gazit et al. (1999) Breast Cancer Res Treat 54(2): 135-146), and hsp27, hsp70, hsp73, and hsp90 (Cardillo et al. (2000) Anticancer Res 20(6B):4579-4583; Strik et al. (2000) Anticancer Res 20(6B):4457-4552).

[00151] As described herein, the promoters in the modular components of the biological circuit chemotactic converters described herein, such as genetic toggle switches, can drive expression of an operably linked recombinase, repressor, or output product, thus regulating expression and consequent activity of said recombinase, repressor, or output product. In some embodiments of the various aspects described herein, promoter sequences are added to the modules and biological circuit chemotactic converters in order to enumerate and input physiological events and stimuli, such as activation of gene networks or exposure to nutrients, toxins, metabolites, or any environmental exposure.

[00152] In some embodiments of the various aspects described herein, the promoter sequence that is added to a module or a biological circuit chemotactic converter is an inducible promoter that allows control of the module or biological circuit chemotactic converter using one or more chemical inducers.

[00153] In some embodiments, the inducible promoter comprises an Anhydrotetracycline (aTc)-inducible promoter as provided in PLtetO-1 (Pubmed Nucleotide# U66309) with the sequence comprising

GCATGCTCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCAC
ATCAGCAGGACGCACTGACCAGGA (**SEQ ID NO: 1**).

[00154] In some embodiments, the inducible promoter is an arabinose-inducible promoter P_{BAD} comprising the sequence

AAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCCTGCGTCTTTTACTGGCTCTT
CTCGCTAACCAACCGGTAACCCCGCTTATTAAGCATTCTGTAACAAAGCGGGACCA
AAGCCATGACAAAAACGCGTAACAAAAGTGTCTATAATCACGGCAGAAAAGTCCACATT
GATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTATCCATAAGATTAGCGG
ATCCTACCTGACGCTTTTATCGCAACTCTCTACTGTTTCTCCATA (**SEQ ID NO: 2**).

[00155] In some embodiments, the inducible promoter is an isopropyl β -D-1-thiogalactopyranoside (IPTG) inducible promoter. In one embodiment, the IPTG-inducible promoter comprises the P_{TAC} sequence found in the vector encoded by PubMed Accession ID #EU546824. In

one embodiment, the IPTG-inducible promoter sequence comprises the P_{Trc-2} sequence
CCATCGAATGGCTGAAATGAGCTGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGA
ATTGTGAGCGGATAACAATTTACACAGGA (**SEQ ID NO: 3**).

[00156] In some embodiments, the IPTG-inducible promoter comprises the P_{Trc-2} sequence found in the vector encoded by PubMed Accession ID #EU546816.

[00157] In some embodiments, the IPTG-inducible promoter comprises the P_{LlacO-1} sequence:ATAAATGTGAGCGGATAACATTGACATTGTGAGCGGATAACAAGATACTGAGC
ACTCAGCAGGACGCACTGACC (**SEQ ID NO: 4**).

[00158] In some embodiments, the IPTG-inducible promoter comprises the P_{AlacO-1} sequence
AAAATTTATCAAAAAGAGTGTTGACTTGTGAGCGGATAACAATGATACTTAGATTCAATT
GTGAGCGGATAACAATTTACACACA (**SEQ ID NO: 5**).

[00159] In some embodiments, the IPTG-inducible promoter comprises the P_{lac/ara-1} sequence
CATAGCATTTTTATCCATAAGATTAGCGGATCCTAAGCTTTACAATTGTGAGCGCTCACA
ATTATGATAGATTCAATTGTGAGCGGATAACAATTTACACACA (**SEQ ID NO: 6**).

[00160] In some embodiments, the inducible promoter sequence comprises the P_{LsIcon} sequence of
GCATGCACAGATAACCATCTGCGGTGATAAATTATCTCTGGCGGTGTTGACATAAATACC
ACTGGCGGTtATAaTGAGCACATCAGCAGG//GTATGCAAAGGA (**SEQ ID NO: 7**), **SEQ ID NO: 1009**, or **SEQ ID NO: 1010**.

[00161] Other non-limiting examples of constitutive and inducible promoters that are useful for the modules and biological circuit chemotactic converters described herein can be found in **SEQ ID NO: 41 – SEQ ID NO: 843**, and **SEQ ID NO: 1005 – SEQ ID NO: 1008**.

[00162] A variety of constitutive promoter sequences are available for those embodiments of the modules and biological circuit chemotactic converters described herein where constitutive expression of an operably linked sequence is required. Exemplary constitutive *E. coli* σ^{70} promoter sequences can be found in **SEQ ID NO: 41 – SEQ ID NO: 125** and **SEQ ID NO: 1005 – SEQ ID NO: 1008**. Exemplary constitutive *E. coli* σ^S promoter sequences can be found in **SEQ ID NO: 126 – SEQ ID NO: 127**. Exemplary constitutive *B. subtilis* σ^{32} promoter sequences can be found in **SEQ ID NO: 128**. Exemplary constitutive *B. subtilis* σ^A promoter sequences can be found in **SEQ ID NO: 129 – SEQ ID NO: 130**. Exemplary constitutive *B. subtilis* σ^B promoter sequences can be found in **SEQ ID NO: 131 – SEQ ID NO: 133**. Exemplary constitutive promoter sequences from miscellaneous prokaryotes can be found in **SEQ ID NO: 134 – SEQ ID NO: 135**. Exemplary constitutive promoter sequences from bacteriophage T7 can be found in **SEQ ID NO: 136 – SEQ ID NO: 150**. An exemplary constitutive promoter sequence from bacteriophage SP6 can be found in **SEQ ID NO: 151**. Exemplary constitutive promoter sequences from yeast can be found in **SEQ ID NO: 152 – SEQ ID NO: 164**. Exemplary constitutive promoter sequences from miscellaneous eukaryotes can be found in **SEQ ID NO: 165 – SEQ ID NO: 166**.

[00163] A variety of inducible (positive and negative) and otherwise regulateable promoter sequences are available for those embodiments of the modules and biological circuit chemotactic converters described herein where inducible (repressible or activatable) expression of an operably linked sequence is required. Exemplary cell signaling promoter sequences can be found in **SEQ ID NO: 167 – SEQ ID NO: 198**. Exemplary metal inducible promoter sequences can be found in **SEQ ID NO: 199 – SEQ ID NO: 204**. Exemplary T7 promoter sequences can be found in **SEQ ID NO: 205 – SEQ ID NO: 224**. Exemplary stress kit promoter sequences can be found in **SEQ ID NO: 225– SEQ ID NO: 241**. Exemplary logic promoter sequences can be found in **SEQ ID NO: 242 – SEQ ID NO: 319**. Exemplary positively regulated *E. coli* σ^{70} promoter sequences can be found in **SEQ ID NO: 320 – SEQ ID NO: 424**. An exemplary positively regulated *E. coli* σ^S promoter sequence can be found in **SEQ ID NO: 425**. An exemplary positively regulated *E. coli* σ^{32} promoter sequence can be found in **SEQ ID NO: 426**. An exemplary positively regulated *E. coli* σ^{54} promoter sequence can be found in **SEQ ID NO: 427**. Exemplary positively regulated *B. subtilis* σ^A promoter sequences can be found in **SEQ ID NO: 428 – SEQ ID NO: 437**. Exemplary miscellaneous prokaryotic inducible promoter sequences can be found in **SEQ ID NO: 438 – SEQ ID NO: 443**. Exemplary yeast positive (activatable) promoter sequences can be found in **SEQ ID NO: 444 – SEQ ID NO: 459**. Exemplary eukaryotic positive (activatable) promoter sequences can be found in **SEQ ID NO: 460 – SEQ ID NO: 464**. Exemplary negatively regulated (repressible) *E. coli* σ^{70} promoter sequences can be found in **SEQ ID NO: 465 – SEQ ID NO: 683**. Exemplary negatively regulated (repressible) *E. coli* σ^S promoter sequences can be found in **SEQ ID NO: 684– SEQ ID NO: 688**. Exemplary negatively regulated (repressible) *E. coli* σ^{32} promoter sequences can be found in **SEQ ID NO: 689– SEQ ID NO: 692**. An exemplary negatively regulated (repressible) *E. coli* σ^{54} promoter sequence can be found in **SEQ ID NO: 693**. Exemplary negatively regulated (repressible) *B. subtilis* σ^A promoter sequences can be found in **SEQ ID NO: 694 – SEQ ID NO: 696**. Exemplary T7 negatively regulated (repressible) promoter sequences can be found in **SEQ ID NO: 697 – SEQ ID NO: 700**. Exemplary yeast negatively regulated (repressible) promoter sequences can be found in **SEQ ID NO: 701 – SEQ ID NO: 703**. Exemplary eukaryotic negatively regulated (repressible) promoter sequences can be found in **SEQ ID NO: 704 – SEQ ID NO: 710**.

[00164] In addition, a variety of combination (positive and negative) and otherwise regulateable promoter sequences are available for those embodiments of the modules and biological circuit chemotactic converters described herein where both repressible and activateable expression of an operably linked sequence is required. Exemplary combination activateable and repressible *E. coli* promoter sequences can be found in **SEQ ID NO: 711 – SEQ ID NO: 833**. An exemplary combination activateable and repressible miscellaneous prokaryotic promoter sequence can be found in **SEQ ID NO: 834**. Exemplary combination activateable and repressible miscellaneous yeast promoter sequences can be found in **SEQ ID NO: 835 – SEQ ID NO: 839**. Exemplary combination

activateable and repressible miscellaneous eukaryotic promoter sequences can be found in **SEQ ID NO: 840 – SEQ ID NO: 843**.

Recombinases and Recombination Recognition Sequences

[00165] In some aspects, sequences encoding recombinases and recombination recognition sequences are used in the various modules and components of the biological circuit chemotactic converters described herein. Sequences encoding recombinases and recombination recognition sequences are used, in part, to provide a genetically inheritable switch in the state of a module, or any of the biological circuit chemotactic converters described herein. The recombinases in the various modules and components of the biological circuit chemotactic converters s described herein, are preferably expressed between their cognate recognition sites, *i.e.*, recombinase recognition site_{for} - recombinase - recombinase recognition site_{rev}. As a result, upon recombinase expression following activation of an upstream promoter, the recombinase causes a single inversion of the nucleic acid sequence between the cognate recognition sites, including its own sequence (*i.e.*, recombinase recognition site_{for} - inverted recombinase - recombinase recognition site_{rev}). Any further transcription from the upstream promoter yields a long antisense RNA molecule of the recombinase gene rather than sense RNA, and therefore no further recombinase protein is produced. Thus, the inversion event is discrete and stable and does not result in a mixture of inverted and non-inverted states.

[00166] The advantages of the use of recombinases that mediate site-specific inversion for use in the various aspects of the biological circuit chemotactic converters described herein are the binary dynamics, the sensitivity of the output, the efficiency of DNA usage, and the persistence of the DNA modification. A “recombinase,” as defined herein, is a site-specific enzyme that recognizes short DNA sequences, which are typically between about 30 bp and 40 bp, and mediates the recombination between these elements that results in the excision, integration, inversion, or exchange of DNA fragments.

[00167] Recombinases can be classified into two distinct families, the integrase and invertase/resolvase families, based on distinct biochemical properties. Members of the integrase family cleave one strand of each of the two DNA molecules involved, then exchange this strand, and subsequently cleave the second DNA strand. Integrase family recombinases use a conserved tyrosine residue to establish a transient covalent bond between the recombinase and the target DNA. Members of the invertase/resolvase family of recombinases cleave all 4 DNA strands and then exchange them, and initiate DNA cleavage by utilizing a serine residue as the catalytic residue. Recombinases have been used for numerous standard biological applications, including the creation of gene knockouts and the solving of sorting problems (N. J. Kilby, Trends Genet 9, 413 (Dec, 1993); K. A. Haynes, J Biol Eng 2, 8 (2008); T. S. Ham, Biotechnol Bioeng 94, 1 (2006); K. A. Datsenko, Proc Natl Acad Sci U S A 97, 6640 (2000)).

[00168] Inversion recombination happens between two short inverted repeated DNA sequences, typically less than 30 bp long. The recombinases bind to these inverted repeated

sequences, which are specific to each recombinase, and are defined herein as “recombinase recognition sequences” or “recombinase recognition sites.” A DNA loop formation, assisted by DNA bending proteins, brings the two repeat sequences together, at which point DNA cleavage and ligation occur. This reaction is ATP independent and requires supercoiled DNA. The end result of such an inversion recombination event is that the stretch of DNA between the repeated site inverts, *i.e.*, the stretch of DNA reverses orientation, such that what was the coding strand is now the non-coding strand and vice versa. In such reactions, the DNA is conserved with no net gain or no loss of DNA.

[00169] The recombinases provided herein are not meant to be an exclusive listing. Other examples of recombinases that are useful in the modules and any of the biological circuit chemotactic converters described herein are known to those of skill in the art, and furthermore, any new recombinase that is discovered or generated can be used in the embodiments described herein.

[00170] In some embodiments of the aspects described herein, the recombinase comprises the sequence of Cre recombinase of Pubmed Gene ID #2777477 (**SEQ ID NO: 1002**), and the corresponding loxP recombinase recognition sequences comprise the sequences of **SEQ ID NO: 8** and **SEQ ID NO: 9**.

[00171] In some embodiments of the aspects described herein,, the recombinase is Flp recombinase comprising the sequences of GenBank ID U46493 or NC_001398 (**SEQ ID NO: 1003**). In other embodiments, the recombinase is an enhanced Flp recombinase that comprises the sequence of **SEQ ID NO: 10**. The corresponding recombinase recognition sequences for the Flp and enhanced Flp recombinases comprise FRT sites with sequences comprising **SEQ ID NO: 11**. In some embodiments, minimal FRT recombinase recognition sites are used, comprising the sequence of **SEQ ID NO: 12**.

[00172] In some embodiments, the recombinase is R recombinase comprising the sequence of GenBank ID # X02398 (**SEQ ID NO: 1004**) and the corresponding recombinase recognition sequence comprises **SEQ ID NO: 13**.

[00173] In some embodiments, the recombinase comprises the bidirectional FimB recombinase of GeneID: 948832 (**SEQ ID NO: 16**) and the corresponding recombinase recognition sequences comprise **SEQ ID NO: 14** and **SEQ ID NO: 15**.

[00174] In some embodiments, the recombinase is the unidirectional FimE recombinase of GeneID: 948836 (**SEQ ID NO: 17**) and the corresponding recombinase recognition sequences comprise **SEQ ID NO: 14** and **SEQ ID NO: 15**.

[00175] In some embodiments, the recombinase is an Int recombinase. In some embodiments, the Int recombinase comprises a sequence that encodes for an Int recombinase selected from the group consisting of intE, HP1 Int, and HK022 Int.

[00176] In some embodiments, the recombinase is the XerC/XerD recombinase comprising the sequence of GeneID: 5387246 (**SEQ ID NO: 18**) and the corresponding recombinase recognition sequences comprise cer and dif.

[00177] In some embodiments, the recombinase is *Salmonella* Hin recombinase comprising the sequence of GeneID: 1254295 (**SEQ ID NO: 19**) and the corresponding recombinase recognition sequences comprise hixL and hixR.

[00178] In regard to the use of the 'Cre recombinase' and corresponding 'Lox' recombination recognition sequences, or the "Cre-Lox system," in the modules and biological circuit chemotactic converters described herein, the Cre protein has been purified to homogeneity (Abremski et al. (1984) J. Mol. Biol. 259:1509) and the *cre* gene has been cloned and expressed in a variety of host cells (Abremski et al. (1983)). Purified Cre protein is available from a number of suppliers (*e.g.*, Stratagene, Novagen and New England Nuclear/Du Pont). Cre catalyzes the cleavage of the lox site within the spacer region and creates a six base-pair staggered cut (Hoess and Abremski (1985) J. Mol. Biol. 181:351). The two 13 bp inverted repeat domains of the lox site represent binding sites for the Cre protein. If two lox sites differ in their spacer regions in such a manner that the overhanging ends of the cleaved DNA cannot reanneal with one another, Cre cannot efficiently catalyze a recombination event using the two different lox sites. For example, it has been reported that Cre cannot recombine (at least not efficiently) a loxP site and a loxP511 site; these two lox sites differ in the spacer region. Two lox sites which differ due to variations in the binding sites (*i.e.*, the 13 bp inverted repeats) may be recombined by Cre provided that Cre can bind to each of the variant binding sites; the efficiency of the reaction between two different lox sites (varying in the binding sites) may be less efficient than between two lox sites having the same sequence (the efficiency will depend on the degree and the location of the variations in the binding sites). For example, the loxC2 site can be efficiently recombined with the loxP site; these two lox sites differ by a single nucleotide in the left binding site.

[00179] In addition to the foregoing examples of sequences that the Cre protein recognizes, Cre also recognizes a number of variant or mutant lox sites (variant relative to the loxP sequence), including the loxB, loxL, loxR, loxA86, and lox.DELTA.117 sites which are found in the *E. coli* chromosome (Hoess et al. (1982)). Other variant lox sites include loxP511 (**SEQ ID NO: 20**); Hoess et al. (1986), *supra*), loxC2 (**SEQ ID NO: 21**); U.S. Pat. No. 4,959,317), lox66 (**SEQ ID NO: 22**), lox 71 (**SEQ ID NO: 23**), and lox BBa_J61046 (**SEQ ID NO: 24**).

[00180] Other alternative site-specific recombinases include: 1) the FLP recombinase of the 2pi plasmid of *Saccharomyces cerevisiae* (Cox (1983) Proc. Natl. Acad. Sci. USA 80:4223) which recognize the *frt* site which, like the loxP site, comprises two 13 bp inverted repeats separated by an 8 bp spacer (**SEQ ID NO: 25**). The FLP gene has been cloned and expressed in *E. coli* (Cox, *supra*) and in mammalian cells (PCT International Patent Application PCT/US92/01899, Publication No.: WO 92/15694, the disclosure of which is herein incorporated by reference) and has been purified (Meyer-Lean et al. (1987) Nucleic Acids Res. 15:6469; Babineau et al (1985) J. Biol. Chem. 260:12313; Gronostajski and Sadowski (1985) J. Biol. Chem. 260:12328); 2) the integrase of *Streptomyces* phage .PHI.C31 that carries out efficient recombination between the attP site of the phage genome and the attB site of the host chromosome (Groth et al., 2000 Proc. Natl. Acad. Sci.

USA, 97: 5995); 3) the Int recombinase of bacteriophage lambda (lambda-int/attP) (with or without Xis) which recognizes att sites (Weisberg et al. In: Lambda II, supra, pp. 211-250); 4) the xerC and xerD recombinases of *E. coli* which together form a recombinase that recognizes the 28 bp dif site (Leslie and Sherratt (1995) EMBO J. 14:1561); 5) the Int protein from the conjugative transposon Tn916 (Lu and Churchward (1994) EMBO J. 13:1541); 6) TpnI and the β -lactamase transposons (Levesque (1990) J. Bacteriol. 172:3745); 7) the Tn3 resolvase (Flanagan et al. (1989) J. Mol. Biol. 206:295 and Stark et al. (1989) Cell 58:779); 8) the SpoIVC recombinase of *Bacillus subtilis* (Sato et al. J. Bacteriol. 172:1092); 9) the Hin recombinase (Galsgow et al. (1989) J. Biol. Chem. 264:10072); 10) the Cin recombinase (Haft et al. (1988) EMBO J. 7:3991); 11) the immunoglobulin recombinases (Malynn et al. Cell (1988) 54:453); and 12) the FIMB and FIME recombinases (Blomfield et al., 1997 Mol. Microbiol. 23:705).

[00181] In the natural *Salmonella* system, the Hin DNA recombinase (BBa_J31000, BBa_J31001) catalyzes an inversion reaction that regulates the expression of alternative flagellin genes by switching the orientation of a promoter located on a 1 kb invertible DNA segment. The asymmetrical palindromic sequences *hixL* and *hixR* flank the invertible DNA segment and serve as the recognition sites for cleavage and strand exchange. A ~70 bp cis-acting *recombinational enhancer (RE)* increases efficiency of protein-DNA complex formation. In some embodiments, rather than *hixL* and *hixR*, *hixC* (BBa_J44000), a composite 26 bp symmetrical *hix* site that shows higher binding affinity for Hin and a 16-fold slower inversion rate than wild type sites *hixL* and *hixR* can be used. In addition, a modified Hin/*hix* DNA recombination system can be used *in vivo* to manipulate at least two adjacent *hixC*-flanked DNA segments. Hin recombinase fused to a C-terminus LVA degradation tag (BBa_J31001) and *hixC* (BBa_J44000) are sufficient for DNA inversion activity. Exemplary sequences for the recombinational enhancer and modified Hin recombinase recognition sequences include Recombinational Enhancer (RE) for Hin/Hix inverting (**SEQ ID NO: 844**) and *hixC* binding site for *Salmonella typhimurium* Hin recombinase (**SEQ ID NO: 845**).

[00182] Bacteriophage λ has long served as a model system for studies of regulated site-specific recombination. In conditions favorable for bacterial growth, the phage genome is inserted into the *Escherichia coli* genome by an 'integrative' recombination reaction, which takes place between DNA attachment sites called *attP* and *attB* in the phage and bacterial genomes, respectively. As a result, the integrated λ DNA is bounded by hybrid attachment sites, termed *attL* and *attR*. In response to the physiological state of the bacterial host or to DNA damage, λ phage DNA excises itself from the host chromosome. This excision reaction recombines *attL* with *attR* to precisely restore the *attP* and *attB* sites on the circular λ and *E. coli* DNAs. The phage-encoded λ integrase protein (Int), a tyrosine recombinase, splices together bacterial and phage attachment sites. Int is required for both integration and excision of the λ prophage.

[00183] λ recombination has a strong directional bias in response to environmental conditions. Accessory factors, whose expression levels change in response to host physiology, control the action of Int and determine whether the phage genome will remain integrated or be excised. Int has two DNA-binding domains: a C-terminal domain, consisting of a catalytic domain and a core-binding (CB) domain, that interacts with the core recombining sites and an N-terminal domain (N-domain) that recognizes the regulatory arm DNA sites. The heterobivalent Int molecules bridge distant core and arm sites with the help of accessory proteins, such as integration host factor (IHF), which bend the DNA at intervening sites, and appose arm and core sequences for interaction with the Int recombinase. Five arm DNA sites in the regions flanking the core of *attP* are differentially occupied during integration and excision reactions. The integration products *attL* and *attR* cannot revert back to *attP* and *attB* without assistance from the phage-encoded factor Xis, which bends DNA on its own or in combination with the host-encoded factor Fis. Xis also inhibits integration, and prevents the *attP* and *attB* products of excision from reverting the *attP* and *attB* products of excision from reverting to *attL* and *attR*. Because the cellular levels of IHF and Fis proteins respond to growth conditions, these host-encoded factors have been proposed as the master signals for integration and excision.

[00184] Additional exemplary λ recombination recognition sequences and recombinases for the biological converter switches described herein include, but are not limited to, Lambda attB (SEQ ID NO: 846), Lambda attP (SEQ ID NO: 847), attR2 recombination site (SEQ ID NO: 848), and attR2 recombination site-reverse orientation (SEQ ID NO: 849) as λ recombination recognition sequences, and Xis lambda (excisionase from *E. coli* phage lambda (removes prophage from host genome)); Int lambda (integrase from *E. coli* phage lambda); Xis (Xis from bacteriophage lambda, assembly standard 21); a~xis (the bacteriophage lambda xis gene ready to have rbs attached and stop codon; and xis, from bacteriophage lambda; assembly standard 21.

[00185] Bacteriophage P22 is a lambdoid phage which infects *Salmonella typhimurium*. P22 can integrate into and excise out of its host chromosome via site-specific recombination. Both integration and excision reactions require the phage-encoded *int* gene, and excision is dependent on the *xis* gene as well.

[00186] P22 Int is a member of the λ integrase family. The Int proteins of λ and P22 are composed of two domains. The catalytic domain binds to the core region of the phage recombination site, *attP*, where the actual recombination reactions occur. The smaller amino-terminal domain binds to arm-type sequences which are located on either side of the core within the attP. The active components of λ integrative and excisive recombination are nucleosome-like structures, called intasomes, in which DNA is folded around several molecules of Int and integration host factor (IHF). It has been demonstrated that one monomer of λ integrase can simultaneously occupy both a core-type binding site and an arm-type binding site. Formation of these bridges is facilitated by IHF, which binds to specific sequences and imparts a substantial bend to the DNA.

[00187] The *attP* regions of P22 and λ are also similar in that both contain arm regions, known as the P and P' arms, which contain Int arm-type binding sites and IHF binding sites. However, the arrangement, spacing, and orientation of the Int and IHF binding sites are distinct. The *attP* region of λ contains two Int arm-type binding sites on the P arm and three on the P' arm. The P arm contains two IHF binding sites, and the P' arm contains a single site. The *attP* region of P22 contains three Int arm-type binding sites on the P arm and two sites on the P' arm. In addition, IHF binding sites, called H and H', are located on each arm of the P22 *attP*. Leong *et al.* showed that the *Escherichia coli* IHF can recognize and bind to these P22 IHF binding sites *in vitro*. It was also shown that the maximum amount of P22 integrative recombination occurred in the presence of *E. coli* IHF *in vitro*, whereas in its absence, recombination was detectable but depressed. However, the requirement for IHF or other possible accessory proteins during P22 site-specific recombination *in vivo* has not been tested. In this study, we assessed the role of IHF in P22 integration and excision *in vivo*.

[00188] Although the *attP* region of P22 contains strong IHF binding sites, *in vivo* measurements of integration and excision frequencies showed that infecting P22 phages can perform site-specific recombination to its maximum efficiency in the absence of IHF. In addition, a plasmid integration assay showed that integrative recombination occurs equally well in wild-type and *ihfA* mutant cells. P22 integrative recombination is also efficient in *Escherichia coli* in the absence of functional IHF. Additional exemplary recombination recognition sequences and recombinases for the biological converter switches described herein include, but are not limited to, recombination recognition sequences P22 "attB", reverse complement (**SEQ ID NO: 850**) and P22 "attP" (**SEQ ID NO: 851**), and recombinases Xis P22 (excisionase from *E. coli* phage P22 (removes prophage from host genome)), and Int P22 (integrase from *E. coli* phage P22).

[00189] The FLP system of the yeast 2mm plasmid is one of the most attractive for genomic manipulation because of its efficiency, simplicity, and demonstrated *in vivo* activity in a wide range of organisms. The Flp system has been used to construct specific genomic deletions and gene duplications, study gene function, promote chromosomal translocations, promote site-specific chromosome cleavage, and facilitate the construction of genomic libraries in organisms including bacteria, yeast, insects, plants, mice, and humans. Site-specific recombination catalyzed by the FLP recombinase occurs readily in bacterial cells.

[00190] The yeast FLP system has been studied intensively. The only requirements for FLP recombination are the FLP protein and the FLP recombination target (FRT) sites on the DNA substrates. The minimal functional FRT site contains only 34 bp. The FLP protein can promote both inter- and intramolecular recombination. Exemplary recombination recognition sequences for use with the yeast FLP system include, but are not limited to, FRT (**SEQ ID NO: 852**) and [FRT] recombination site for flp recombinase in BBb (**SEQ ID NO: 853**).

[00191] The separation and segregation of newly replicated *E. coli* circular chromosomes can also be prevented by the formation of circular chromosome dimers, which can arise during crossing

over by homologous recombination. In *E. coli*, these dimers, which arise about once every six generations, are resolved to monomers by the action of the FtsK–XerCD–*dif* chromosome dimer resolution machinery. Two site-specific recombinases of the tyrosine recombinase family, XerCD, act at a 28 bp recombination site, *dif*, located in the replication terminus region of the *E. coli* chromosome to remove the crossover introduced by dimer formation, thereby converting dimers to monomers. A complete dimer resolution reaction during recombination at *dif* requires the action of the C-terminal domain of FtsK (FtsK_C). FtsK is a multifunctional protein whose N-terminal domain acts in cell division, while the C-terminal domain functions in chromosome segregation. Therefore, FtsK is well suited to coordinate chromosome segregation and cell division. A purified protein, FtsK_{50C}, containing a functional C-terminal domain, can translocate DNA in an ATP-dependent manner and activate Xer recombination at the recombination site *dif*, thereby reconstituting *in vitro* the expected *in vivo* activities of the C-terminal domain of the complete FtsK protein. Additional exemplary recombination recognition sequences for use with the XerCD system are include, but are not limited to *dif* site with forward orientation (SEQ ID NO: 854) and *dif* site with reverse orientation (SEQ ID NO: 855).

[00192] The *fim* switch (*fimS*) comprises a 314 bp DNA element that can be inverted by site-specific recombinases FimB and FimE. In the natural system, *fimSc* contains a promoter, that when switched to the on orientation, drives transcription of the *fim* operon. The *fim* operon is needed for export and structural assembly of type 1 fimbriae. FimB and FimE, required to invert *fimS*, are members of the λ integrase family of site-specific recombinases. Recombination of *fimS* is distinct from the related Xer-mediated recombination in that the recombinases act independently to invert *fimS*. Each inverted repeat (IR) is flanked by overlapping FimB and FimE binding sites, and following occupancy of these sites they recombine the switch within the IR sequence. As for λ phage chromosomal integration and excision, *fim* recombination also requires accessory proteins, specifically integration host factor (IHF) and the leucine-responsive regulatory protein (Lrp). These proteins are believed to contribute to the overall architecture of the *fim* switch that facilitates synapse of the 9 bp IRs.

[00193] FimB catalyses inversion in both directions, although with a slight bias for the off-to-on orientation, while FimE predominantly catalyses on-to-off inversion. Control of FimE expression is important in bringing about its orientation bias; as the *fim* switch is located at the end of *fimE*, the orientation of *fimS* determines the length and 3' sequence of the *fimE* transcript. As a consequence, *fimE* mRNA is likely to be subject to more rapid 3' to 5' degradation when the switch is in the off orientation than when it is in the on orientation. In addition, FimE preferentially binds to *fimS* in the on orientation, as has been demonstrated *in vitro* and *in vivo*, which adds to the directional bias. A further difference between FimB and FimE is that FimB inversion frequencies are markedly lower than those exhibited by FimE, both *in vitro* and *in vivo*. Additional exemplary

recombination recognition sequences and recombinases for use with the FimB and FimE system include, but are not limited to, fimE IRR (**SEQ ID NO: 856**) and fimE IRL (**SEQ ID NO: 857**).

Ribosome Binding Sites

[00194] Ribosome binding sites (RBS) are sequences that promote efficient and accurate translation of mRNAs for protein synthesis, and are also provided for use in the modules and biological circuit chemotactic converters described herein to enable modulation of the efficiency and rates of synthesis of the proteins encoded by the converters, such as recombinases and repressors. An RBS affects the translation rate of an open reading frame in two main ways - i) the rate at which ribosomes are recruited to the mRNA and initiate translation is dependent on the sequence of the RBS, and ii) the RBS can also affect the stability of the mRNA, thereby affecting the number of proteins made over the lifetime of the mRNA. Accordingly, one or more ribosome binding site sequences (RBS) can be added to the modules and biological circuit chemotactic converters described herein to control expression of proteins, such as recombinases or protein output products.

[00195] Translation initiation in prokaryotes is a complex process involving the ribosome, the mRNA, and several other proteins, such as initiation factors, as described in Laursen BS, et al., *Microbiol Mol Biol Rev* 2005 Mar; 69(1) 101-23. Translation initiation can be broken down into two major steps - i) binding of the ribosome and associated factors to the mRNA, and ii) conversion of the bound ribosome into a translating ribosome lengthening processing along the mRNA. The rate of the first step can be increased by making the RBS highly complementary to the free end of the 16s rRNA and by ensuring that the start codon is AUG. The rate of ribosome binding can also be increased by ensuring that there is minimal secondary structure in the neighborhood of the RBS. Since binding between the RBS and the ribosome is mediated by base-pairing interactions, competition for the RBS from other sequences on the mRNA, can reduce the rate of ribosome binding. The rate of the second step in translation initiation, conversion of the bound ribosome into an initiation complex is dependent on the spacing between the RBS and the start codon being optimal (5-6bp).

[00196] Thus, a "ribosome binding site" ("RBS"), as defined herein, is a segment of the 5' (upstream) part of an mRNA molecule that binds to the ribosome to position the message correctly for the initiation of translation. The RBS controls the accuracy and efficiency with which the translation of mRNA begins. In prokaryotes (such as *E. coli*) the RBS typically lies about 7 nucleotides upstream from the start codon (*i.e.*, the first AUG). The sequence itself in general is called the "Shine-Dalgarno" sequence after its discoverers, regardless of the exact identity of the bases. Strong Shine-Dalgarno sequences are rich in purines (A's, G's), and the "Shine-Dalgarno consensus" sequence -- derived statistically from lining up many well-characterized strong ribosome binding sites -- has the sequence AGGAGG. The complementary sequence (CCUCCU) occurs at the 3'-end of the structural RNA ("16S") of the small ribosomal subunit and it base-pairs with the Shine-Dalgarno sequence in the mRNA to facilitate proper initiation of protein synthesis. In some embodiments of the aspects

described herein, a ribosome binding site (RBS) is added to a biological circuit chemotactic converter to regulate expression of a recombinase or other protein encoded by the converter,

[00197] For protein synthesis in eukaryotes and eukaryotic cells, the 5' end of the mRNA has a modified chemical structure ("cap") recognized by the ribosome, which then binds the mRNA and moves along it ("scans") until it finds the first AUG codon. A characteristic pattern of bases (called a "Kozak sequence") is sometimes found around that codon and assists in positioning the mRNA correctly in a manner reminiscent of the Shine-Dalgarno sequence, but does not involve base pairing with the ribosomal RNA.

[00198] RBSs can include only a portion of the Shine-Dalgarno sequence. When looking at the spacing between the RBS and the start codon, the aligned spacing rather than just the absolute spacing is important. In essence, if only a portion of the Shine-Dalgarno sequence is included in the RBS, the spacing that matters is between wherever the center of the full Shine-Dalgarno sequence would be and the start codon rather than between the included portion of the Shine-Dalgarno sequence and the start codon.

[00199] While the Shine-Dalgarno portion of the RBS is critical to the strength of the RBS, the sequence upstream of the Shine-Dalgarno sequence is also important. One of the ribosomal proteins, S1, is known to bind to adenine bases upstream from the Shine-Dalgarno sequence. As a result, in some embodiments of the modules and biological circuit chemotactic converters described herein, an RBS can be made stronger by adding more adenines to the sequence upstream of the RBS. A promoter may add some bases onto the start of the mRNA that may affect the strength of the RBS by affecting S1 binding.

[00200] In addition, the degree of secondary structure can affect the translation initiation rate. This fact can be used to produce regulated translation initiation rates, as described in Isaacs FJ *et al.*, Nat Biotechnol 2004 Jul; 22(7) 841-7.

[00201] In addition to affecting the translation rate per unit time, an RBS can affect the level of protein synthesis in a second way. That is because the stability of the mRNA affects the steady state level of mRNA, *i.e.*, a stable mRNA will have a higher steady state level than an unstable mRNA that is being produced as an identical rate. Since the primary sequence and the secondary structure of an RBS (for example, the RBS could introduce an RNase site) can affect the stability of the mRNA, the RBS can affect the amount of mRNA and hence the amount of protein that is synthesized.

[00202] A "regulated RBS" is an RBS for which the binding affinity of the RBS and the ribosome can be controlled, thereby changing the RBS strength. One strategy for regulating the strength of prokaryotic RBSs is to control the accessibility of the RBS to the ribosome. By occluding the RBS in RNA secondary structure, translation initiation can be significantly reduced. By contrast, by reducing secondary structure and revealing the RBS, translation initiation rate can be increased. Isaacs and coworkers engineered mRNA sequences with an upstream sequence partially

complementary to the RBS. Base-pairing between the upstream sequence and the RBS 'locks' the RBS off. A 'key' RNA molecule that disrupts the mRNA secondary structure by preferentially base-pairing with the upstream sequence can be used to expose the RBS and increase translation initiation rate.

[00203] Accordingly, in some embodiments of the aspects described herein, a ribosome binding site (RBS) for use in the biological circuit chemotactic converters described herein comprises a sequence that is selected from the group consisting of **SEQ ID NO: 26- SEQ ID NO: 33**. In some embodiments of the aspects described herein, a ribosome binding site (RBS) comprises a sequence selected from the ribosome binding site sequences provided in **SEQ ID NO: 858- SEQ ID NO: 898**. In some embodiments of the aspects described herein, a ribosome binding site (RBS) comprises a sequence selected from the community ribosome binding site sequences provided in **SEQ ID NO: 899- SEQ ID NO: 906**. In some embodiments of the aspects described herein, a ribosome binding site (RBS) comprises a sequence selected from the miscellaneous prokaryotic ribosome binding site sequences provided in **SEQ ID NO: 907- SEQ ID NO: 994**. In some embodiments of the aspects described herein, a ribosome binding site (RBS) comprises a sequence selected from the regulated prokaryotic ribosome binding site sequences provided in **SEQ ID NO: 995- SEQ ID NO: 996**. In some embodiments of the aspects described herein, a ribosome binding site (RBS) comprises a sequence selected from the regulated yeast ribosome binding site sequences provided in **SEQ ID NO: 997- SEQ ID NO: 998**. In some embodiments of the aspects described herein, a ribosome binding site (RBS) comprises a sequence selected from the eukaryotic ribosome binding site sequences provided in **SEQ ID NO: 999 - SEQ ID NO: 1001**. In some embodiments of the aspects described herein, novel ribosome binding sites can be generated using automated design of synthetic ribosome sites, as described in Salis HM *et al.*, Nature Biotechnology 27, 946 - 950 (2009).

Terminators

[00204] Terminators are sequences that usually occur at the end of a gene or operon and cause transcription to stop, and are also provided for use in the modules and biological circuit chemotactic converters described herein to regulate transcription and prevent transcription from occurring in an unregulated fashion, *i.e.*, a terminator sequence prevents activation of downstream modules by upstream promoters. A "terminator" or "termination signal", as described herein, is comprised of the DNA sequences involved in specific termination of an RNA transcript by an RNA polymerase. Thus, in certain embodiments a terminator that ends the production of an RNA transcript is contemplated. A terminator can be necessary *in vivo* to achieve desirable message levels.

[00205] In prokaryotes, terminators usually fall into two categories (1) rho-independent terminators and (2) rho-dependent terminators. Rho-independent terminators are generally composed of palindromic sequence that forms a stem loop rich in G-C base pairs followed by several T bases. Without wishing to be bound by a theory, the conventional model of transcriptional termination is that the stem loop causes RNA polymerase to pause, and transcription of the poly-A tail causes the RNA:DNA duplex to unwind and dissociate from RNA polymerase.

[00206] The most commonly used type of terminator is a forward terminator. When placed downstream of a nucleic acid sequence that is usually transcribed, a forward transcriptional terminator will cause transcription to abort. In some embodiments, bidirectional transcriptional terminators are provided. Such terminators will usually cause transcription to terminate on both the forward and reverse strand. Finally, in some embodiments, reverse transcriptional terminators are provided that terminate transcription on the reverse strand only.

[00207] In eukaryotic systems, the terminator region can also comprise specific DNA sequences that permit site-specific cleavage of the new transcript so as to expose a polyadenylation site. This signals a specialized endogenous polymerase to add a stretch of about 200 A residues (polyA) to the 3' end of the transcript. RNA molecules modified with this polyA tail appear to more stable and are translated more efficiently. Thus, in those embodiments involving eukaryotes, it is preferred that a terminator comprises a signal for the cleavage of the RNA, and it is more preferred that the terminator signal promotes polyadenylation of the message. The terminator and/or polyadenylation site elements can serve to enhance message levels and/or to minimize read through between modules of the biological converter switches. As disclosed herein, terminators contemplated for use in the modules, biological circuit chemotactic converters, and methods of use thereof can include any known terminator of transcription described herein or known to one of ordinary skill in the art. Such terminators include, but are not limited to, the termination sequences of genes, such as for example, the bovine growth hormone terminator, or viral termination sequences, such as for example, the SV40 terminator. In certain embodiments, the termination signal encompasses a lack of transcribable or translatable sequence, such as due to a sequence truncation. The terminator used can be unidirectional or bidirectional.

[00208] Terminators for use in the modules and biological circuit chemotactic converters described herein can be selected from the non-limiting examples of Tables 1-5.

Table 1: Examples of Forward Terminators					
Name	Description	Direction	Efficiency Fwd. Rev.		Length
BBa_B0010	T1 from <i>E. coli</i> rrnB	Forward			80
BBa_B0012	TE from coliphage T7	Forward	0.309[CC]	-0.368[CC]	41
BBa_B0013	TE from coliphage T7 (+/-)	Forward	0.6[CC]	-1.06[CC]	47
BBa_B0015	double terminator (B0010-B0012)	Forward	0.984[CC] 0.97[JK]	0.295[CC] 0.62[JK]	129
BBa_B0017	double terminator (B0010-B0010)	Forward			168
BBa_B0053	Terminator (His)	Forward			72
BBa_B0055	-- No description --				78
BBa_B1002	Terminator (artificial, small, %T~85%)	Forward	0.98[CH]		34

Table 1: Examples of Forward Terminators					
BBa_B1003	Terminator (artificial, small, %T~80)	Forward	0.83[CH]		34
BBa_B1004	Terminator (artificial, small, %T~55)	Forward	0.93[CH]		34
BBa_B1005	Terminator (artificial, small, %T~25%)	Forward	0.86[CH]		34
BBa_B1006	Terminator (artificial, large, %T~>90)	Forward	0.99[CH]		39
BBa_B1010	Terminator (artificial, large, %T~<10)	Forward	0.95[CH]		40
BBa_I11013	Modification of biobricks part BBa_B0015				129
BBa_I51003	-- No description --				110
BBa_J61048	[rnpB-T1] Terminator	Forward	0.98[JCA]		113

Table 2: Examples of Bidirectional Terminators					
Name	Description	Direction	<u>Efficiency</u> Fwd. Rev.		Length
BBa_B0011	LuxICDABEG (+/-)	Bidirectional	0.419[CC]/0.95[JK]	0.636[CC]/0.86[JK]	46
BBa_B0014	double terminator (B0012-B0011)	Bidirectional	0.604[CC]/0.96[JK]	0.86[JK]	95
BBa_B0021	LuxICDABEG (+/-), reversed	Bidirectional	0.636[CC]/0.86[JK]	0.419[CC]/0.95[JK]	46
BBa_B0024	double terminator (B0012-B0011), reversed	Bidirectional	0.86[JK]	0.604[CC]/0.96[JK]	95
BBa_B0050	Terminator (pBR322, +/-)	Bidirectional			33
BBa_B0051	Terminator (yciA/tonA, +/-)	Bidirectional			35
BBa_B1001	Terminator (artificial, small, %T~90)	Bidirectional	0.81[CH]		34
BBa_B1007	Terminator (artificial, large, %T~80)	Bidirectional	0.83[CH]		40
BBa_B1008	Terminator (artificial, large, %T~70)	Bidirectional			40
BBa_B1009	Terminator (artificial, large, %T~40%)	Bidirectional			40
BBa_K259006	GFP-Terminator	Bidirectional	0.604[CC]/0.96[JK]	0.86[JK]	823

Table 3: Examples of Reverse Terminators					
Name	Description	Direction	<u>Efficiency</u> Fwd. Rev.		Length
BBa_B0020	Terminator (Reverse B0010)	Reverse			82
BBa_B0022	TE from coliphage T7, reversed	Reverse	-0.368[CC]	0.309[CC]	41
BBa_B0023	TE from coliphage T7, reversed	Reverse	-1.06[CC]	0.6[CC]	47

Table 3: Examples of Reverse Terminators

BBa_B0025	double terminator (B0015), reversed	Reverse	0.295[CC]/0.62[JK]	0.984[CC]/0.97[JK]	129
BBa_B0052	Terminator (rrnC)	Forward			41
BBa_B0060	Terminator (Reverse B0050)	Bidirectional			33
BBa_B0061	Terminator (Reverse B0051)	Bidirectional			35
BBa_B0063	Terminator (Reverse B0053)	Reverse			72

Table 4: Examples of Yeast Terminators

Name	Description	Direction	Efficiency		Length
			Fwd.	Rev.	
BBa_J63002	ADH1 terminator from <i>S. cerevisiae</i>	Forward			225
BBa_K110012	STE2 terminator	Forward			123
BBa_Y1015	CycE1				252

Table 5: Examples of Eukaryotic Terminators

Name	Description	Direction	Efficiency		Chassis	Length
			Fwd.	Rev.		
BBa_J52016	eukaryotic -- derived from SV40 early poly A signal sequence	Forward				238
BBa_J63002	ADH1 terminator from <i>S. cerevisiae</i>	Forward				225
BBa_K110012	STE2 terminator	Forward				123
BBa_Y1015	CycE1					252

Degradation Tags

[00209] In some embodiments of the aspects described herein, a nucleic sequence encoding a protein degradation tag can be added to the modules and biological circuit chemotactic converters described herein to enhance protein degradation of a protein, such as a recombinase or protein output product. As defined herein, a “degradation tag” is a genetic addition to the end of a nucleic acid sequence that modifies the protein that is expressed from that sequence, such that the protein undergoes faster degradation by cellular degradation mechanisms. Thus, such protein degradation tags ‘mark’ a protein for degradation, thus decreasing a protein's half-life.

[00210] One of the useful aspects of degradation tags is the ability to detect and regulate gene activity in a time-sensitive manner. Such protein degradation tags can operate through the use of protein-degrading enzymes, such as proteases, within the cell. In some embodiments, the tags encode for a sequence of about eleven amino acids at the C-terminus of a protein, wherein said sequence is normally generated in *E. coli* when a ribosome gets stuck on a broken (“truncated”) mRNA. Without a

normal termination codon, the ribosome can't detach from the defective mRNA. A special type of RNA known as ssrA ("small stable RNA A") or tmRNA ("transfer-messenger RNA") rescues the ribosome by adding the degradation tag followed by a stop codon. This allows the ribosome to break free and continue functioning. The tagged, incomplete protein can get degraded by the proteases ClpXP or ClpAP. Although the initial discovery of the number of amino acids encoding for an ssrA/tmRNA tag was eleven, the efficacy of mutating the last three amino acids of that system has been tested. Thus, the tags AAV, ASV, LVA, and LAA are classified by only three amino acids.

[00211] In some embodiments of the aspects described herein, the protein degradation tag is an ssrA tag. In some embodiments of the aspects described herein, the ssrA tag comprises a sequence that is selected from the group consisting of sequences that encode for the peptides RPAANDENYALAA (**SEQ ID NO: 34**), RPAANDENYALVA (**SEQ ID NO: 35**), RPAANDENYAAAV (**SEQ ID NO: 36**), and RPAANDENYAASV (**SEQ ID NO: 37**).

[00212] In some embodiments of the aspects described herein, the protein degradation tag is an LAA variant comprising the sequence GCAGCAAACGACGAAAACCTACGCTTTAGCAGCTTAA (**SEQ ID NO:38**). In one embodiment, the protein degradation tag is an AAV variant comprising the sequence GCAGCAAACGACGAAAACCTACGCTGCAGCAGTTTAA (**SEQ ID NO: 39**). In some embodiments, the protein degradation tag is an ASV variant comprising the sequence GCAGCAAACGACGAAAACCTACGCTGCATCAGTTTAA (**SEQ ID NO: 40**).

Output Product Sequences and Output Products

[00213] Also provided herein are a variety of biological outputs for use in the various modules and biological circuit chemotactic converters described herein. These biological outputs, or "output products," as defined herein, refer to products that can be used as markers of specific states of the modules and biological circuit chemotactic converters described herein. An output sequence can encode for a protein or an RNA molecule that is used to track or mark the state of the cell upon receiving a particular input for a biological circuit chemotactic converters. Such output products can be used to distinguish between various states of a cell. For example, upon an biological circuit chemotactic converter in a cell or cellular system receiving a specific input signal, the output gene product of the biological circuit chemotactic converter can be a sensor receptor. In some embodiments, the input signal causes the biological circuit chemotactic converter to switch from expression of one sensor to another sensor, in response to changes in chemicals in the external environment.

Reporter Outputs

[00214] In some embodiments of the aspects described herein, the output products are "reporters." As defined herein, "reporters" refer to proteins that can be used to measure gene expression. Reporters generally produce a measurable signal such as fluorescence, color, or luminescence. Reporter protein coding sequences encode proteins whose presence in the cell or

organism is readily observed. For example, fluorescent proteins cause a cell to fluoresce when excited with light of a particular wavelength, luciferases cause a cell to catalyze a reaction that produces light, and enzymes such as β -galactosidase convert a substrate to a colored product. In some embodiments, reporters are used to quantify the strength or activity of the signal received by the modules or biological converter switches of the invention. In some embodiments, reporters can be fused in-frame to other protein coding sequences to identify where a protein is located in a cell or organism.

[00215] There are several different ways to measure or quantify a reporter depending on the particular reporter and what kind of characterization data is desired. In some embodiments, microscopy can be a useful technique for obtaining both spatial and temporal information on reporter activity, particularly at the single cell level. In other embodiments, flow cytometers can be used for measuring the distribution in reporter activity across a large population of cells. In some embodiments, plate readers may be used for taking population average measurements of many different samples over time. In other embodiments, instruments that combine such various functions, can be used, such as multiplex plate readers designed for flow cytometers, and combination microscopy and flow cytometric instruments.

[00216] Fluorescent proteins are convenient ways to visualize or quantify the output of a module or a biological circuit chemotactic converter described herein. Fluorescence can be readily quantified using a microscope, plate reader or flow cytometer equipped to excite the fluorescent protein with the appropriate wavelength of light. Since several different fluorescent proteins are available, multiple gene expression measurements can be made in parallel. Non-limiting examples of fluorescent proteins are provided in Table 6.

Table 6: Examples of Fluorescent Protein Reporters						
Name	Protein	Description	Tag	Emission	Excitation	Length
BBa_E0030	EYFP	enhanced yellow fluorescent protein derived from <i>A. victoria</i> GFP	None	527	514	723
BBa_E0020	ECFP	engineered cyan fluorescent protein derived from <i>A. victoria</i> GFP	None	476	439	723
BBa_E1010	mRFP1	**highly** engineered mutant of red fluorescent protein from <i>Discosoma striata</i> (coral)	None	607	584	681
BBa_E2050	mOrange	derivative of mRFP1, yeast-optimized	None	562	548	744
BBa_E0040	GFPmut3b	green fluorescent protein derived from jellyfish <i>Aequorea victoria</i> wild-type GFP (SwissProt: P42212)	None	511	501	720
BBa_J52021		dnTraf6-linker-GFP				1446
BBa_J52026		dnMyD88-linker-GFP				1155
BBa_I715022		Amino Portion of RFP				462
BBa_I715023		Carboxyl portion of RFP				220

Table 6: Examples of Fluorescent Protein Reporters						
BBa_I712028		CherryNLS - synthetic construct monomeric red fluorescent protein with nuclear localization sequence				733
BBa_K125500		GFP fusion brick				718
BBa_K106000		GFP, AarI BD part				714
BBa_K106004		mCherry, AarI AB part				708
BBa_K106005		mCherry, AarI BD part				708
BBa_K106028		GFP, AarI AB part				714
BBa_K165005		Venus YFP, yeast optimized for fusion				744
BBa_K157005		Split-Cerulean-cCFP				261
BBa_K157006		Split-Cerulean-nCFP				483
BBa_K157007		Split-Venus-cYFP				261
BBa_K157008		Split-Venus-nYFP				486
BBa_K125810		slr2016 signal sequence + GFP fusion for secretion of GFP				779
BBa_K082003	GFP	GFP(+LVA)				756
BBa_K156009		OFp (orange fluorescent protein)				864
BBa_K156010		SBFP2 (strongly enhanced blue fluorescent protein)				720
BBa_K106671		GFP, AarI AD part				714
BBa_K294055	GFPmut3b	GFP RFP Hybrid	None	511	501	720
BBa_K192001		CFP +tgt +lva				858
BBa_K180001	GFPmut3b	Green fluorescent protein (+LVA)	LVA			754
BBa_K283005		lpp_ompA_eGFP_streptavidin				1533
BBa_K180008	mCherry	mCherry (rights owned by Clontech)				708
BBa_K180009	mBanana	mBanana (rights owned by Clontech)				708

[00217] Luminescence can be readily quantified using a plate reader or luminescence counter. Luciferases can be used as output products for various embodiments described herein, for example, measuring low levels of gene expression, because cells tend to have little to no background luminescence in the absence of a luciferase. Non-limiting examples of luciferases are provided in Table 7.

Table 7: Examples of Luciferases		
Name	Description	Length

Table 7: Examples of Luciferases		
BBa_J52011	dnMyD88-linker-Rluc	1371
BBa_J52013	dnMyD88-linker-Rluc-linker-PEST191	1872
BBa_I712019	Firefly luciferase - luciferase from Photinus pyralis	1653

[00218] In other embodiments, enzymes that produce colored substrates can be quantified using spectrophotometers or other instruments that can take absorbance measurements including plate readers. Like luciferases, enzymes like β -galactosidase can be used for measuring low levels of gene expression because they tend to amplify low signals. Non-limiting examples of such enzymes are provided in Table 8.

Table 8: Examples of Enzymes that Produce Colored Substrates		
Name	Description	Length
BBa_I732006	lacZ alpha fragment	234
BBa_I732005	lacZ (encoding beta-galactosidase, full-length)	3075
BBa_K147002	xylE	924

[00219] Another reporter output product for use in the different aspects described herein includes fluoresceine-A-binding (BBa_K157004).

Transcriptional Outputs:

[00220] In some embodiments of the aspects described herein, the output product of a given module or biological circuit chemotactic converter is itself a transcriptional activator or repressor, the production of which by an output product sequence can result in a further change in state of the cell, and provide additional input signals to subsequent or additional modules or biological circuit chemotactic converters. Transcriptional regulators either activate or repress transcription from cognate promoters. Transcriptional activators typically bind nearby to transcriptional promoters and recruit RNA polymerase to directly initiate transcription. Repressors bind to transcriptional promoters and sterically hinder transcriptional initiation by RNA polymerase. Other transcriptional regulators serve as either an activator or a repressor depending on where it binds and cellular conditions. Non-limiting examples of transcriptional regulators as output products are provided in Table 9.

Table 9: Examples of Transcriptional Regulators						
Name	Protein	Description	Tag	Direction	Uniprot	Length
BBa_C0079	lasR-LVA	lasR activator from P. aeruginosa PAO1(+LVA)	LVA	Forward	P25084	756
BBa_C0077	cinR	cinR activator from Rhizobium leguminosarum (+LVA)	LVA	Forward	~Q84HT2	762

Table 9: Examples of Transcriptional Regulators						
BBa_C0179	lasR	lasR activator from <i>P. aeruginosa</i> PAO1(no LVA)	None	Forward	P25084	723
BBa_J07009	ToxR	toxicity-gene activator from <i>Vibrio cholerae</i>	None	Forward	P15795	630
BBa_K118001		appY coding sequence encoding a DNA-binding transcriptional activator				753
BBa_K137113		rcaA				624
BBa_K131022		LuxO D47E, <i>Vibrio harveyi</i>				1362
BBa_K131023		LuxO D47A, <i>Vibrio harveyi</i>				1362
BBa_K082006		LuxR-G2F				753
BBa_K294205		This is a coding sequence of heat shock protein from <i>E.coli</i>				402
BBa_S04301	lasR-LVA	C0079:B0015	LVA	Forward	P25084	918
BBa_K266002	lasR-LVA	LasR + Term	LVA	Forward	P25084	918
BBa_C0012	LacI	lacI repressor from <i>E. coli</i> (+LVA)	LVA	Forward	P03023	1128
BBa_C0040	TetR	tetracycline repressor from transposon Tn10 (+LVA)	LVA	Forward	P04483	660
BBa_C0050	CI HK022	cI repressor from phage HK022 (+LVA?)	LVA	Forward	P18680	744
BBa_C0051	CI lambda	cI repressor from <i>E. coli</i> phage lambda (+LVA)	LVA	Forward	P03034	750
BBa_C0052	CI 434-LVA	cI repressor from phage 434 (+LVA)	LVA	Forward	P16117	669
BBa_C0053	C2 P22	c2 repressor from <i>Salmonella</i> phage P22 (+LVA)	LVA	Forward	P69202	687
BBa_C0073	mnt-weak	mnt repressor (weak) from <i>Salmonella</i> phage P22 (+LVA)	LVA	Forward	P03049	288
BBa_C0075	cI TP901	TP901 cI repressor from phage TP901-1 (+LVA)	LVA	Forward	none	579
BBa_C0074	penI	penI repressor from <i>Bacillus licheniformis</i> (+LVA)	LVA	Forward	P06555	423
BBa_C0072	mnt	mnt repressor (strong) from <i>Salmonella</i> phage P22 (+LVA)	LVA	Forward	P03049	288
BBa_C2001	Zif23-GCN4	Zif23-GCN4 engineered repressor (+LVA, C2000 codon-optimized for <i>E.coli</i>)	LVA	Forward	P03069	300
BBa_C0056	CI 434	cI repressor from phage 434 (no LVA)	None	Forward	P16117	636
BBa_J06501	LacI-mut2	LacI repressor (temperature-sensitive mut 265) (+LVA)	LVA	Forward	~ P03023	1153

Table 9: Examples of Transcriptional Regulators						
BBa_J06500	LacI-mut1	LacI repressor (temperature-sensitive mut 241) (+LVA)	LVA	Forward	~ P03023	1153
BBa_C2006		MalE.FactorXa.Zif268-GCN4				1428
BBa_I715032		lacIq reverse				1128
BBa_I732100		LacI				1086
BBa_I732101		LRLa				1086
BBa_I732105		ARL2A0101				1086
BBa_I732106		ARL2A0102				1086
BBa_I732107		ARL2A0103				1086
BBa_I732110		ARL2A0203				1086
BBa_I732112		ARL2A0301				1086
BBa_I732115		ARL4A0604				1086
BBa_K091001		LsrR gene		Forward		954
BBa_K091121		LacI wild-type gene				1083
BBa_K091122		LacI_I12 protein				1083
BBa_K143033		LacI (Lva ⁻ , N-terminal deletion) regulatory protein				1086
BBa_K142000		lacI IS mutant (IPTG unresponsive) R197A				1128
BBa_K142001		lacI IS mutant (IPTG unresponsive) R197F				1128
BBa_K142002		lacI IS mutant (IPTG unresponsive) T276A				1128
BBa_K142003		lacI IS mutant (IPTG unresponsive) T276F				1128
BBa_K106666		Lac Repressor, AarI AB part				1104
BBa_K106667		Lac Repressor, AarI BD part				1107
BBa_K142004		lacI IS mutant (IPTG unresponsive) R197A T276A				1128
BBa_K106668		Tet Repressor, AarI AB part				618
BBa_K106669		Tet Repressor, AarI BD part				621
BBa_K142005		lacI IS mutant (IPTG unresponsive) R197A T276F				1128
BBa_K142006		lacI IS mutant (IPTG unresponsive) R197F T276A				1128
BBa_K142007		lacI IS mutant (IPTG unresponsive) R197F T276F				1128

Table 9: Examples of Transcriptional Regulators						
BBa_K082004	LacI	LacI- wild type				1083
BBa_K082005	LacI	LacI-Mutant				1083
BBa_C0062	LuxR	luxR repressor/activator, (no LVA?)	None	Forward	P12746	756
BBa_C0071	rhIR-LVA	rhIR repressor/activator from <i>P. aeruginosa</i> PA3477 (+LVA)	LVA	Forward	P54292	762
BBa_C0080	araC	araC arabinose operon regulatory protein (repressor/activator) from <i>E. coli</i> (+LVA)	LVA	Forward	P0A9E0	915
BBa_C0171	rhIR	rhIR repressor/activator from <i>P. aeruginosa</i> PA3477 (no LVA)	None	Forward	P54292	729
BBa_K108021		Fis				297

Selection Markers

[00221] In other embodiments of the various aspects described herein, genes encoding selection markers are used as output product sequences. “Selection markers,” as defined herein, refer to protein coding sequences that confer a selective advantage or disadvantage to a biological unit, such as a cell. For example, a common type of prokaryotic selection marker is one that confers resistance to a particular antibiotic. Thus, cells that carry the selection marker can grow in media despite the presence of antibiotic. For example, most plasmids contain antibiotic selection markers so that it is ensured that the plasmid is maintained during cell replication and division, as cells that lose a copy of the plasmid will soon either die or fail to grow in media supplemented with antibiotic. A second common type of selection marker, often termed a positive selection marker, are those that are toxic to the cell. Positive selection markers are frequently used during cloning to select against cells transformed with the cloning vector and ensure that only cells transformed with a plasmid containing the insert. Non-limiting examples of selection marker output products are provided in Table 10.

Name	Protein	Description	UniProt	KEGG	Length
BBa_T9150	PyrF	orotidine 5	P08244	eco:b1281;	741
BBa_J31002	AadA-bkw	kanamycin resistance backwards (KanB) [cf. BBa_J23012 & BBa_J31003]	P0AG05	none	816
BBa_J31003	AadA2	kanamycin resistance forward (KanF) [cf. BBa_J23012 & BBa_J31002]	P0AG05	none	816
BBa_J31004	CAT-bkw	chloramphenicol acetyltransferase (backwards, CmB) [cf. BBa_J31005]	P62577	none	660
BBa_J31006	TetA(C)-bkw	tetracycline resistance protein TetA(C) (backwards) [cf. BBa_J31007]	P02981		1191
BBa_J31005	CAT	chloramphenicol acetyltransferase (forwards, CmF) [cf. BBa_J31004]	P62577	none	660
BBa_J31007	TetA(C)	tetracycline resistance protein TetA(C)	P02981		1191

		(forward), [cf. BBa_J31006]			
BBa_K145151		ccdB coding region			306
BBa_K143031		Aad9 Spectinomycin Resistance Gene			771
BBa_K156011		aadA (streptomycin 3'-adenyltransferase)			789

Enzyme Outputs

[00222] An output product sequence can encode an enzyme for use in different embodiments the modules and biological circuit chemotactic converters described herein. In some embodiments, an enzyme output can be used as a response to a particular input. For example, in response to a particular input received by a biological circuit chemotactic converter of the invention, such as a certain range of toxin concentration present in the environment, a module of the biological circuit chemotactic converter can “turn on” a modular component that encodes as an output gene product an enzyme that can degrade or otherwise destroy the toxin.

[00223] In some embodiments of the aspects described herein, output product sequences encode “biosynthetic enzymes” that catalyze the conversion of substrates to products. For example, such biosynthetic enzymes can be combined together along with or within the modules and biological circuit chemotactic converters described herein to construct pathways that produce or degrade useful chemicals and materials, in response to specific signals. These combinations of enzymes can reconstitute either natural or synthetic biosynthetic pathways. These enzymes have applications in specialty chemicals, biofuels, and bioremediation.

[00224] For example, N-Acyl Homoserine lactones (AHLs or N-AHLs) are a class of signaling molecules involved in bacterial quorum sensing. “Quorum sensing” refers to a method of communication between bacteria that enables the coordination of group based behavior based on population density. In synthetic biology, genetic parts derived from quorum sensing systems have been used to create patterns on a lawn of bacteria and to achieve synchronized cell behavior. AHL can diffuse across cell membranes and is stable in growth media over a range of pH values. AHL can bind to transcriptional activators such as LuxR and stimulate transcription from cognate promoters. Several similar quorum sensing systems exists across different bacterial species; thus, there are several known enzymes that synthesize or degrade different AHL molecules that can be used for the modules and biological circuit chemotactic converters described herein.

Table 11: Examples of AHLs							
Name	Protein	Description	Direction	Uniprot	KEGG	E.C.	Length
BBa_C0061	luxI-LVA	autoinducer synthetase for AHL	Forward	P12747	none	none	618
BBa_C0060	aiiA-LVA	autoinducer inactivation enzyme from Bacillus; hydrolyzes acetyl	Forward	Q1WNZ5	none	3.1.1.-	789

Table 11: Examples of AHLs							
		homoserine lactone					
BBa_C0070	rhII-LVA	autoinducer synthetase for N-butyryl-HSL (BHL) and HHL	Forward	Q02QW5	none	none	642
BBa_C0076	cinI	autoinducer synthetase	Forward	Q1MDW1	none	none	702
BBa_C0078	lasI	autoinducer synthetase for PAI from <i>Pseudomonas aeruginosa</i>	Forward	P33883	pae:PA1432	none	642
BBa_C0161	luxI	autoinducer synthetase for AHL (no LVA)	Forward	P12747	none	none	585
BBa_C0170	rhII	autoinducer synthetase for N-butyryl-HSL (BHL) and HHL (no LVA)	Forward	Q02QW5	none	none	609
BBa_C0178	lasI	autoinducer synthetase for PAI from <i>Pseudomonas aeruginosa</i> (no LVA)	Forward	P33883	pae:PA1432	none	609
BBa_K091109		LuxS					516
BBa_C0060	aiiA-LVA	autoinducer inactivation enzyme from <i>Bacillus</i> ; hydrolyzes acetyl homoserine lactone	Forward	Q1WNZ5	none	3.1.1.-	789
BBa_C0160	aiiA	autoinducer inactivation enzyme aiiA (no LVA)	Forward	Q1WNZ5	none	3.1.1.-	756

[00225] “Isoprenoids,” also known as “terpenoids,” refer to a large and highly diverse class of natural organic chemicals with many functions in plant primary and secondary metabolism. Most are multicyclic structures that differ from one another not only in functional groups but also in their basic carbon skeletons. Isoprenoids are synthesized from common prenyl diphosphate precursors through the action of terpene synthases and terpene-modifying enzymes such as cytochrome P450 monooxygenases. Plant terpenoids are used extensively for their aromatic qualities. They play a role in traditional herbal remedies and are under investigation for antibacterial, antineoplastic, and other pharmaceutical functions. Much effort has been directed toward their production in microbial hosts.

[00226] There are two primary pathways for making isoprenoids: the mevalonate pathway and the non-mevalonate pathway. Exemplary isoprenoids for use as output products in the modules and biological circuit chemotactic converters described herein are provided in Table 12.

Table 12: Examples of Isoprenoids		
Name	Description	Length

Table 12: Examples of Isoprenoids		
BBa_K118000	dxs coding sequence encoding 1-deoxyxylulose-5-phosphate synthase	1866
BBa_K115050	A-coA -> AA-coA	1188
BBa_K115056	IPP -> OPP or DMAPP -> OPP	552
BBa_K115057	OPP -> FPP	903
BBa_K118002	crtB coding sequence encoding phytoene synthase	933
BBa_K118003	crtI coding sequence encoding phytoene dehydrogenase	1482
BBa_K118008	crtY coding sequence encoding lycopene B-cyclase	1152

[00227] Odorants are volatile compounds that have an aroma detectable by the olfactory system. Odorant enzymes convert a substrate to an odorant product. Exemplary odorant enzymes for use as output products in the modules and biological circuit chemotactic converters described herein are provided in Table 13.

Table 13: Examples of Odorant Enzymes						
Name	Protein	Description	Uniprot	KEGG	E.C.	Length
BBa_J45001	SAMT	SAM:salicylic acid carboxyl methyltransferase; converts salicylic acid to methyl salicylate (winter)	Q8H6N2	none	none	1155
BBa_J45002	BAMT	SAM:benzoic acid carboxyl methyltransferase; converts benzoic acid to methyl benzoate (floral odor)	Q9FYZ9	none	2.1.1.-	1098
BBa_J45004	BSMT1	SAM:benzoic acid/salicylic acid carboxyl methyltransferase I; converts salicylic acid to methyl sali	Q84UB5	none	none	1074
BBa_J45008	BAT2	branched-chain amino acid transaminase (BAT2); converts leucine to alpha-ketoisocaproate	P47176	sce:YJR148W	2.6.1.42	1134
BBa_J45014	ATF1-1148 mutant	alcohol acetyltransferase I; converts isoamyl alcohol to isoamyl acetate (banana odor)	P40353	sce:YOR377W	2.3.1.84	1581
BBa_J45017	PchA & PchB	isochorismate pyruvate-lyase and isochorismate synthase (pchBA); converts chorismate to salicylate				1736
BBa_I742107		COMT				1101

[00228] The following are exemplary enzymes involved in the biosynthesis of plastic, specifically polyhydroxybutyrate, for use as output products in the modules and biological circuit chemotactic converters described herein.

Table 14: Examples of Plastic Biosynthesis Enzymes		
Name	Description	Length
BBa_K125504	phaE BioPlastic polyhydroxybutyrate synthesis pathway (origin PCC6803 slr1829)	996
BBa_K125501	phaA BioPlastic polyhydroxybutyrate synthesis pathway (origin PCC6803 slr1994)	1233
BBa_K125502	phaB BioPlastic polyhydroxybutyrate synthesis pathway (origin PCC6803 slr1993)	726
BBa_K125503	phaC BioPlastic polyhydroxybutyrate synthesis pathway (origin PCC6803 slr1830)	1140
BBa_K156012	phaA (acetyl-CoA acetyltransferase)	1182
BBa_K156013	phaB1 (acetylacetyl-CoA reductase)	741
BBa_K156014	phaC1 (Poly(3-hydroxybutyrate) polymerase)	

[00229] The following are exemplary enzymes involved in the biosynthesis of butanol and butanol metabolism for use as output products in the modules and biological circuit chemotactic converters described herein.

Table 15: Examples of Butanol Biosynthesis Enzymes		
Name	Description	Length
BBa_I725011	B-hydroxy butyryl coA dehydrogenase	870
BBa_I72512	Enoyl-coa hydratase	801
BBa_I725013	Butyryl CoA dehydrogenase	1155
BBa_I725014	Butyraldehyde dehydrogenase	2598
BBa_I725015	Butanol dehydrogenase	1188

[00230] Bisphenol A is a toxin that has been shown to leech from certain types of plastic. Recent studies have shown this chemical to have detrimental effects in animal studies and that it can be harmful to humans as well. Exemplary bisphenol A degradation enzymes from *Sphingomonas bisphenolicum* include BisdB and BisdA and can be used, for example, as an output product in the modules and biological circuit chemotactic converters described herein, for the remediation of bisphenol A contamination.

[00231] Other exemplary miscellaneous enzymes for use as output products in the modules and biological circuit chemotactic converters described herein are provided in Table 16.

Table 16: Examples of Miscellaneous Biosynthetic Enzymes						
Name	Description	Direction	Uniprot	KEGG	E.C.	Length
BBa_K118022	cex coding sequence encoding					1461

Table 16: Examples of Miscellaneous Biosynthetic Enzymes						
	Cellulomonas fimi exoglucanase					
BBa_K118023	cenA coding sequence encoding Cellulomonas fimi endoglucanase A					1353
BBa_K118028	beta-glucosidase gene bglX (chu_2268) from Cytophaga hutchinsonii					2280
BBa_C0083	aspartate ammonia-lyase	Forward	P0AC38	eco:b4139	4.3.1.1	1518
BBa_I15008	heme oxygenase (ho1) from Synechocystis	Forward	P72849	syn:slr1184	1.14.99.3	726
BBa_I15009	phycocyanobilin:ferredoxin oxidoreductase (PcyA) from synechocystis	Forward	Q55891	syn:slr0116	1.3.7.5	750
BBa_T9150	orotidine 5	Forward	P08244	eco:b1281;	4.1.1.23	741
BBa_I716153	hemB					975
BBa_I716154	hemC					942
BBa_I716155	hemD					741
BBa_I716152	hemA (from CFT703)					1257
BBa_I742141	sam5 (coumarate hydroxylase) coding sequence					1542
BBa_I742142	sam8 (tyrosine-ammonia lyase) coding sequence					1536
BBa_I723024	PhzM					1019
BBa_I723025	PhzS					1210
BBa_K137005	pabA (from pABA synthesis)					585
BBa_K137006	pabB (from pABA synthesis)					1890
BBa_K137009	folB (dihydroneopterin aldolase)					354
BBa_K137011	folKE (GTP Cyclohydrolase I + pyrophosphokinase)					1053
BBa_K137017	Galactose Oxidase					1926
BBa_K118015	glgC coding sequence encoding ADP-glucose pyrophosphorylase					1299
BBa_K118016	glgC16 (glgC with G336D substitution)					1299
BBa_K123001	BisdB					1284
BBa_K108018	PhbAB					1997
BBa_K108026	XylA					1053
BBa_K108027	XylM					1110

Table 16: Examples of Miscellaneous Biosynthetic Enzymes						
BBa_K108028	XylB					1101
BBa_K108029	XylS					966
BBa_K147003	ohbA					531
BBa_K123000	BisdA					330
BBa_K284999	Deletar este					1431
BBa_I716253	HPI, katG					2181
BBa_K137000	katE					2265
BBa_K137014	katE +LAA					2298
BBa_K137067	katG					2184
BBa_K078102	dxnB					886
BBa_K078003	one part of the initial dioxygenase of the dioxin degradation pathway					1897

[00232] Other enzymes of use as output products in the modules and biological circuit chemotactic converters described herein include enzymes that phosphorylate or dephosphorylate either small molecules or other proteins, or enzymes that methylate or demethylate other proteins or DNA.

Table 17: Examples of Phosphorylation and Methylation-Related Enzymes							
Name	Protein	Description	Direction	Uniprot	KEGG	E.C.	Length
BBa_C0082	tar-envZ	Receptor, tar-envZ	Forward				1491
BBa_J58104		Fusion protein Trg-EnvZ for signal transduction					1485
BBa_J58105		Synthetic periplasmic binding protein that docks a vanillin molecule					891
BBa_I752001		CheZ coding sequence (Chemotaxis protein)					639
BBa_K091002		LsrK gene	Forward				1593
BBa_K147000		cheZ					835
BBa_K118015		glgC coding sequence encoding ADP-glucose pyrophosphorylase					1299
BBa_K118016		glgC16 (glgC with G336D substitution)					1299
BBa_K094100		cheZ gene					695
BBa_K136046		envZ*					1353

Table 17: Examples of Phosphorylation and Methylation-Related Enzymes							
BBa_K283008	chez	chez_Histag					713
BBa_C0024	CheB	CheB chemotaxis coding sequence (protein glutamate methylesterase)	Forward	P07330	JW1872	3.1.1.61	1053
BBa_K108020		Dam					837

[00233] Also useful as output products in the modules and biological circuit chemotactic converters described herein are receptors, ligands, and lytic proteins. Receptors tend to have three domains: an extracellular domain for binding ligands such as proteins, peptides or small molecules, a transmembrane domain, and an intracellular or cytoplasmic domain which frequently can participate in some sort of signal transduction event such as phosphorylation. In some embodiments, transporter, channel, or pump gene sequences are used as output products. Transporters are membrane proteins responsible for transport of substances across the cell membrane. Channels are made up of proteins that form transmembrane pores through which selected ions can diffuse. Pumps are membrane proteins that can move substances against their gradients in an energy-dependent process known as active transport. In some embodiments, nucleic acid sequences encoding proteins and protein domains whose primary purpose is to bind other proteins, ions, small molecules, and other ligands are used as output products in the modules and biological circuit chemotactic converters described herein. Exemplary receptors, ligands, and lytic proteins are listed in Table 18.

Table 18: Examples of Receptors, Ligands, and Lytic Proteins						
Name	Protein	Description	Tag	Direction	UniProt	Length
BBa_J07009	ToxR	toxicity-gene activator from <i>Vibrio cholerae</i>	None	Forward	P15795	630
BBa_K133063		(TIR)TLR3				453
BBa_K133064		(TIR)TLR9				585
BBa_K133065		(TMTIR)TLR3				600
BBa_K133069		(TMTIR)TLR3stop				603
BBa_K133067		(TMTIR)TLR4				621
BBa_K133060		(TMTIR)TLR9				645
BBa_K209400		AarI B-C part, hM4D				1434
BBa_K209401		AarI B-C part, Rs1.3				1407
BBa_I712002		CCR5				1059
BBa_I712003		CCR5-NUb				1194
BBa_I712010		CD4 sequence without signal peptide				1299
BBa_I712017		Chemokine (CXC motif) receptor 4,				1191

Table 18: Examples of Receptors, Ligands, and Lytic Proteins						
		fused to N-terminal half of ubiquitin.				
BBa_I15010	Cph8	cph8 (Cph1/EnvZ fusion)	None	Forward		2238
BBa_I728500		CPX Terminal Surface Display Protein with Polystyrene-Binding Peptide				654
BBa_J52035		dnMyD88				420
BBa_K259000		fhuA - Outer membrane transporter for ferrichrome-iron				2247
BBa_K259001		fiu B Outer Membrane Ferric Iron Transporter				2247
BBa_J58104		Fusion protein Trg-EnvZ for signal transduction				1485
BBa_K137112		IamB				1339
BBa_C0082	tar-envZ	Receptor, tar-envZ	LVA	Forward		1491
BBa_J58105		Synthetic periplasmic binding protein that docks a vanillin molecule				891
BBa_I712012		TIR domain of TLR3				456
BBa_K143037		YtvA Blue Light Receptor for B.subtilis				789
BBa_J07006		malE				1191
BBa_J07017		FecA protein				2325
BBa_K141000	UCP1	Ucp1				924
BBa_K141002		Ucp 175 deleted				921
BBa_K141003		Ucp 76 deleted				921
BBa_K190028		GlpF				846
BBa_I746200		FepA L8T Mutant - Large Diffusion pore for <i>E. coli</i> outer membrane.				2208
BBa_I765002		ExbB membrane spanning protein in TonB-ExbB-ExbD complex [<i>Escherichia coli</i> K12]				735
BBa_I765003		TonB ferric siderophore transport system, periplasmic binding protein TonB [<i>Pseudomonas entomophila</i>]				735
BBa_K090000		Glutamate gated K ⁺ channel				1194
BBa_K284000		Lactate Permease from <i>Kluyveromyces fragilis</i>				1873
BBa_K284997		Deletar este				1069
BBa_J22101		Lac Y gene				1288
BBa_K079015		LacY transporter protein from <i>E. coli</i>				1254

Table 18: Examples of Receptors, Ligands, and Lytic Proteins						
BBa_K119003		RcnA (YohM)				833
BBa_K137001		LacY				1254
BBa_I712024		CD4				1374
BBa_K133061		CD4 ecto				1113
BBa_K136046		envZ*				1353
BBa_K157002		Transmembrane region of the EGF-Receptor (ErbB-1)				87
BBa_K227006		puc BA coding region of R. sphaeroides		forward		336
BBa_M12067		E1				264
BBa_I721002		Lead Binding Protein				399
BBa_K126000		TE33 Fab L chain				648
BBa_K133070		gyrEC				660
BBa_K133062		gyrHP				660
BBa_K157003		Anti-NIP singlechain Fv-Fragment				753
BBa_K211001		RI7				987
BBa_K211002		RI7-odr10 chimeric GPCR				1062
BBa_K103004		protein Z _{SPA-1}				190
BBa_K128003		p1025				101
BBa_K133059		RGD				9
BBa_K283010		Streptavidin				387
BBa_K103004		protein Z _{SPA-1}				190
BBa_K128003		p1025				101
BBa_K133059		RGD				9
BBa_K283010		Streptavidin				387
BBa_K112000	Holin	T4 holin, complete CDS, berkeley standard				657
BBa_K112002	Holin	T4 holin, without stop codon, berkeley standard				654
BBa_K112004		a~T4 holin in BBb				661
BBa_K112006		T4 antiholin in BBb				294
BBa_K112009		in BBb				288
BBa_K112010		a~T4 antiholin in BBb				298
BBa_K112012		T4 lysozyme in BBb				495
BBa_K112015		in BBb				489

Table 18: Examples of Receptors, Ligands, and Lytic Proteins						
BBa_K112016		a~T4 lysozyme in BBb				499
BBa_K117000		Lysis gene (promotes lysis in colicin-producing bacteria strain)				144
BBa_K124014		Bacteriophage Holin Gene pS105				317
BBa_K108001		SRRz				1242
BBa_K112300		{lambda lysozyme} in BBb format				477
BBa_K112304		{a~lambda lysozyme} in BBb format				481
BBa_K112306		{lambda holin} in BBb format				318
BBa_K112310		{a~lambda holin}; adheres to Berkeley standard				322
BBa_K112312		{lambda antiholin}; adheres to Berkeley standard				324
BBa_K112316		{a~lambda antiholin}; adheres to Berkeley standard				328
BBa_K124017		Bacteriophage Lysis Cassette S105, R, and Rz				1257
BBa_K112806		[T4 endolysin]				514
BBa_K284001		Lysozyme from Gallus gallus				539

[00234] Particularly useful for the different embodiments of the various aspects described herein are chemotaxis sensors. Exemplary, non-limiting chemotaxis sensors are provided in Table 19. Such chemotaxis sensors can be further engineered by, in some embodiments, swapping in new sensory domains to allow for chemotaxis towards new chemicals, such as, for example, biological toxins or environmental pollutants. For example, the sensing and transmembrane signaling modules of different chemoreceptors can be swapped to generate functional chimeric receptors with directed changes in ligand specificity. In some embodiments, engineering of chemotactic sensors can be achieved by using other known sensing proteins from two-component pathways from various organisms, including prokaryotes and eukaryotes, and swapping these sensory domains to form novel engineered chemotaxis sensor proteins.

Table 19: Examples of Chemotaxis Sensors/Receptors								
Protein	Gene	Location	Monomer Mass (kDa)	Oligomer Number	Monomers/cell	Monomer Concentration (μ M)	Affinity (K_d) for signaling complex (μ M)	References
Maltose-binding protein	<i>malE</i>	Periplasm	41	1	45,000	400	250	1,2,11
Galactose/galucose-	<i>mglB</i>	Periplasm	33	1	20,000	200		1,2

Table 19: Examples of Chemotaxis Sensors/Receptors

binding protein								
Ribose-binding protein	<i>rbsB</i>	Periplasm	30	1	40,000	400		1,2
Dipeptide-binding protein	<i>dpp</i>	Periplasm	57	1				2
Ni ²⁺ -binding protein	<i>nikA</i>	Periplasm	57	1	20,000	200		3
Aspartate receptor	<i>tar</i>	Membrane	60	2	1,500	2		6,13,16,17
Serine receptor	<i>tsr</i>	Membrane	59	(2)	3,000	4		13,16,17
Ribose/galactose receptor	<i>trg</i>	Membrane	59	2	150	0.2		5,16,17
Dipeptide receptor	<i>tap</i>	Membrane	58	(2)	150	0.2		2,4,16,17
Citrate receptor	<i>tcp</i>	Membrane	59	(2)				7

References: (1) Koman et al 1979; (2) Macnab 1987; (3) DePina et al 1995; (4) Stock et al 1991; (5) Hazelbauer & Adler 1971; (6) Milligan & Koshland 1988; (7) Yamamoto & Imae 1993; (8) Simms and Subbaramaiah 1991; (9) Kihara et al 1989; (10) Tang & Blair 1995; (11) Manson et al 1985; (12) Li et al 1995; (13) Gegner et al 1992; (14) Sanatnia et al 1995; (15) Blat & Eisenbach 1996b; (16) Bray et al 1993; (17) Bray & Bouret 1995.

[00235] Non-limiting examples of engineered chimeric chemotactic receptors, with swapped domains, for use in the sensor modules and biological circuit chemotactic converters described herein are provided in Table20.

Table 20: Examples of Engineered Chimeric Chemotactic Receptors

Sensor Domain	Signaling Domain	Source	Details
Tar	Tsr	Krikos et al. (1985) PNAS 82: 1326-1330.	Tar /Tsr chimeras Produced functional chemotactic receptors.
NarX	Tar	Ward et al. (2002) Mol. Microbial. 44(3): 709-719.	NarX sensor - Tar signalling domain chimera produced a chemotactic response to nitrate and nitrites.
Tar	Human Insulin Receptor	Moe et al. (1989) PNAS 86: 5683-5687.	Tar sensor - Human insulin receptor cytosolic domain produced a tyrosine kinase chimera
McpB	McpC	Kristich and Ordal, (2004) J.Bact. 186: 5950-5955.	Domains swapped between B.subtilis chemoreceptors
Trg	EnvZ	Baumgartner et al., (1993) J. Bact. 176: 1157-1163.	Recognition of sugar-binding protein ligands by Trg chemoreceptor and ompC promoter activation by EnvZ.

Table 20: Examples of Engineered Chimeric Chemotactic Receptors

Cph1	EnvZ	Levskaia et al. (2005). Nature 438: 441-442	Photoreceptor Cph1 from Synechocystis - EnvZ cytoplasmic domain chimera produces light-activated receptor
------	------	---	---

RNA Molecule Components and Output Products

[00236] Double-stranded (dsRNA) has been shown to direct the sequence-specific silencing of mRNA through a process known as RNA interference (RNAi). The process occurs in a wide variety of organisms, including mammals and other vertebrates. Accordingly, in some embodiments of the aspects described herein, sequences encoding RNA molecules can be used as components or output products in the biological circuit chemotactic converters. Such RNA molecules can be double-stranded or single-stranded and are designed, in some embodiments, to mediate RNAi, *e.g.*, with respect to another output product or, for example, to inhibit a recombinase present in a SIMM (see, for example, Figure 8). In those embodiments where a sequence encodes an RNA molecule that acts to mediate RNAi, the sequence can be said to encode an "iRNA molecule."

[00237] In some embodiments, an iRNA molecule can have any architecture described herein. *e.g.*, it can be incorporate an overhang structure, a hairpin or other single strand structure or a two-strand structure, as described herein. An "iRNA molecule" as used herein, is an RNA molecule which can by itself, or which can be cleaved into an RNA agent that can, downregulate the expression of a target sequence, *e.g.*, a recombinase sequence in a SIMM or an output product encoded by another module or biological circuit chemotactic converter, as described herein. While not wishing to be bound by theory, an iRNA molecule can act by one or more of a number of mechanisms, including post-transcriptional cleavage of a target mRNA sometimes referred to in the art as RNAi, or pre-transcriptional or pre-translational mechanisms. An iRNA molecule can include a single strand or can include more than one strand, *e.g.*, it can be a double stranded iRNA molecule.

[00238] The sequence encoding an iRNA molecule should include a region of sufficient homology to a target sequence, and be of sufficient length in terms of nucleotides, such that the iRNA molecule, or a fragment thereof, can mediate down regulation of the target sequence. Thus, the iRNA molecule is or includes a region that is at least partially, and in some embodiments fully, complementary to a target RNA sequence. It is not necessary that there be perfect complementarity between the iRNA molecule and the target sequence, but the correspondence must be sufficient to enable the iRNA molecule t, or a cleavage product thereof, to direct sequence specific silencing, *e.g.*, by RNAi cleavage of the target RNA sequence, *e.g.*, mRNA.

[00239] Complementarity, or degree of homology with the target strand, is most critical in the antisense strand. While perfect complementarity, particularly in the antisense strand, is often desired some embodiments can include, particularly in the antisense strand, one or more but preferably 6, 5, 4, 3, 2, or fewer mismatches (with respect to the target RNA). The mismatches, particularly in the antisense strand, are most tolerated in the terminal regions and if present are preferably in a terminal

region or regions, *e.g.*, within 6, 5, 4, or 3 nucleotides of the 5' and/or 3' terminus. The sense strand need only be sufficiently complementary with the antisense strand to maintain the overall double strand character of the molecule.

[00240] iRNA molecules for use in the modules and biological converter switches described herein include: molecules that are long enough to trigger the interferon response (which can be cleaved by Dicer (Bernstein et al. 2001. *Nature*, 409:363-366) and enter a RISC (RNAi-induced silencing complex); and, molecules that are sufficiently short that they do not trigger the interferon response (which molecules can also be cleaved by Dicer and/or enter a RISC), *e.g.*, molecules that are of a size which allows entry into a RISC, *e.g.*, molecules which resemble Dicer-cleavage products. Molecules that are short enough that they do not trigger an interferon response are termed "sRNA molecules" or "shorter iRNA molecules" herein. Accordingly, a sRNA molecule or shorter iRNA molecule, as used herein, refers to an iRNA molecule, *e.g.*, a double stranded RNA molecule or single strand molecule, that is sufficiently short that it does not induce a deleterious interferon response in a mammalian cell, such as a human cell, *e.g.*, it has a duplexed region of less than 60 but preferably less than 50, 40, or 30 nucleotide pairs. The sRNA molecule, or a cleavage product thereof, can downregulate a target sequence, *e.g.*, by inducing RNAi with respect to a target RNA sequence.

[00241] Each strand of an sRNA molecule can be equal to or less than 30, 25, 24, 23, 22, 21, or 20 nucleotides in length. The strand is preferably at least 19 nucleotides in length. For example, each strand can be between 21 and 25 nucleotides in length. Preferred sRNA molecules have a duplex region of 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotide pairs, and one or more overhangs, preferably one or two 3' overhangs, of 2-3 nucleotides.

[00242] A "single strand iRNA molecule" as used herein, is an iRNA molecule that is made up of a single molecule. It may include a duplexed region, formed by intra-strand pairing, *e.g.*, it may be, or include, a hairpin or pan-handle structure. Single strand iRNA molecules are preferably antisense with regard to the target sequence. A single strand iRNA molecule should be sufficiently long that it can enter the RISC and participate in RISC mediated cleavage of a target mRNA. A single strand iRNA molecule for use in the modules and biological converter switches described herein is at least 14, and more preferably at least 15, 20, 25, 29, 35, 40, or 50 nucleotides in length. It is preferably less than 200, 100, or 60 nucleotides in length.

[00243] Hairpin iRNA molecules can have a duplex region equal to or at least 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotide pairs. The duplex region is preferably equal to or less than 200, 100, or 50, in length. Preferred ranges for the duplex region are 15-30, 17 to 23, 19 to 23, and 19 to 21 nucleotides pairs in length. The hairpin preferably has a single strand overhang or terminal unpaired region, preferably the 3', and preferably of the antisense side of the hairpin. Preferred overhangs are 2-3 nucleotides in length.

[00244] A "double stranded (ds) iRNA molecule" as used herein, refers to an iRNA molecule that includes more than one, and preferably two, strands in which interchain hybridization can form a

region of duplex structure. The antisense strand of a double stranded iRNA molecule should be equal to or at least, 14, 15, 16 17, 18, 19, 25, 29, 40, or 60 nucleotides in length. It should be equal to or less than 200, 100, or 50, nucleotides in length. Preferred ranges are 17 to 25, 19 to 23, and 19 to 21 nucleotides in length. The sense strand of a double stranded iRNA molecule should be equal to or at least 14, 15, 16 17, 18, 19, 25, 29, 40, or 60 nucleotides in length. It should be equal to or less than 200, 100, or 50, nucleotides in length. Preferred ranges are 17 to 25, 19 to 23, and 19 to 21 nucleotides in length. The double strand portion of a double stranded iRNA molecule should be equal to or at least, 14, 15, 16 17, 18, 19, 20, 21, 22, 23, 24, 25, 29, 40, or 60 nucleotide pairs in length. It should be equal to or less than 200, 100, or 50, nucleotides pairs in length. Preferred ranges are 15-30, 17 to 23, 19 to 23, and 19 to 21 nucleotides pairs in length.

[00245] In some embodiments, the ds iRNA molecule is sufficiently large that it can be cleaved by an endogenous molecule, *e.g.*, by Dicer, to produce smaller ds iRNA agents, *e.g.*, sRNAs agents

[00246] It is preferred that the sense and antisense strands be chosen such that the ds iRNA molecule includes a single strand or unpaired region at one or both ends of the molecule. Thus, a ds iRNA agent contains sense and antisense strands, preferable paired to contain an overhang, *e.g.*, one or two 5' or 3' overhangs but preferably a 3' overhang of 2-3 nucleotides. Most embodiments have a 3' overhang. Preferred sRNA molecule have single-stranded overhangs, preferably 3' overhangs, of 1 or preferably 2 or 3 nucleotides in length at each end. The overhangs can be the result of one strand being longer than the other, or the result of two strands of the same length being staggered. 5' ends are preferably phosphorylated.

[00247] Preferred lengths for the duplexed region is between 15 and 30, most preferably 18, 19, 20, 21, 22, and 23 nucleotides in length, *e.g.*, in the sRNA molecule range discussed above. sRNA molecules can resemble in length and structure the natural Dicer processed products from long dsRNAs. Hairpin, or other single strand structures which provide the required double stranded region, and preferably a 3' overhang are also encompassed within the term sRNA molecule, as used herein.

[00248] The iRNA molecules described herein, including ds iRNA molecules and sRNA molecules, can mediate silencing of a target RNA, *e.g.*, mRNA, *e.g.*, a transcript of a sequence that encodes a protein expressed in one or more modules or biological converter switches as described herein. For convenience, such a target mRNA is also referred to herein as an mRNA to be silenced or translationally regulated. Such a sequence is also referred to as a target sequence. As used herein, the phrase "mediates RNAi" refers to the ability to silence, in a sequence specific manner, a target RNA molecule or sequence. While not wishing to be bound by theory, it is believed that silencing uses the RNAi machinery or process and a guide RNA, *e.g.*, an sRNA agent of 21 to 23 nucleotides.

[00249] In other embodiments of the aspects described herein, RNA molecules for use in the modules and biological circuit chemotactic converters described herein comprise natural or engineered microRNA sequences. Also provided herein are references and resources, such as

programs and databases found on the World Wide Web, that can be used for obtaining information on endogenous microRNAs and their expression patterns, as well as information in regard to cognate microRNA sequences and their properties.

[00250] Mature microRNAs (also referred to as miRNAs) are short, highly conserved, endogenous non-coding regulatory RNAs (18 to 24 nucleotides in length), expressed from longer transcripts (termed “pre-microRNAs”) encoded in animal, plant and virus genomes, as well as in single-celled eukaryotes. Endogenous miRNAs found in genomes regulate the expression of target genes by binding to complementary sites, termed herein as “microRNA target sequences,” in the mRNA transcripts of target genes to cause translational repression and/or transcript degradation. miRNAs have been implicated in processes and pathways such as development, cell proliferation, apoptosis, metabolism and morphogenesis, and in diseases including cancer (S. Griffiths-Jones *et al.*, “miRBase: tools for microRNA genomics.” *Nuc. Acid. Res.*, 2007; 36, D154-D158). Expression of a microRNA target sequence refers to transcription of the DNA sequence that encodes the microRNA target sequence to RNA. In some embodiments, a microRNA target sequence is operably linked to or driven by a promoter sequence. In some embodiments, a microRNA target sequence comprises part of another sequence that is operably linked to a promoter sequence, such as a sequence encoding an output product or a recombinase, and is said to be linked to, attached to, or fused to, the sequence encoding the output product or the recombinase.

[00251] The way microRNA and their targets interact in animals and plants is different in certain aspects. Translational repression is thought to be the primary mechanism in animals, with transcript degradation the dominant mechanism for plant target transcripts. The difference in mechanisms lies in the fact that plant miRNA exhibits perfect or nearly perfect base pairing with the target but in the case of animals, the pairing is rather imperfect. Also, miRNAs in plants bind to their targets within coding regions cleaving at single sites, whereas most of the miRNA binding sites in animals are in the 3' un-translated regions (UTR). In animals, functional miRNA:miRNA target sequence duplexes are found to be more variable in structure and they contain only short complementary sequence stretches, interrupted by gaps and mismatches. In animal miRNA: miRNA target sequence interactions, multiplicity (one miRNA targeting more than one gene) and cooperation (one gene targeted by several miRNAs) are very common but rare in the case of plants. All these make the approaches in miRNA target prediction in plants and animals different in details (V. Chandra *et al.*, “MTar: a computational microRNA target prediction architecture for human transcriptome.” *BMC Bioinformatics* 2010, 11(Suppl 1):S2).

[00252] Experimental evidence shows that the miRNA target sequence needs enough complementarities in either the 3' end or in the 5' end for its binding to a miRNA. Based on these complementarities of miRNA: miRNA target sequence target duplex, the miRNA target sequence can be divided into three main classes. They are the 5' dominant seed site targets (5' seed-only), the 5' dominant canonical seed site targets (5' dominant) and the 3' complementary seed site targets (3'

canonical). The 5' dominant canonical targets possess high complementarities in 5' end and a few complementary pairs in 3' end. The 5' dominant seed-only targets possess high complementarities in 5' end (of the miRNA) and only a very few or no complementary pairs in 3' end. The seed-only sites have a perfect base pairing to the seed portion of 5' end of the miRNA and limited base pairing to 3' end of the miRNA. The 3' complimentary targets have high complementarities in 3' end and insufficient pairings in 5' end. The seed region of the miRNA is a consecutive stretch of seven or eight nucleotides at 5' end. The 3' complementary sites have an extensive base pairing to 3' end of the miRNA that compensate for imperfection or a shorter stretch of base pairing to a seed portion of the miRNA. All of these site types are used to mediate regulation by miRNAs and show that the 3' complimentary class of target site is used to discriminate among individual members of miRNA families *in vivo*. A genome-wide statistical analysis shows that on an average one miRNA has approximately 100 evolutionarily conserved target sites, indicating that miRNAs regulate a large fraction of protein-coding genes.

[00253] At present, miRNA databases include miRNAs for human, *Caenorhabditis elegans*, *D. melanogaster*, *Danio rerio* (zebrafish), *Gallus gallus* (chicken), and *Arabidopsis thaliana*. miRNAs are even present in simple multicellular organisms, such as poriferans (sponges) and cnidarians (starlet sea anemone). Many of the bilaterian animal miRNAs are phylogenetically conserved; 55% of *C. elegans* miRNAs have homologues in humans, which indicates that miRNAs have had important roles throughout animal evolution. Animal miRNAs seem to have evolved separately from those in plants because their sequences, precursor structure and biogenesis mechanisms are distinct from those in plants (Kim VN *et al.*, "Biogenesis of small RNAs in animals," Nat Rev Mol Cell Biol. 2009 Feb;10(2):126-39).

[00254] miRNAs useful as components and output products for designing the modules and biological circuit chemotactic converters described herein can be found at a variety of databases as known by one of skill in the art, such as those described at "miRBase: tools for microRNA genomics," Nuc. Acid. Res., 2007: 36 (Database Issue), D154-D158; "miRBase: microRNA sequences, targets and gene nomenclature," Nuc. Acid. Res., 2006 34 (Database Issue):D140-D144; and "The microRNA Registry," Nuc. Acid. Res., 2004 32 (Database Issue):D109-D111), which are incorporated herein in their entirety by reference.

[00255] Accordingly, in some embodiments of the aspects described herein, a module or biological circuit chemotactic converter can further comprise a sequence encoding an RNA molecule, such as an iRNA molecule or microRNA molecule. In such embodiments, the sequence encoding the RNA molecule can be operably linked to a promoter sequence, or comprise part of another sequence, such as a sequence encoding a protein output. In those embodiments where the RNA molecule comprises part of, is linked to, attached to, or fused to, the sequence encoding, *e.g.*, an output product, transcription of the sequence results in expression of both the mRNA of the output product and expression of the RNA molecule.

Definitions

[00256] The terms “nucleic acids” and “nucleotides” refer to naturally occurring or synthetic or artificial nucleic acid or nucleotides. The terms “nucleic acids” and “nucleotides” comprise deoxyribonucleotides or ribonucleotides or any nucleotide analogue and polymers or hybrids thereof in either single- or doublestranded, sense or antisense form. As will also be appreciated by those in the art, many variants of a nucleic acid can be used for the same purpose as a given nucleic acid. Thus, a nucleic acid also encompasses substantially identical nucleic acids and complements thereof. Nucleotide analogues include nucleotides having modifications in the chemical structure of the base, sugar and/or phosphate, including, but not limited to, 5-position pyrimidine modifications, 8-position purine modifications, modifications at cytosine exocyclic amines, substitution of 5-bromo-uracil, and the like; and 2'-position sugar modifications, including but not limited to, sugar-modified ribonucleotides in which the 2'-OH is replaced by a group selected from H, OR, R, halo, SH, SR, NH₂, NHR, NR₂, or CN. shRNAs also can comprise non-natural elements such as non-natural bases, *e.g.*, ionosin and xanthine, nonnatural sugars, *e.g.*, 2'-methoxy ribose, or non-natural phosphodiester linkages, *e.g.*, methylphosphonates, phosphorothioates and peptides.

[00257] The term “nucleic acid sequence” or “oligonucleotide” or “polynucleotide” are used interchangeably herein and refers to at least two nucleotides covalently linked together. The term “nucleic acid sequence” is also used inter-changeably herein with “gene,” “cDNA,” and “mRNA,” unless expressly specified otherwise. As will be appreciated by those in the art, the depiction of a single nucleic acid sequence also defines the sequence of the complementary nucleic acid sequence. Thus, a nucleic acid sequence also encompasses the complementary strand of a depicted single strand. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (*e.g.*, degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. As will also be appreciated by those in the art, a single nucleic acid sequence provides a probe that can hybridize to the target sequence under stringent hybridization conditions. Thus, a nucleic acid sequence also encompasses a probe that hybridizes under stringent hybridization conditions. The term “nucleic acid sequence” refers to a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5'-to the 3'-end. It includes chromosomal DNA, self-replicating plasmids, infectious polymers of DNA or RNA and DNA or RNA that performs a primarily structural role. “Nucleic acid sequence” also refers to a consecutive list of abbreviations, letters, characters or words, which represent nucleotides. Nucleic acid sequences can be single stranded or double stranded, or can contain portions of both double stranded and single stranded sequence. The nucleic acid sequence can be DNA, both genomic and cDNA, RNA, or a hybrid, where the nucleic acid sequence can contain combinations of deoxyribo- and ribo- nucleotides, and combinations of bases including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine and isoguanine. Nucleic acid sequences can be obtained by chemical synthesis methods or by recombinant methods. A nucleic acid

sequence will generally contain phosphodiester bonds, although nucleic acid analogs can be included that can have at least one different linkage, *e.g.*, phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages and peptide nucleic acid backbones and linkages in the nucleic acid sequence. Other analog nucleic acids include those with positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Pat. Nos. 5, 235,033 and 5, 034,506, which are incorporated by reference. Nucleic acid sequences containing one or more non-naturally occurring or modified nucleotides are also included within one definition of nucleic acid sequences. The modified nucleotide analog can be located for example at the 5'-end and/or the 3'-end of the nucleic acid sequence. Representative examples of nucleotide analogs can be selected from sugar- or backbone-modified ribonucleotides. It should be noted, however, that also nucleobase- modified ribonucleotides, *i.e.* ribonucleotides, containing a non naturally occurring nucleobase instead of a naturally occurring nucleobase such as uridines or cytidines modified at the 5-position, *e.g.* 5-(2-amino)propyl uridine, 5-bromo uridine; adenosines and guanosines modified at the 8-position, *e.g.* 8-bromo guanosine; deaza nucleotides, *e.g.* 7 deaza-adenosine; O- and N- alkylated nucleotides, *e.g.* N6-methyl adenosine are suitable. The 2' OH- group can be replaced by a group selected from H, OR, R, halo, SH, SR, NH₂, NHR, NR₂ or CN, wherein R is C- C₆ alkyl, alkenyl or alkynyl and halo is F, Cl, Br or I. Modifications of the ribose-phosphate backbone can be done for a variety of reasons, *e.g.*, to increase the stability and half- life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring nucleic acids and analogs can be used; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs can be used. Nucleic acid sequences include but are not limited to, nucleic acid sequence encoding proteins, for example that act as transcriptional repressors, antisense molecules, ribozymes, small inhibitory nucleic acid sequences, for example but not limited to RNAi, shRNAi, siRNA, micro RNAi (mRNAi), antisense oligonucleotides etc.

[00258] The term "oligonucleotide" as used herein refers to an oligomer or polymer of ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) or mimetics thereof, as well as oligonucleotides having non-naturally-occurring portions which function similarly. Such modified or substituted oligonucleotides are often preferred over native forms because of desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for nucleic acid target and increased stability in the presence of nucleases. An oligonucleotide preferably includes two or more nucleomonomers covalently coupled to each other by linkages (*e.g.*, phosphodiester) or substitute linkages.

[00259] In its broadest sense, the term "substantially complementary," when used herein with respect to a nucleotide sequence in relation to a reference or target nucleotide sequence, means a nucleotide sequence having a percentage of identity between the substantially complementary nucleotide sequence and the exact complementary sequence of said reference or target nucleotide sequence of at least 60%, at least 70%, at least 80% or 85%, at least 90%, at least 93%, at least 95%

or 96%, at least 97% or 98%, at least 99% or 100% (the later being equivalent to the term "identical" in this context). For example, identity is assessed over a length of at least 10 nucleotides, or at least 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 or up to 50 nucleotides of the entire length of the nucleic acid sequence to said reference sequence (if not specified otherwise below). Sequence comparisons are carried out using default GAP analysis with the University of Wisconsin GCG, SEQWEB application of GAP, based on the algorithm of Needleman and Wunsch (Needleman and Wunsch (1970) J Mol. Biol. 48: 443-453; as defined above). A nucleotide sequence "substantially complementary" to a reference nucleotide sequence hybridizes to the reference nucleotide sequence under low stringency conditions, preferably medium stringency conditions, most preferably high stringency conditions (as defined above).

[00260] In its broadest sense, the term "substantially identical," when used herein with respect to a nucleotide sequence, means a nucleotide sequence corresponding to a reference or target nucleotide sequence, wherein the percentage of identity between the substantially identical nucleotide sequence and the reference or target nucleotide sequence is at least 60%, at least 70%, at least 80% or 85%, at least 90%, at least 93%, at least 95% or 96%, at least 97% or 98%, at least 99% or 100% (the later being equivalent to the term "identical" in this context). For example, identity is assessed over a length of 10-22 nucleotides, such as at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 or up to 50 nucleotides of a nucleic acid sequence to said reference sequence (if not specified otherwise below). Sequence comparisons are carried out using default GAP analysis with the University of Wisconsin GCG, SEQWEB application of GAP, based on the algorithm of Needleman and Wunsch (Needleman and Wunsch (1970) J Mol. Biol. 48: 443-453; as defined above). A nucleotide sequence that is "substantially identical" to a reference nucleotide sequence hybridizes to the exact complementary sequence of the reference nucleotide sequence (*i.e.* its corresponding strand in a double-stranded molecule) under low stringency conditions, preferably medium stringency conditions, most preferably high stringency conditions (as defined above). Homologues of a specific nucleotide sequence include nucleotide sequences that encode an amino acid sequence that is at least 24% identical, at least 35% identical, at least 50% identical, at least 65% identical to the reference amino acid sequence, as measured using the parameters described above, wherein the amino acid sequence encoded by the homolog has the same biological activity as the protein encoded by the specific nucleotide. The term "substantially non-identical" refers to a nucleotide sequence that does not hybridize to the nucleic acid sequence under stringent conditions.

[00261] As used herein, the term "gene" refers to a nucleic acid sequence comprising an open reading frame encoding a polypeptide, including both exon and (optionally) intron sequences. A "gene" refers to coding sequence of a gene product, as well as non-coding regions of the gene product, including 5'UTR and 3'UTR regions, introns and the promoter of the gene product. These definitions generally refer to a single-stranded molecule, but in specific embodiments will also encompass an additional strand that is partially, substantially or fully complementary to the single-

stranded molecule. Thus, a nucleic acid sequence can encompass a double-stranded molecule or a double-stranded molecule that comprises one or more complementary strand(s) or "complement(s)" of a particular sequence comprising a molecule. As used herein, a single stranded nucleic acid can be denoted by the prefix "ss", a double stranded nucleic acid by the prefix "ds", and a triple stranded nucleic acid by the prefix "ts."

[00262] The term "nucleic acid construct" as used herein refers to a nucleic acid at least partly created by recombinant methods. The term "DNA construct" refers to a polynucleotide construct consisting of deoxyribonucleotides. The construct can be single or double stranded. The construct can be circular or linear. A person of ordinary skill in the art is familiar with a variety of ways to obtain and generate a DNA construct. Constructs can be prepared by means of customary recombination and cloning techniques as are described, for example, in Maniatis T, Fritsch EF and Sambrook J (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor (NY); Silhavy et al. (1984) *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY); Ausubel et al. (1987) *Current Protocols in Molecular Biology*, Greene Publishing Assoc and Wiley Interscience; Gelvin et al. (Eds) (1990) *Plant Molecular Biology Manual*; Kluwer Academic Pub-lisher, Dordrecht, The Netherlands.

[00263] The terms "polypeptide", "peptide", "oligopeptide", "polypeptide", "gene product", "expression product" and "protein" are used interchangeably herein to refer to a polymer or oligomer of consecutive amino acid residues.

[00264] The term "operable linkage" or "operably linked" are used interchangeably herein, are to be understood as meaning, for example, the sequential arrangement of a regulatory element (*e.g.* a promoter) with a nucleic acid sequence to be expressed and, if appropriate, further regulatory elements (such as, *e.g.*, a terminator) in such a way that each of the regulatory elements can fulfill its intended function to allow, modify, facilitate or otherwise influence expression of the linked nucleic acid sequence. The expression may result depending on the arrangement of the nucleic acid sequences in relation to sense or antisense RNA. To this end, direct linkage in the chemical sense is not necessarily required. Genetic control sequences such as, for example, enhancer sequences, can also exert their function on the target sequence from positions which are further away, or indeed from other DNA molecules. In some embodiments, arrangements are those in which the nucleic acid sequence to be expressed recombinantly is positioned behind the sequence acting as promoter, so that the two sequences are linked covalently to each other. The distance between the promoter sequence and the nucleic acid sequence to be expressed recombinantly can be any distance, and in some embodiments is less than 200 base pairs, especially less than 100 base pairs, less than 50 base pairs. In some embodiments, the nucleic acid sequence to be transcribed is located behind the promoter in such a way that the transcription start is identical with the desired beginning of the chimeric RNA of the invention. Operable linkage, and an expression construct, can be generated by means of customary recombination and cloning techniques as described (*e.g.*, in Maniatis T, Fritsch EF and Sambrook J

(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor (NY); Silhavy et al. (1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY); Ausubel et al. (1987) Current Protocols in Molecular Biology, Greene Publishing Assoc and Wiley Interscience; Gelvin et al. (Eds) (1990) Plant Molecular Biology Manual; Kluwer Academic Publisher, Dordrecht, The Netherlands). However, further sequences may also be positioned between the two sequences. The insertion of sequences may also lead to the expression of fusion proteins, or serves as ribosome binding sites. In some embodiments, the expression construct, consisting of a linkage of promoter and nucleic acid sequence to be expressed, can exist in a vector integrated form and be inserted into a plant genome, for example by transformation.

[00265] The terms "promoter," "promoter element," or "promoter sequence" are equivalents and as used herein, refers to a DNA sequence which when operatively linked to a nucleotide sequence of interest is capable of controlling the transcription of the nucleotide sequence of interest into mRNA. A promoter is typically, though not necessarily, located 5' (*i.e.*, upstream) of a nucleotide sequence of interest (*e.g.*, proximal to the transcriptional start site of a structural gene) whose transcription into mRNA it controls, and provides a site for specific binding by RNA polymerase and other transcription factors for initiation of transcription. A polynucleotide sequence is "heterologous to" an organism or a second polynucleotide sequence if it originates from a foreign species, or, if from the same species, is modified from its original form. For example, a promoter operably linked to a heterologous coding sequence refers to a coding sequence from a species different from that from which the promoter was derived, or, if from the same species, a coding sequence which is not naturally associated with the promoter (*e.g.* a genetically engineered coding sequence or an allele from a different ecotype or variety). Suitable promoters can be derived from genes of the host cells where expression should occur or from pathogens for the host cells (*e.g.*, tissue promoters or pathogens like viruses).

[00266] If a promoter is an "inducible promoter", as defined herein, then the rate of transcription is modified in response to an inducing agent or inducer. In contrast, the rate of transcription is not regulated by an inducer if the promoter is a constitutive promoter. The term "constitutive" when made in reference to a promoter means that the promoter is capable of directing transcription of an operably linked nucleic acid sequence in the absence of a stimulus (*e.g.*, heat shock, chemicals, agents, light, etc.). Typically, constitutive promoters are capable of directing expression of a nucleic acid sequence in substantially any cell and any tissue. In contrast, the term "regulateable" or "inducible" promoter referred to herein is one which is capable of directing a level of transcription of an operably linked nucleic acid sequence in the presence of a stimulus (*e.g.*, heat shock, chemicals, light, agent etc.) which is different from the level of transcription of the operably linked nucleic acid sequence in the absence of the stimulus.

[00267] A promoter can be regulated in a tissue-specific or tissue preferred manner such that it is only active in transcribing the associated coding region in a specific tissue type(s). The term "tissue specific" as it applies to a promoter refers to a promoter that is capable of directing selective expression of a nucleotide sequence of interest to a specific type of tissue (*e.g.*, liver) in the relative absence of expression of the same nucleotide sequence of interest in a different type of tissue (*e.g.*, kidney). Tissue specificity of a promoter may be evaluated by, for example, operably linking a reporter gene to the promoter sequence to generate a reporter construct, introducing the reporter construct into the genome of an organism, *e.g.* an animal model such that the reporter construct is integrated into every tissue of the resulting transgenic animal, and detecting the expression of the reporter gene (*e.g.*, detecting mRNA, protein, or the activity of a protein encoded by the reporter gene) in different tissues of the transgenic animal. The detection of a greater level of expression of the reporter gene in one or more tissues relative to the level of expression of the reporter gene in other tissues shows that the promoter is specific for the tissues in which greater levels of expression are detected. The term "cell type specific" as applied to a promoter refers to a promoter, which is capable of directing selective expression of a nucleotide sequence of interest in a specific type of cell in the relative absence of expression of the same nucleotide sequence of interest in a different type of cell within the same tissue. The term "cell type specific" when applied to a promoter also means a promoter capable of promoting selective expression of a nucleotide sequence of interest in a region within a single tissue. Cell type specificity of a promoter may be assessed using methods well known in the art, *e.g.*, GUS activity staining or immunohistochemical staining. The term "minimal promoter" as used herein refers to the minimal nucleic acid sequence comprising a promoter element while also maintaining a functional promoter. A minimal promoter may comprise an inducible, constitutive or tissue-specific promoter.

[00268] The term "expression" as used herein refers to the biosynthesis of a gene product, preferably to the transcription and/or translation of a nucleotide sequence, for example an endogenous gene or a heterologous gene, in a cell. For example, in the case of a heterologous nucleic acid sequence, expression involves transcription of the heterologous nucleic acid sequence into mRNA and, optionally, the subsequent translation of mRNA into one or more polypeptides. Expression also refers to biosynthesis of an RNA molecule, such as an iRNA molecule, which refers to expression and transcription of an iRNAi molecule such as siRNA, shRNA, microRNA, and antisense DNA but does not require translation to polypeptide sequences. The term "expression construct" and "nucleic acid construct" as used herein are synonyms and refer to a nucleic acid sequence capable of directing the expression of a particular nucleotide sequence, such as the heterologous target gene sequence in an appropriate host cell (*e.g.*, a prokaryotic cell, eukaryotic cell, or mammalian cell). If translation of the desired heterologous target gene is required, it also typically comprises sequences required for proper translation of the nucleotide sequence. The coding region may code for a protein of interest but may also code for a functional RNA of interest, for example antisense RNA, dsRNA, or a nontranslated

RNA, in the sense or antisense direction. The nucleic acid construct as disclosed herein can be chimeric, meaning that at least one of its components is heterologous with respect to at least one of its other components.

[00269] The term "leakiness" or "leaky" as used in reference to "promoter leakiness" refers to some level of expression of the nucleic acid sequence which is operatively linked to the promoter, even when the promoter is not intended to result in expression of the nucleic acid sequence (*i.e.* when the promoter is in the "off" state, a background level of expression of the nucleic acid sequence which is operatively linked to such promoter exists). In one illustrative example using inducible promoters, for example a Tet-on promoter, a leaky promoter is where some level of the nucleic acid sequence expression (which is operatively linked to the Tet-on promoter) still occurs in the absence of the inducer agent, tetracycline. Typically, most inducible promoters and tissue-specific promoters have approximately 10%-30% or 10-20% unintended or background nucleic acid sequence expression when the promoter is not active, for example, the background of leakiness of nucleic acid sequence expression is about 10%-20% or about 10-30%. As an illustrative example using a tissue-specific promoter, a "leaky promoter" is one in which expression of the nucleic acid sequence occurs in tissue where a tissue-specific promoter is not active, *i.e.* expression occurs in a non-specific tissue. Stated in another way using a kidney-specific promoter as an example; if at least some level of the nucleic acid sequence expression occurs in at least one tissue other than the kidney, where the nucleic acid sequence is operably linked to a kidney specific promoter, the kidney specific promoter would be considered a leaky promoter

[00270] The term "enhancer" refers to a cis-acting regulatory sequence involved in the transcriptional activation of a nucleic acid sequence. An enhancer can function in either orientation and can be upstream or downstream of the promoter. As used herein, the term "gene product(s)" is used to refer to include RNA transcribed from a gene, or a polypeptide encoded by a gene or translated from RNA. A protein and/or peptide or fragment thereof can be any protein of interest, for example, but not limited to; mutated proteins; therapeutic proteins; truncated proteins, wherein the protein is normally absent or expressed at lower levels in the cell. Proteins can also be selected from a group comprising; mutated proteins, genetically engineered proteins, peptides, synthetic peptides, recombinant proteins, chimeric proteins, antibodies, midibodies, tribodies, humanized proteins, humanized antibodies, chimeric antibodies, modified proteins and fragments thereof.

[00271] As used herein, the term "comprising" means that other elements can also be present in addition to the defined elements presented. The use of "comprising" indicates inclusion rather than limitation. Accordingly, the terms "comprising" means "including principally, but not necessary solely". Furthermore, variation of the word "comprising", such as "comprise" and "comprises", have correspondingly the same meanings. The term "consisting essentially of" means "including principally, but not necessary solely at least one", and as such, is intended to mean a "selection of one or more, and in any combination". Stated another way, the term "consisting essentially of" means that

an element can be added, subtracted or substituted without materially affecting the novel characteristics of the invention. This applies equally to steps within a described method as well as compositions and components therein. In other embodiments, the inventions, compositions, methods, and respective components thereof, described herein are intended to be exclusive of any element not deemed an essential element to the component, composition or method ("consisting of"). For example, a biological converter switch that comprises a sequence encoding a recombinase and a recombinase recognition sequence encompasses both the recombinase and a recombinase recognition sequence of a larger sequence. By way of further example, a composition that comprises elements A and B also encompasses a composition consisting of A, B and C.

[00272] As used in this specification and the appended claims, the singular forms "a," "an," and "the" include plural references unless the context clearly dictates otherwise. Thus for example, references to "the method" includes one or more methods, and/or steps of the type described herein and/or which will become apparent to those persons skilled in the art upon reading this disclosure and so forth.

[00273] It is understood that the foregoing detailed description and the following examples are illustrative only and are not to be taken as limitations upon the scope of the invention. Various changes and modifications to the disclosed embodiments, which will be apparent to those of skill in the art, may be made without departing from the spirit and scope of the present invention. Further, all patents, patent applications, publications, and websites identified are expressly incorporated herein by reference for the purpose of describing and disclosing, for example, the methodologies described in such publications that might be used in connection with the present invention. These publications are provided solely for their disclosure prior to the filing date of the present application. Nothing in this regard should be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or for any other reason. All statements as to the date or representation as to the contents of these documents are based on the information available to the applicants and do not constitute any admission as to the correctness of the dates or contents of these documents.

[00274] The present invention can be defined in any of the following numbered paragraphs:

1. A biological circuit chemotactic converter, the biological circuit chemotactic converter comprising: at least one input module (IM)_n, at least one genetic toggle switch (TS)_n, at least one logic module (LM)_n, and at least two sensor modules (SM_A and SM_B)_{n+1};
 wherein the input module comprises an inducible promoter sequence (iP_A) operably linked to a repressor sequence encoding a repressor protein (R_A);
 wherein the genetic toggle switch comprises a first repressible promoter sequence (rP₁) that drives expression of a second repressor sequence (R₂), and a second repressible promoter sequence (rP₂) that drives expression of a first repressor sequence (R₁), (rP₁-R₂ and rP₂-R₁);
 wherein the first and second repressor sequences encode first and second repressors proteins,

and wherein the first repressible promoter sequence is inhibited by the first repressor protein and the second repressible promoter sequence is inhibited by the second repressor protein; wherein the logic module comprises a repressible promoter sequence (rP_B) operably linked to a repressor sequence (R_B);

wherein each sensor module comprises a repressible promoter sequence (rP_C and rP_D) operably linked to a nucleic acid sequence encoding a sensor molecule (sensor A and sensor B);

and wherein $n \geq 1$.

2. The biological circuit chemotactic converter of paragraph 1, wherein the inducible promoter sequence of the input module is induced by a biological agent, a chemical agent, a metal ion, a toxin, or a pollutant.
3. The biological circuit chemotactic converter of any of paragraphs 1-2, wherein the repressor protein encoded by the first repressor sequence of the toggle switch (R_1) and the repressor protein encoded by the input module (R_A) are the same repressor protein.
4. The biological circuit chemotactic converter of any of paragraphs 1-3, wherein the repressor proteins encoded by the logic module and the input module are different repressor proteins.
5. The biological circuit chemotactic converter of any of paragraphs 1-4, wherein the repressor proteins encoded by the logic module and the genetic toggle switch are different repressor proteins.
6. The biological circuit chemotactic converter of any of paragraphs 1-5, wherein the second repressible promoter sequence of the genetic toggle switch (rP_2), the repressible promoter sequence of the logic module (rP_B), and the repressible promoter sequence of one sensor module (rP_D) are repressed by the same repressor protein.
7. The biological circuit chemotactic converter of any of paragraphs 1-6, wherein the second repressible promoter sequence of the genetic toggle switch (rP_2), the repressible promoter sequence of the logic module (rP_B), and the repressible promoter sequence of one sensor module (rP_D) comprise the same repressible promoter sequence.
8. The biological circuit chemotactic converter of any of paragraphs 1-7, wherein the repressible promoter sequence of one sensor module (rP_C) is repressed by the repressor encoded by the logic module (R_B).
9. The biological circuit chemotactic converter of any of paragraphs 1-8, wherein the at least two sensor modules encode for different sensor molecules.
10. The biological circuit chemotactic converter of any of paragraphs 1-9, wherein n is an integer value between and including 1 and 100.

11. The biological circuit chemotactic converter of any of paragraphs 1-10, wherein n is an integer value between and including 1 and 10.
12. The biological circuit chemotactic converter of any of paragraphs 1-11, wherein n is an integer value selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10.
13. The biological circuit chemotactic converter of any of paragraphs 1-12, further comprising at least one ribosome binding sequence.
14. The biological circuit chemotactic converter of any of paragraphs 1-13, further comprising at least one terminator sequence.
15. The biological circuit chemotactic converter of any of paragraphs 1-14, further comprising at least one degradation tag sequence.
16. The biological circuit chemotactic converter of any of paragraphs 1-15, further comprising at least one sequence encoding an iRNA molecule specific for at least one protein encoded by the biological circuit chemotactic converter.
17. The biological circuit chemotactic converter of any of paragraphs 1-16, wherein at least one repressor protein is an engineered zinc-finger protein.
18. The biological circuit chemotactic converter of any of paragraphs 1-17, where the promoter sequence of the inducible and repressible promoters is selected from any of the promoter sequences of SEQ ID NOs: 1-7, SEQ ID NOs: 167-843, and SEQ ID NOs: 1009-1010.
19. The biological circuit chemotactic converter of any of paragraphs 1-18, where the sensor molecule encoded by the nucleic acid sequence of the sensor domain is a chemotaxis receptor or chemotaxis sensor selected from Tables 19 and 20.
20. A system comprising at least one biological circuit chemotactic converter of any of paragraphs 1-19.
21. A method of modulating chemotaxis in a biological system using the biological circuit chemotactic converter or system thereof of any of paragraphs 1-20.
22. The method of paragraph 21, wherein the biological system is a cell.
23. The method of paragraph 22, wherein the biological system is an artificial cell, a prokaryotic cell, or a eukaryotic cell.
24. Any of the above-described paragraphs wherein the biological circuit chemotactic converters, and methods of use thereof consist essentially of the specified components.

EXAMPLES

[00275] Synthetic biologists have engineered a wide range of functionality into living systems, including the genetic toggle switch¹ and repressilator², genetic switches^{1, 3-8}, oscillators^{2, 9-11}, digital logic evaluators^{12, 13}, counters¹⁴, filters¹⁵⁻¹⁷, sensors¹⁸⁻²⁰, and cell-cell communicators^{16, 18}, with varying degrees of success. These functionalities have been demonstrated in proof-of-principle applications, such as in bacteria that invade cancer cells²¹, engineered bacteriophages that break up biofilms²² or enhance antibiotic treatments²³, and synthetic microbial pathways that enable the production of antimalarial drug precursors²⁴. However, for all of its successes, many more challenges remain in advancing synthetic biology from low-level gene circuitry to higher-order networks. Herein, we describe the development of novel and useful biological circuit chemotactic converters for use as, and as part of, engineered biological circuits, and the utility of such biological circuit chemotactic converters in various application areas.

Synthetic Gene Networks

[00276] The engineering of mechanical, electrical, and chemical systems is enabled by frameworks for handling complexity and reliable means of probing and manipulating system states—tools that are largely lacking in the engineering of biology. Developing properly functioning biological circuits can involve complicated protocols for DNA construction, rudimentary model-guided and rational design, and repeated rounds of trial and error followed by fine-tuning. Limitations in characterizing kinetic processes and interactions between synthetic components and other unknown constituents *in vivo* make troubleshooting frustrating and prohibitively time-consuming.

[00277] Some advances have been made in streamlining the physical construction of artificial biological systems, in the form of resources and methods for building larger engineered DNA systems from smaller defined parts²⁵⁻²⁸. Additionally, large-scale DNA sequencing and synthesis technologies are enabling researchers to directly program whole genes, genetic circuits and eventually genomes, as well as to re-encode DNA sequences with optimal codons and minimal restriction sites.

[00278] Despite certain advances in molecular construction, building synthetic gene networks that function as desired remains extremely challenging. Achieving robustness, an important property of engineered systems, requires careful incorporation of the concepts of feedback, redundancy, and decoupling²⁹. For example, robustness in a synthetic gene oscillator composed of linked negative and positive feedback loops was found to be dependent on a time delay in the negative feedback loop¹⁰. To achieve robustness, synthetic biologists need to adapt designs from natural systems, incorporate feedback and multiple redundant circuits, and seek out evolutionary methods to refine circuit performance. Accelerated, large-scale diversification³⁰ and the use of characterized component libraries in conjunction with *in silico* models for *a priori* design²⁷ help to fine-tune network performance toward the desired output. However, synthetic biologists are often limited by a dearth of interoperable biological parts and reliable means of characterizing information flow through engineered gene networks.

[00279] Thus, as described herein, we have developed next-generation synthetic biology modules and circuits that (1) advance and expand the toolkit of available parts and modules; (2) provide probes for reliably quantifying state values of all the internal nodes of synthetic (and natural) biomolecular systems, and (3) create platforms for characterizing interactions between higher-order networks that have desired functionality for satisfying practical applications. Herein, we describe the development of several next-generation synthetic gene networks and application areas for their use.

Interoperable Parts for Synthetic Gene Networks

[00280] While there has been no shortage of novel circuit topologies to construct, limitations in the number of interoperable and well-characterized parts have constrained the development of more complex biological systems^{25, 27, 31}. The situation is complicated by the many potential interactions between biological parts, which are derived from a variety of sources within different cellular backgrounds, and which are not well understood or characterized. As a result, the majority of synthetic circuits are still constructed *ad hoc* from a small number of commonly used components with a significant amount of trial and error.

[00281] The biological circuit chemotactic converters we describe herein expand the toolkit of parts for synthetic biological circuits. Promoter libraries with defined transcriptional activities can be created and characterized, plugged into *in silico* models, and used to develop novel synthetic gene networks that function without significant *post-hoc* adjustments^{27, 32-35}. However, libraries of other biological components, such as transcription factors and recombinases, remain lacking. Automated techniques for assembling and testing multiple potentially interacting parts together within cells, as described herein, can greatly assist in the characterization and development of useful part libraries.

[00282] We have discovered and show that nucleic-acid-based parts, such as RNA devices, are essential tools for libraries of interoperable parts and devices, since they can be rationally programmed based on sequence specificity^{6, 36, 37}. Novel circuit interconnections can be established using small RNA molecules, such as siRNAs, to control expression of specific components. Further, nucleic-acid-based parts can also be designed to function orthogonally to the wildtype cellular machinery, thus enabling dedication of defined cellular resources to engineered functions³⁸⁻⁴⁰. Artificial codons and unnatural amino acids, which serve to infuse these biological toolkits with even more orthogonal parts, have enabled new methods for studying existing proteins and the realization of proteins with novel functions⁴¹.

[00283] As described herein, we have found that engineered zinc fingers constitute a flexible system for targeting specific DNA sequences, which significantly expands the synthetic biology toolkit for uses in, for example, targeted recombination, control of transcriptional activity, and making circuit interconnections. This technology has seen the greatest use in the creation of zinc-finger nucleases that generate targeted double-strand breaks for genomic modifications⁴². Zinc fingers have also been used to create artificial transcription factors by fusing zinc-finger domains with activation or repression domains^{43, 44}. As described herein, libraries of externally controllable transcriptional

activators or repressors can be created by engineering protein or RNA ligand-responsive regulators, which control the transcription or translation of zinc-finger-based artificial transcription factors themselves¹⁹. These libraries enable the construction of basic circuits, such as genetic toggle switches^{27,45}, as well as more complex gene networks, such as the biological converter switches, described herein. In fact, several of the higher-order networks we describe herein rely on having multiple reliable and interoperable transcriptional activators and repressors for proper functioning. The greater availability of such transcriptional activators and repressors permits greater flexibility and sensitivity of the biological circuit chemotactic converters described herein.

Probes for Synthetic Gene Networks

[00284] Significant advances have been made in the development of new technologies for manipulating biological systems and probing their internal states. At the single-molecule level, for instance, optical tweezers and atomic force microscopes have provided alternative, direct ways to probe the biophysical states of single DNA, RNA, and protein molecules as they undergo conformational changes and other dynamical processes⁴⁶⁻⁵⁰. However, we lack similar tools for tracking the *in vivo* operation of synthetic gene circuits in a high-throughput fashion. Making dynamic measurements of biological networks involves placing sensors at multiple internal nodes, akin to how current and voltage are measured in electrical systems. Furthermore, external manipulation of synthetic biomolecular systems is typically accomplished by the addition of chemical inducers, which can suffer from crosstalk⁵¹, be difficult to remove, and be consumed over time. As a result, inputs are often troublesome to control dynamically. Thus, there is a need in the art to develop and characterize probes for such uses.

[00285] Microfluidic devices enable increasingly precise manipulation and measurement of cells, especially since inputs can be modulated over time⁵². These systems allow the rapid addition and removal of chemical inducers, enabling more sophisticated, time-dependent inputs than conventional step functions, while also allowing researchers to track single cells for long periods of time. These developments make possible the wider use in synthetic biology of well-established engineering approaches for analyzing circuits and other systems. For example, frequency-domain analysis, a technique used commonly in electrical engineering, can be utilized with microfluidics to characterize the transfer functions and noise behaviour of synthetic biological circuits⁵³⁻⁵⁵, and small-signal linearization of nonlinear gene circuits can be achieved by applying oscillatory perturbations with microfluidics and measuring responses at the single-cell level^{54,55}.

[00286] Indeed, microfluidics provides one useful platform for perturbing synthetic gene circuits with well-controlled inputs and observing the outputs in high-resolution fashion. However, without the proper “sensors” (*i.e.*, for quantitatively and simultaneously probing all the internal nodes of a given gene circuit), microfluidic technology alone is not sufficient to bring full, engineering-like characterization to synthetic gene networks.

[00287] Thus far, probes enabling quantitative measurements of synthetic gene circuits have primarily focused on the transcriptome, such as the use of fluorescent proteins for *in vivo* quantitation of promoter activity or protein expression. With the advent of novel mass-spectrometry-based methods that provide global, absolute protein concentrations in cells⁵⁶, quantitative transcriptome data can now be merged with proteome data, improving our ability to characterize and model the dynamics of synthetic gene networks. Global proteomic data can also assist synthetic biologists in understanding the metabolic burden that artificial circuits place on the cells they inhabit. Further efforts to devise fluorescent-based and other reporters for the simultaneous monitoring of transcriptome and proteome dynamics *in vivo* are needed to close the loop on full circuit accounting. Some promising tools under development include tracking protein function by incorporating unnatural amino acids that exhibit fluorescence^{57, 58}, quantum dots⁵⁹, and radiofrequency-controlled nanoparticles⁶⁰.

[00288] As the field awaits entire-circuit probes, there are, in the meantime, several accessible technologies for increasing the throughput and pace of piece-wise gene-circuit characterization. Recent advances in engineering light-inducible biological parts and systems⁶¹⁻⁶³ have unlocked the potential for optical-based circuit characterization. For instance, by coupling a synthetic gene network of interest to light-inducible systems as well as fluorescent protein outputs, both control and monitoring could be accomplished via the reliable and high-speed optics that are typically associated with fluorescence microscopy. This prospect, particularly in the context of microfluidic devices, can facilitate the focusing of optical inputs and read-outs to single cells.

[00289] Using electrical signals, in lieu of chemical or optical signals, for control and monitoring of biological system also present high-speed advantages. Recently, advances have been made in integrating silicon electronics with lipid bilayers containing transmembrane pores to perform electronic signal conduction⁶⁴. This technology can eventually allow direct communication and control between engineered cells and electronic circuits via ionic flow. The incorporation of these and other technologies to perturb and monitor the *in vivo* performance of synthetic gene networks can enable us to achieve desired functionality faster and more reliably.

Next-Generation Gene Networks

[00290] Advancing synthetic gene circuits into the realm of higher-order networks with programmable functionality is one of the ultimate goals of synthetic biology. Described herein are several next-generation gene circuits and their real-world applications.

Whole-Cell Biosensors and Response Systems

[00291] Programmable cells that act as whole-cell biosensors have been created by interfacing engineered gene networks with the cell's natural regulatory circuitry¹⁸ or with other biological components such as light-responsive elements^{62, 63}. In some aspects, novel or re-engineered sensory modalities and components, such as those described herein, can be developed to expand the range of applications that programmable cells can address. In some embodiments, proteins or RNAs are

engineered to detect a range of small molecules^{91, 92}. In other embodiments, protein-based synthetic signaling cascades are designed by rationally rewiring protein-protein interactions and output responses of prokaryotic two-component signal transduction systems⁹³.

[00292] The detection of electrical signals or production of biological energy (*e.g.*, mimicking the operation of electrical electrocytes⁹⁴) can also be enabled, in some aspects, by incorporating natural or synthetic ion channels into engineered cells. Additionally, magneto-responsive bacteria can play useful roles in environmental and medical applications, in some aspects⁹⁵. For example, in some embodiments, synthetic bacteria, designed to form magnetosomes and seek out cancer cells, can be used to enhance imaging, and in other embodiments, magnetic bacteria can be engineered to interact with nanoparticles to enhance the targeting of cancer cells. Moreover, in some embodiments, the introduction of mechanosensitive ion channels (such as MscL from *M. tuberculosis* and MscS from *Escherichia coli*) can be used to endow designer cells with the ability to detect mechanical forces⁹⁶. Such cells can be useful *in vivo* sensors for studying, in some embodiments, cellular differentiation signals or the effects of external stresses on the body.

[00293] In some aspects, programmable cells possessing novel sensory modules described herein can be integrated with mechanical, electrical, and chemical systems to detect, process, and respond to external stimuli, and be utilized for a variety of environmental and medical applications. For example, in some embodiments, bacteria can be engineered to seek out hazardous chemicals in the environment, perform cleanup, and return to their origin to report on the number of hazardous sites encountered via analysis by microfluidic devices. To achieve such complex tasks, in some embodiments, chemotactic bacteria can be programmed to swim from waypoint to waypoint. In such an example, a dish containing gradients of several chemoattractants constitutes the navigational course (Figure 1A).

[00294] At the core of this aspect is a synthetic gene network made up of a series of sequential toggle switches that control the expression of receptors needed for bacterial chemotaxis towards chemoattractants⁹⁷ (Figure 1B). In such aspects, the programmable cells initially express only a single chemoattractant receptor, and therefore migrate up only one of the chemoattractant gradients⁹⁷. To determine that a waypoint has been achieved, in some embodiments, a threshold-based toggle switch is turned ON upon reaching a sufficiently high concentration of the chemoattractant. When the first toggle switch is ON, production of the first chemoattractant receptor is suppressed and production of a second receptor allowed, resulting in cells swimming up the second chemical gradient. In some embodiments, the ON switch additionally primes the next toggle switch in the series to be switched ON when the second waypoint is reached. When that second toggle flips ON, the previous switch is flipped OFF to ensure that only one chemoattractant is being followed at a time. The final chemoattractant leads the bacteria back to its origin so that the engineered cells complete a multi-stop round trip.

Analog-to-Digital and Digital-to-Analog Converters

[00295] Electrical engineers have utilized digital processing to achieve reliability and flexibility, even though the world in which digital circuits operate is inherently analog. Although synthetic biological circuits are unlikely to match the computing power of digital electronics, circuits inspired by digital and analog electronics can increase the reliability and programmability of biological behaviours.

[00296] In some aspects, biological analog-to-digital converters are provided that translate external analog inputs, such as inducer concentrations or exposure times, into internal digital representations for biological processing. In one embodiment of this aspect, a bank of genetic switches with adjustable thresholds is provided, as depicted in Figure 3A. In some embodiments, the switches are made out of libraries of artificial transcription factors, as described herein. In such embodiments, discretization of analog inputs into levels of digital output is achieved. In some embodiments, depending on the level of analog inputs received by the biological analog-to-digital converter, different pathways are activated. In some aspects, the biological analog-to-digital converter are used in cells, for use as, for example, biosensors in medical and environmental settings. For example, in one embodiment, whole-cell biosensors¹⁸, resident in the gut, are engineered to generate different reporter molecules depending on the detected level of a physiological input, such as gastrointestinal bleeding, which can be measured in a sample, such as a stool sample. In some embodiments, expressing different reporter molecules rather than a continuous gradient of a single reporter molecule is used to yield more reliable and easily interpretable outputs.

[00297] In other aspects, biological digital-to-analog converter switches are provided that translate digital representations back into analog outputs, as represented in Figure 3B. In some embodiments, such biological digital-to-analog converters are used to reliably set internal system states. For example, in one embodiment, instead of fine-tuning transcriptional activity with varying amounts of chemical inducers, a digital-to-analog converter, composed of a bank of genetic switches, each of which is sensitive to a different inducer, is used to provide better control. If each activated switch enabled transcription from promoters of varying strengths ($P_{\text{output},3} > P_{\text{output},2} > P_{\text{output},1}$), then digital combinations of inducers are used to program defined levels of transcriptional activities, as shown in Figure 3B. In some aspects, such biological digital-to-analog converter circuits are used in biotechnology applications. For example, in some embodiments, biological digital-to-analog converter circuits are used where reliable expression of different pathways is needed for programming different modes of operation in engineered cells. In other embodiments, digital-to-analog converters can be used for probing synthetic circuits.

Tunable Filters and Noise Generators

[00298] Fine-tuning the performance of a synthetic gene network typically means re-engineering its components, be it by replacing or mutating its parts. In some aspects described herein, networks are provided, the responses of which are tuned without re-engineering of their parts, such as, in some embodiments, the biological version of a tunable electronic filter, which enable more

sophisticated cellular-based signal processing. Synthetic transcriptional cascades exhibit low-pass filter characteristics¹⁵, and artificial gene circuits with negative autoregulation are capable of pushing the noise spectra of their output to higher frequencies, where it can be filtered by the low-pass characteristics of a downstream gene cascade⁶⁵. In some embodiments, tunable genetic filters with respect to time can be implemented by tuning RNA and/or protein degradation in autoregulated negative feedback circuits, as shown in Figure 4^{53, 66-68}. In certain aspects, these circuits are useful in studying and shaping noise spectra to optimize the performance of artificial gene networks.

[00299] Recently, an externally tunable, bacterial bandpass-filter was described¹⁷, which uses low-pass and high-pass filters in series to derive bandpass activity with respect to enzymes and inducer molecules. These types of filters, when coupled, in certain embodiments, to quorum-sensing modules, can be used for spatial patterning applications^{16, 17}. In some embodiments, these filters can also be extended to complex multicellular pattern formation, by engineering a suite of different cells, each carrying filters that respond to different inputs. In some embodiments of this aspect, synthetic gene circuits based on tunable filters can be used as platforms for studying cellular differentiation and development, as artificial pattern generation is a model for how natural systems form complex structures^{16, 17}.

[00300] Recent developments in stem cell biology have unlocked important, potential roles for synthetic gene networks⁶⁹. First, it has been shown that stochastic fluctuations in protein expression in embryonic stem cells are important for determining differentiation fates⁷⁰. Stochasticity can be harnessed in differentiation to force population-wide heterogeneity and provide system robustness.

[00301] Accordingly, in some embodiments, the effects of stochasticity in stem cell differentiation can be studied with synthetic gene circuits as described herein that act as tunable noise generators. In some embodiments, the mean value and variance of an output can be effectively tuned with two external signals, one for regulating the transcription and the other for regulating the translation of a gene, or in some embodiments, with three external signals, the third for regulating DNA copy number⁷¹. By varying noise levels while keeping mean expression levels constant, in some embodiments of the aspect, the thresholds at which gene expression noise yields beneficial versus detrimental effects on stem cell differentiation can be elucidated⁷².

[00302] The discovery of induced pluripotent stem cells (iPSCs), based on the controlled expression of four transcription factors (OCT4, SOX2, KLF2, MYC) in adult fibroblasts, has created a source of patient-specific progenitor cells for engineering⁶⁹. In some aspects, genetic noise generators and basic control circuits are used to dissect the mechanism for inducing pluripotency in differentiated adult cells, by sequentially controlling the mean expression levels of the four iPSC-dependent transcription factors. In some embodiments, timing circuits²⁷ can be used for higher-efficiency stem cell reprogramming.

[00303] Lineage commitment to trophectoderm, ectoderm, mesoderm, and endoderm pathways are controlled by distinct sets of genes⁷⁰, and many interacting factors, including growth factors, extracellular matrices and mechanical forces, play important roles in cellular differentiation⁷³. Accordingly, in some aspects, synthetic gene cascades can be used to program cellular commitment with increased fidelity, for use in applications in biotechnology and regenerative medicine.

Adaptive Learning Networks

[00304] Endogenous biomolecular networks in bacteria can exhibit anticipatory behaviour for related perturbations in environmental stimuli^{74, 75}. This type of behaviour and the associated underlying design principles can, in certain aspects, be harnessed to endow transcriptional networks with the ability to learn⁷⁶, much like synaptic interconnections between neurons. In some embodiments, a basic design that enable this functionality involves two transcriptional activators (Activator A and Activator B), each of which is expressed in the presence of a different stimulus (Figure 5A). In one embodiment, both transcriptional activators drive the expression of effector proteins (Effector A and Effector B), which control distinct genetic pathways. When both transcriptional factors are active, indicating the simultaneous presence of the two stimuli, a toggle switch is flipped ON, creating an associative memory. Subsequently, if either of the transcription factors is activated, AND logic between the ON toggle switch and one transcriptional activator produces the effector protein that controls the pathways of the other activators. Based on such a design, in other aspects and embodiments, cells are programmed to associate simultaneous inputs and exhibit anticipatory behaviour by activating the pathways of associated stimuli, even in, in some embodiments, the presence of only one of the stimuli.

[00305] In another aspect of a learning network, bacteria can be engineered that are taught “winner-take-all” behaviour in detecting stimuli, similar to cortical neural processing⁷⁷. In such embodiments, bacteria are exposed to different types of chemical stimuli (Inducers A-C) (Figure 5B). In some embodiments, an exogenously added inducer (Inducer “Learn”) acts as a trigger for learning and serves as one input into multiple, independent transcriptional AND gates, which possess secondary inputs for detecting the presence of each of the different chemical stimuli. Each gate drives an individual toggle switch that, when flipped, suppresses the flipping of the other switches. This creates a winner-take-all system in which the presence of the most abundant chemical stimuli is recorded. Furthermore, in other embodiments, the toggle switch outputs are fed as inputs into transcriptional AND gates, which once again possess secondary inputs for detecting the presence of the different stimuli. If these gates drive different fluorescent reporters when activated, then the overall system associates a single type of stimuli with the learning trigger and responds with an output only in the presence of the single type of stimuli in the future. This system can, in other embodiments, be adapted to make chemotactic bacteria that “remember” a particular location or landmark, and only respond to the gradient of one chemoattractant.

[00306] In other aspects of learning networks, synthetic gene circuits are designed to adapt on their own, that is, without external mutagenesis or exogenous nucleic acids. For example, in some embodiments, transcription-based interconnections are dynamically reconfigured based on the expression of DNA recombinases¹⁴. In another aspect, error-prone RNA polymerases, which create mutant RNAs that can be reverse-transcribed and joined back into the genome, based on double-stranded breaks created by zinc-finger nucleases, are used. Specificity for where the mutations would occur is achieved, in some embodiments, by using promoters that are uniquely read by error-prone RNA polymerases, such as in some embodiments, T7 promoters with a T7 error-prone RNA polymerase, and, in some embodiments, zinc-finger nucleases that define where homologous recombination can occur⁷⁸. In such an aspect, mutagenesis frequency can be targeted to specific regions of the genome.

Protein-Based Computational Circuits

[00307] In addition to DNA- and RNA-based circuits, protein-based synthetic systems have the potential to enable flexible and fast computation⁷⁹⁻⁸¹. Protein-based designs can operate on much quicker time scales than genetic circuits because their operation is independent of the transcription and translation machinery. Accordingly, in some aspects described herein, protein-based circuits are described that act as rapidly responding logic gates, smart sensors, or memory elements.

[00308] In some embodiments, synthetic amyloids serve as novel components for epigenetic memory circuits. Li and Lindquist demonstrated, by fusing a yeast prion determinant from Sup35 to the rat glucocorticoid receptor, a transcription factor regulated by steroid hormone, that the state of transcriptional activity from the fused protein can be affected and inherited stably in an epigenetic fashion⁸². Given the increasing number of identified prionogenic proteins⁸³, in some embodiments, amyloid-based memory systems can be created that transmit functionality from one generation to the next (Figure 6). In some embodiments, aggregation can be induced by the transient overexpression of the prionogenic domain (PD), while, in some embodiments, disaggregation can be achieved by expressing protein remodelling factors such as chaperones (HSP104).

[00309] Since genetic (nucleic acid-based) circuits and proteins function on different time scales, in some aspects, synthetic networks that couple both modalities are generated. For example, in some embodiments, the output of protein-based computation can be stored in recombinase-based memory elements^{5, 8, 14}. In some embodiments, the two types of networks are coupled to harness their varied filtering capabilities. As a non-limiting example, the mitogen-activated protein kinase cascade contains both positive-feedback and negative-feedback loops that enable rapid activation followed by deactivation⁸⁴, thus acting like a high-pass filter. As another non-limiting example, transcription- and translation-based networks operate on slower time scales rendering them effective low-pass filters. Therefore, in some aspects, synthetic kinase/phosphatase circuits that in turn drive gene-based networks are used to design bandpass filters.

Intercell Signaling Circuits and Pulse-Based Processing for Genetic Oscillators

[00310] Robust genetic oscillators with tunable periods have been developed via a combination of experimental and computational efforts^{10, 11, 86}. In addition to providing design principles guiding the evolution of naturally-occurring biological clocks and circadian rhythms, these synthetic oscillators can have significant utility in biotechnology applications, such as, in some embodiments, optimizing the synthesis and delivery of biologic drugs. As a non-limiting example, glucocorticoid secretion has a circadian and ultradian pattern of release, resulting in transcriptional pulsing in cells that contain glucocorticoid receptors⁸⁵. Therefore, pulsatile administration of synthetic hormones can have therapeutic benefit compared with synthetic hormones applied in a non-ultradian schedule. In some aspects, the biological converter switches described herein can be used as an alternative to electronic-based periodic drug delivery systems. Accordingly, in some embodiments, bacteria that reside in the human gut are engineered to synthesize an active drug at fixed time intervals. In other embodiments, intercell signalling circuits are developed and implemented for synchronizing and entraining synthetic genetic oscillators^{87, 88}. Such circuits are based, in some embodiments, on modular components from bacterial quorum sensing systems. In some embodiments, light-sensitive^{62, 63} entrainment circuits are engineered for synchronizing mammalian synthetic genetic oscillators. This can aid in the construction of oscillators that can faithfully follow circadian rhythms.

[00311] Spike- or pulse-based processing is present in neurons and has been adapted for use in hybrid computation in electrical systems, where interspike times are viewed as analog parameters and spike counts are viewed as digital parameters⁸⁹. In synthetic gene circuits, spike-based processing can be used in novel methods for encoding information in engineered cells. In some embodiments, instead of transmitting information between cells via absolute levels of quorum-sensing molecules, the frequency of a robust genetic oscillator can be modulated. This is useful, for example, in delivering information over longer distances, as frequency information can be less susceptible to decay over distance than absolute molecule levels. Representing signals in this fashion is analogous to frequency modulation encoding in electrical engineering.

Engineered Circuits for Biological Containment

[00312] Biological containment can be achieved by passive or active techniques. In aspects related to passive biological containment, cells are engineered to be dependent on exogenous supplementation to compensate for gene defects, while in aspects related to active containment, cells are engineered to directly express toxic compounds when located outside their target environments⁹⁰. Genetic counters or timers for programmed cell death can be used as an active containment tool, in some embodiments. In some embodiments, counting circuits can trigger cell suicide after a defined number of cell cycles or a sequence of events. Recently, we developed two designs for synthetic counters – a recombinase-based cascade of memory units and a riboregulated transcriptional cascade – that can be adapted for this purpose¹⁴. In some embodiments, promoters that are cell cycle-dependent and replace output reporter proteins with toxic proteins can be incorporated into the

counters (Figure 7). Circuits of this sort enable cells to be programmed to have limited, prescribed lifetimes.

[00313] In some aspects, redundant circuits are provided herein that implement digital logic allowing for the conditional survival of engineered cells only within their desired environments, and thus reduce the failure rate of biological containment. By developing a broad set of interoperable parts, in some embodiments, multiple layers of control circuits are built for increased reliability. As in electrical and mechanical engineering, quantitative analysis of failure rates in biological systems enable improved systems-level design and robustness of synthetic gene networks. In some embodiments, synthetic containment circuits can be subjected to stressful conditions that lead to increase mutation rates and improper functioning.

Designer Circuits and Systems for Microbiome Engineering

[00314] The human microbiome is fertile ground for the application of engineered organisms as scientific tools and therapeutic agents. There are unique bacterial populations, residing in distinct locations in the human body, that are perturbed in disease states^{98, 99}. Each represents an opportunity for re-engineering the human microbiome and designing targeted therapeutics for a range of conditions, including, in some embodiments, dermatologic, genitourinary, gastrointestinal, and immunologic diseases.

[00315] Recently, there have been examples in which engineered bacteria were designed to infiltrate cellular communities for the purposes of delivering probes, gene circuits, or chemicals^{100, 101}. In a similar fashion, in some aspects, bacteriophages carrying synthetic gene circuits as described herein can be used to transform existing microbiome bacteria with new functionalities. For example, anaerobic bacteria are known to migrate to hypoxic and necrotic regions of solid tumors¹⁰². To enhance the specificity of bacterial cancer therapy, in some embodiments, bacteriophages are designed to infect cancer-targeting bacteria. In some embodiments, these bacteriophages encode conditional expression of chemotherapeutic agents using synthetic logic gates or switches, such as ones that are coupled to environmental sensors.

[00316] Bhatia and colleagues recently developed nanoparticles that perform Boolean logic based on proteolytic activity¹⁰³. In some aspects, viruses that infect tumor cells or bacteria are used to carry synthetic gene circuits, as described herein, that regulate, in a programmable fashion, the expression of enzymes that trigger nanoparticle activity. In some embodiments, such synthetic gene networks are used in targeted therapies against cancer or infectious diseases that exploit the human microbiome.

Switchboard for Dynamically Controlling the Expression of Multiple Genes

[00317] Engineered cells have been used to produce recombinant proteins and chemicals for the biotechnology industry, and one of the major applications of synthetic biology to date has been in enhancing microbial production of biofuels¹⁰⁴ and biomaterials¹⁰⁵⁻¹⁰⁸. Improving production from cells involves numerous engineering decisions related to the entire organism, including, for example,

codon optimization, choosing whether or not to export recombinant proteins¹⁰⁹, rational or evolutionary methods for improving metabolic yields¹¹⁰, and optimization of growth conditions. Often some or all of the genes required for production are non-optimal for bacterial expression and contain repetitive sequences that are unstable in bacterial hosts. Whole-gene synthesis techniques are increasingly being used to optimize coding sequences for recombinant production¹⁰⁸.

[00318] These approaches as well as more traditional knockout techniques, introduce hard-wired changes into the genomes of interest. However, for many industrial and bioprocess applications, for example, there is a need to dynamically modulate and control the expression of multiple genes, depending upon the state of the bioreactor. In some aspects, synthetic switchboards based on the modules, switches and methods described herein can be used. In some embodiments, a synthetic switchboard can tune the expression of many different genes simultaneously and independently. Such a switchboard can comprise, for example, a series of adjustable threshold genetic switches, riboregulators or riboswitches, or a combination thereof, and can be designed to respond, in some embodiments, to different environmental and intracellular variables, such as pH, light intensity, and the metabolic state of the cell. The switchboard design can, in some aspects, integrate novel sensory modalities with tunable, interoperable genetic circuits. Such systems have broad functionality for the biotechnology industry. In some embodiments, a synthetic network is programmed, for example, to shift carbon flux between different pathways depending upon cellular conditions, thereby optimizing the production of biofuels, specialty chemicals, and other materials.

[00319] The novel, next-generation synthetic gene networks described herein advance our understanding of natural systems, provide new biological modules and tools allowing for the construction of more complex systems, and satisfy real-world applications in fields such as medicine, biotechnology, bioremediation, and bioenergy.

References

1. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403, 339-342 (2000).
2. Elowitz, M.B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335-338 (2000).
3. Isaacs, F.J., Hasty, J., Cantor, C.R. & Collins, J.J. Prediction and measurement of an autoregulatory genetic module. *Proc Natl Acad Sci U S A* 100, 7714-7719 (2003).
4. Kramer, B.P. et al. An engineered epigenetic transgene switch in mammalian cells. *Nat Biotechnol* 22, 867-870 (2004).
5. Ham, T.S., Lee, S.K., Keasling, J.D. & Arkin, A.P. A tightly regulated inducible expression system utilizing the *fim* inversion recombination switch. *Biotechnol Bioeng* 94, 1-4 (2006).
6. Deans, T.L., Cantor, C.R. & Collins, J.J. A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell* 130, 363-372 (2007).
7. Ajo-Franklin, C.M. et al. Rational design of memory in eukaryotic cells. *Genes Dev* 21, 2271-2276 (2007).
8. Ham, T.S., Lee, S.K., Keasling, J.D. & Arkin, A.P. Design and construction of a double inversion recombination switch for heritable sequential genetic memory. *PLoS ONE* 3, e2815 (2008).
9. Fung, E. et al. A synthetic gene-metabolic oscillator. *Nature* 435, 118-122 (2005).
10. Stricker, J. et al. A fast, robust and tunable synthetic gene oscillator. *Nature* 456, 516-519 (2008).

11. Tigges, M., Marquez-Lago, T.T., Stelling, J. & Fussenegger, M. A tunable synthetic mammalian oscillator. *Nature* 457, 309-312 (2009).
12. Rinaudo, K. et al. A universal RNAi-based logic evaluator that operates in mammalian cells. *Nat Biotechnol* 25, 795-801 (2007).
13. Win, M.N. & Smolke, C.D. Higher-order cellular information processing with synthetic RNA devices. *Science* 322, 456-460 (2008).
14. Friedland, A.E. et al. Synthetic gene networks that count. *Science* 324, 1199-1202 (2009).
15. Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci U S A* 102, 3581-3586 (2005).
16. Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* 434, 1130-1134 (2005).
17. Sohka, T. et al. An externally tunable bacterial band-pass filter. *Proc Natl Acad Sci U S A* 106, 10135-10140 (2009).
18. Kobayashi, H. et al. Programmable cells: interfacing natural and engineered gene networks. *Proc Natl Acad Sci U S A* 101, 8414-8419 (2004).
19. Bayer, T.S. & Smolke, C.D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat Biotechnol* 23, 337-343 (2005).
20. Win, M.N. & Smolke, C.D. A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proc Natl Acad Sci U S A* 104, 14283-14288 (2007).
21. Anderson, J.C., Clarke, E.J., Arkin, A.P. & Voigt, C.A. Environmentally controlled invasion of cancer cells by engineered bacteria. *J Mol Biol* 355, 619-627 (2006).
22. Lu, T.K. & Collins, J.J. Dispersing biofilms with engineered enzymatic bacteriophage. *Proc Natl Acad Sci U S A* 104, 11197-11202 (2007).
23. Lu, T.K. & Collins, J.J. Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. *Proc Natl Acad Sci U S A* 106, 4629-4634 (2009).
24. Ro, D.K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440, 940-943 (2006).
25. Guido, N.J. et al. A bottom-up approach to gene regulation. *Nature* 439, 856-860 (2006).
26. Shetty, R.P., Endy, D. & Knight, T.F., Jr. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng* 2, 5 (2008).
27. Ellis, T., Wang, X. & Collins, J.J. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat Biotechnol* 27, 465-471 (2009).
28. Czar, M.J., Cai, Y. & Peccoud, J. Writing DNA with GenoCAD. *Nucleic Acids Res* 37, W40-47 (2009).
29. Kitano, H. Biological robustness. *Nat Rev Genet* 5, 826-837 (2004).
30. Wang, H.H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460, 894-898 (2009).
31. Purnick, P.E. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol* 10, 410-422 (2009).
32. Jensen, P.R. & Hammer, K. Artificial promoters for metabolic optimization. *Biotechnol Bioeng* 58, 191-195 (1998).
33. Alper, H., Fischer, C., Nevoigt, E. & Stephanopoulos, G. Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A* 102, 12678-12683 (2005).
34. Hammer, K., Mijakovic, I. & Jensen, P.R. Synthetic promoter libraries--tuning of gene expression. *Trends Biotechnol* 24, 53-55 (2006).
35. Cox, R.S., 3rd, Surette, M.G. & Elowitz, M.B. Programming gene expression with combinatorial promoters. *Mol Syst Biol* 3, 145 (2007).
36. Isaacs, F.J. et al. Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol* 22, 841-847 (2004).
37. Win, M.N., Liang, J.C. & Smolke, C.D. Frameworks for programming biological function through RNA parts and devices. *Chem Biol* 16, 298-310 (2009).
38. Rackham, O. & Chin, J.W. A network of orthogonal ribosome x mRNA pairs. *Nat Chem Biol* 1, 159-166 (2005).
39. Wang, K., Neumann, H., Peak-Chew, S.Y. & Chin, J.W. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat Biotechnol* 25, 770-777 (2007).

40. An, W. & Chin, J.W. Synthesis of orthogonal transcription-translation networks. *Proc Natl Acad Sci U S A* 106, 8477-8482 (2009).
41. Wang, Q., Parrish, A.R. & Wang, L. Expanding the genetic code for biological studies. *Chem Biol* 16, 323-336 (2009).
42. Maeder, M.L. et al. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell* 31, 294-301 (2008).
43. Beerli, R.R., Dreier, B. & Barbas, C.F., 3rd Positive and negative regulation of endogenous genes by designed transcription factors. *Proc Natl Acad Sci U S A* 97, 1495-1500 (2000).
44. Park, K.S. et al. Phenotypic alteration of eukaryotic cells using randomized libraries of artificial transcription factors. *Nat Biotechnol* 21, 1208-1214 (2003).
45. Canton, B., Labno, A. & Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol* 26, 787-793 (2008).
46. Svoboda, K. & Block, S.M. Biological applications of optical forces. *Annu Rev Biophys Biomol Struct* 23, 247-285 (1994).
47. Bustamante, C., Bryant, Z. & Smith, S.B. Ten years of tension: single-molecule DNA mechanics. *Nature* 421, 423-427 (2003).
48. Khalil, A.S. et al. Single M13 bacteriophage tethering and stretching. *Proc Natl Acad Sci U S A* 104, 4892-4897 (2007).
49. Neuman, K.C. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Methods* 5, 491-505 (2008).
50. Khalil, A.S. et al. Kinesin's cover-neck bundle folds forward to generate force. *Proc Natl Acad Sci U S A* 105, 19247-19252 (2008).
51. Lee, S.K. et al. Directed evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters. *Appl Environ Microbiol* 73, 5711-5715 (2007).
52. Gulati, S. et al. Opportunities for microfluidic technologies in synthetic biology. *J R Soc Interface* 6 Suppl 4, S493-506 (2009).
53. Simpson, M.L., Cox, C.D. & Sayler, G.S. Frequency domain analysis of noise in autoregulated gene circuits. *Proc Natl Acad Sci U S A* 100, 4551-4556 (2003).
54. Mettetal, J.T., Muzzey, D., Gomez-Urbe, C. & van Oudenaarden, A. The frequency dependence of osmo-adaptation in *Saccharomyces cerevisiae*. *Science* 319, 482-484 (2008).
55. Bennett, M.R. et al. Metabolic gene regulation in a dynamically changing environment. *Nature* 454, 1119-1122 (2008).
56. Malmstrom, J. et al. Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. *Nature* 460, 762-765 (2009).
57. Summerer, D. et al. A genetically encoded fluorescent amino acid. *Proc Natl Acad Sci U S A* 103, 9785-9789 (2006).
58. Wang, J., Xie, J. & Schultz, P.G. A genetically encoded fluorescent amino acid. *J Am Chem Soc* 128, 8738-8739 (2006).
59. Michalet, X. et al. Quantum dots for live cells, *in vivo* imaging, and diagnostics. *Science* 307, 538-544 (2005).
60. Hamad-Schifferli, K., Schwartz, J.J., Santos, A.T., Zhang, S. & Jacobson, J.M. Remote electronic control of DNA hybridization through inductive coupling to an attached metal nanocrystal antenna. *Nature* 415, 152-155 (2002).
61. Boyden, E.S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat Neurosci* 8, 1263-1268 (2005).
62. Levskaya, A. et al. Synthetic biology: engineering *Escherichia coli* to see light. *Nature* 438, 441-442 (2005).
63. Levskaya, A., Weiner, O.D., Lim, W.A. & Voigt, C.A. Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* (2009).
64. Misra, N. et al. Bioelectronic silicon nanowire devices using functional membrane proteins. *Proc Natl Acad Sci U S A* 106, 13780-13784 (2009).
65. Austin, D.W. et al. Gene network shaping of inherent noise spectra. *Nature* 439, 608-611 (2006).
66. Mar, D.J., Chow, C.C., Gerstner, W., Adams, R.W. & Collins, J.J. Noise shaping in populations of coupled model neurons. *Proc Natl Acad Sci U S A* 96, 10450-10455 (1999).

67. McGinness, K.E., Baker, T.A. & Sauer, R.T. Engineering controllable protein degradation. *Mol Cell* 22, 701-707 (2006).
68. Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.G. & Wandless, T.J. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell* 126, 995-1004 (2006).
69. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861-872 (2007).
70. Macarthur, B.D., Ma'ayan, A. & Lemischka, I.R. Systems biology of stem cell fate and cellular reprogramming. *Nat Rev Mol Cell Biol* (2009).
71. Lu, T., Ferry, M., Weiss, R. & Hasty, J. A molecular noise generator. *Phys Biol* 5, 36006 (2008).
72. Blake, W.J. et al. Phenotypic consequences of promoter-mediated transcriptional noise. *Mol Cell* 24, 853-865 (2006).
73. Discher, D.E., Mooney, D.J. & Zandstra, P.W. Growth factors, matrices, and forces combine and control stem cells. *Science* 324, 1673-1677 (2009).
74. Tagkopoulos, I., Liu, Y.C. & Tavazoie, S. Predictive behavior within microbial genetic networks. *Science* 320, 1313-1317 (2008).
75. Mitchell, A. et al. Adaptive prediction of environmental changes by microorganisms. *Nature* 460, 220-224 (2009).
76. Fernando, C.T. et al. Molecular circuits for associative learning in single-celled organisms. *J R Soc Interface* 6, 463-469 (2009).
77. Lee, D.K., Itti, L., Koch, C. & Braun, J. Attention activates winner-take-all competition among visual filters. *Nat Neurosci* 2, 375-381 (1999).
78. Brakmann, S. & Grzeszik, S. An error-prone T7 RNA polymerase mutant generated by directed evolution. *Chembiochem* 2, 212-219 (2001).
79. Yeh, B.J., Rutigliano, R.J., Deb, A., Bar-Sagi, D. & Lim, W.A. Rewiring cellular morphology pathways with synthetic guanine nucleotide exchange factors. *Nature* 447, 596-600 (2007).
80. Dueber, J.E., Mirsky, E.A. & Lim, W.A. Engineering synthetic signaling proteins with ultrasensitive input/output control. *Nat Biotechnol* 25, 660-662 (2007).
81. Bashor, C.J., Helman, N.C., Yan, S. & Lim, W.A. Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science* 319, 1539-1543 (2008).
82. Li, L. & Lindquist, S. Creating a protein-based element of inheritance. *Science* 287, 661-664 (2000).
83. Alberti, S., Halfmann, R., King, O., Kapila, A. & Lindquist, S. A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* 137, 146-158 (2009).
84. Bhalla, U.S., Ram, P.T. & Iyengar, R. MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network. *Science* 297, 1018-1023 (2002).
85. Stavreva, D.A. et al. Ultradian hormone stimulation induces glucocorticoid receptor-mediated pulses of gene transcription. *Nat Cell Biol* 11, 1093-1102 (2009).
86. Tsai, T.Y. et al. Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science* 321, 126-129 (2008).
87. McMillen, D., Kopell, N., Hasty, J. & Collins, J.J. Synchronizing genetic relaxation oscillators by intercell signaling. *Proc Natl Acad Sci U S A* 99, 679-684 (2002).
88. Garcia-Ojalvo, J., Elowitz, M.B. & Strogatz, S.H. Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing. *Proc Natl Acad Sci U S A* 101, 10955-10960 (2004).
89. Sarpeshkar, R. & O'Halloran, M. Scalable hybrid computation with spikes. *Neural Comput* 14, 2003-2038 (2002).
90. Molin, S. et al. Suicidal genetic elements and their use in biological containment of bacteria. *Annu Rev Microbiol* 47, 139-166 (1993).
91. Looger, L.L., Dwyer, M.A., Smith, J.J. & Hellinga, H.W. Computational design of receptor and sensor proteins with novel functions. *Nature* 423, 185-190 (2003).
92. Win, M.N., Klein, J.S. & Smolke, C.D. Codeine-binding RNA aptamers and rapid determination of their binding constants using a direct coupling surface plasmon resonance assay. *Nucleic Acids Res* 34, 5670-5682 (2006).
93. Skerker, J.M. et al. Rewiring the specificity of two-component signal transduction systems. *Cell* 133, 1043-1054 (2008).

94. Xu, J. & Lavan, D.A. Designing artificial cells to harness the biological ion concentration gradient. *Nat Nanotechnol* 3, 666-670 (2008).
95. Jogler, C. & Schuler, D. Genomics, genetics, and cell biology of magnetosome formation. *Annu Rev Microbiol* 63, 501-521 (2009).
96. Booth, I.R., Edwards, M.D., Black, S., Schumann, U. & Miller, S. Mechanosensitive channels in bacteria: signs of closure? *Nat Rev Microbiol* 5, 431-440 (2007).
97. Falke, J.J., Bass, R.B., Butler, S.L., Chervitz, S.A. & Danielson, M.A. The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu Rev Cell Dev Biol* 13, 457-512 (1997).
98. Gao, Z., Tseng, C.H., Strober, B.E., Pei, Z. & Blaser, M.J. Substantial alterations of the cutaneous bacterial biota in psoriatic lesions. *PLoS One* 3, e2719 (2008).
99. Grice, E.A. et al. Topographical and temporal diversity of the human skin microbiome. *Science* 324, 1190-1192 (2009).
100. Steidler, L. et al. Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10. *Science* 289, 1352-1355 (2000).
101. Braat, H. et al. A phase I trial with transgenic bacteria expressing interleukin-10 in Crohn's disease. *Clin Gastroenterol Hepatol* 4, 754-759 (2006).
102. Wei, M.Q., Mengesha, A., Good, D. & Anne, J. Bacterial targeted tumour therapy-dawn of a new era. *Cancer Lett* 259, 16-27 (2008).
103. von Maltzahn, G. et al. Nanoparticle self-assembly gated by logical proteolytic triggers. *J Am Chem Soc* 129, 6064-6065 (2007).
104. Stephanopoulos, G. Challenges in engineering microbes for biofuels production. *Science* 315, 801-804 (2007).
105. Slotta, U. et al. Spider silk and amyloid fibrils: a structural comparison. *Macromol Biosci* 7, 183-188 (2007).
106. Rammensee, S., Slotta, U., Scheibel, T. & Bausch, A.R. Assembly mechanism of recombinant spider silk proteins. *Proc Natl Acad Sci U S A* 105, 6590-6595 (2008).
107. Teule, F. et al. A protocol for the production of recombinant spider silk-like proteins for artificial fiber spinning. *Nat Protoc* 4, 341-355 (2009).
108. Widmaier, D.M. et al. Engineering the *Salmonella* type III secretion system to export spider silk monomers. *Mol Syst Biol* 5, 309 (2009).
109. Choi, J.H. & Lee, S.Y. Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol* 64, 625-635 (2004).
110. Klein-Marcuschamer, D. & Stephanopoulos, G. Assessing the potential of mutational strategies to elicit new phenotypes in industrial strains. *Proc Natl Acad Sci U S A* 105, 2319-2324 (2008).

CLAIMS

We claim:

1. A biological circuit chemotactic converter, the biological circuit chemotactic converter comprising: at least one input module (IM)_n, at least one genetic toggle switch (TS)_n, at least one logic module (LM)_n, and at least two sensor modules (SM_A and SM_B)_{n+1};
 wherein the input module comprises an inducible promoter sequence (iP_A) operably linked to a repressor sequence encoding a repressor protein (R_A);
 wherein the genetic toggle switch comprises a first repressible promoter sequence (rP₁) that drives expression of a second repressor sequence (R₂), and a second repressible promoter sequence (rP₂) that drives expression of a first repressor sequence (R₁), (rP₁-R₂ and rP₂-R₁);
 wherein the first and second repressor sequences encode first and second repressors proteins, and wherein the first repressible promoter sequence is inhibited by the first repressor protein and the second repressible promoter sequence is inhibited by the second repressor protein;
 wherein the logic module comprises a repressible promoter sequence (rP_B) operably linked to a repressor sequence (R_B);
 wherein each sensor module comprises a repressible promoter sequence (rP_C and rP_D) operably linked to a nucleic acid sequence encoding a sensor molecule (sensor A and sensor B);
 and wherein $n \geq 1$.
2. The biological circuit chemotactic converter of claim 1, wherein the inducible promoter sequence of the input module is induced by a biological agent, a chemical agent, a metal ion, a toxin, or a pollutant.
3. The biological circuit chemotactic converter of claim 1 or 2, wherein the repressor protein encoded by the first repressor sequence of the toggle switch (R₁) and the repressor protein encoded by the input module (R_A) are the same repressor protein.
4. The biological circuit chemotactic converter of any of claims 1-3, wherein the repressor proteins encoded by the logic module and the input module are different repressor proteins.
5. The biological circuit chemotactic converter of any of claims 1-4, wherein the repressor proteins encoded by the logic module and the genetic toggle switch are different repressor proteins.
6. The biological circuit chemotactic converter of any of claims 1-5, wherein the second repressible promoter sequence of the genetic toggle switch (rP₂), the repressible promoter sequence of the logic module (rP_B), and the repressible promoter sequence of one sensor module (rP_D) are repressed by the same repressor protein.

7. The biological circuit chemotactic converter of any of claims 1-6, wherein the second repressible promoter sequence of the genetic toggle switch (rP_2), the repressible promoter sequence of the logic module (rP_B), and the repressible promoter sequence of one sensor module (rP_D) comprise the same repressible promoter sequence.
8. The biological circuit chemotactic converter of any of claims 1-7, wherein the repressible promoter sequence of one sensor module (rP_C) is repressed by the repressor encoded by the logic module (R_B).
9. The biological circuit chemotactic converter of any of claims 1-8, wherein the at least two sensor modules encode for different sensor molecules.
10. The biological circuit chemotactic converter of any of claims 1-9, wherein n is an integer value between and including 1 and 100.
11. The biological circuit chemotactic converter of any of claims 1-10, wherein n is an integer value between and including 1 and 10.
12. The biological circuit chemotactic converter of any of claims 1-11, wherein n is an integer value selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10.
13. The biological circuit chemotactic converter of any of claims 1-12, further comprising at least one ribosome binding sequence.
14. The biological circuit chemotactic converter of any of claims 1-13, further comprising at least one terminator sequence.
15. The biological circuit chemotactic converter of any of claims 1-14, further comprising at least one degradation tag sequence.
16. The biological circuit chemotactic converter of any of claims 1-15, further comprising at least one sequence encoding an iRNA molecule specific for at least one protein encoded by the biological circuit chemotactic converter.
17. The biological circuit chemotactic converter of any of claims 1-16, wherein at least one repressor protein is an engineered zinc-finger protein.
18. The biological circuit chemotactic converter of any of claims 1-17, where the promoter sequence of the inducible and repressible promoters is selected from any of the promoter sequences of SEQ ID NOs: 1-7, SEQ ID NOs: 167-843, and SEQ ID NOs: 1009-1010.
19. The biological circuit chemotactic converter of any of claims 1-18, where the sensor molecule encoded by the nucleic acid sequence of the sensor domain is a chemotaxis receptor or chemotaxis sensor selected from Tables 19 and 20.

20. A system comprising at least one biological circuit chemotactic converter of any of claims 1-19.
21. A method of modulating chemotaxis in a biological system using the biological circuit chemotactic converter of any of claims 1-20.
22. The method of claim 21, wherein the biological system is a cell.
23. The method of claim 22, wherein the biological system is an artificial cell, a prokaryotic cell, or a eukaryotic cell.

1/15

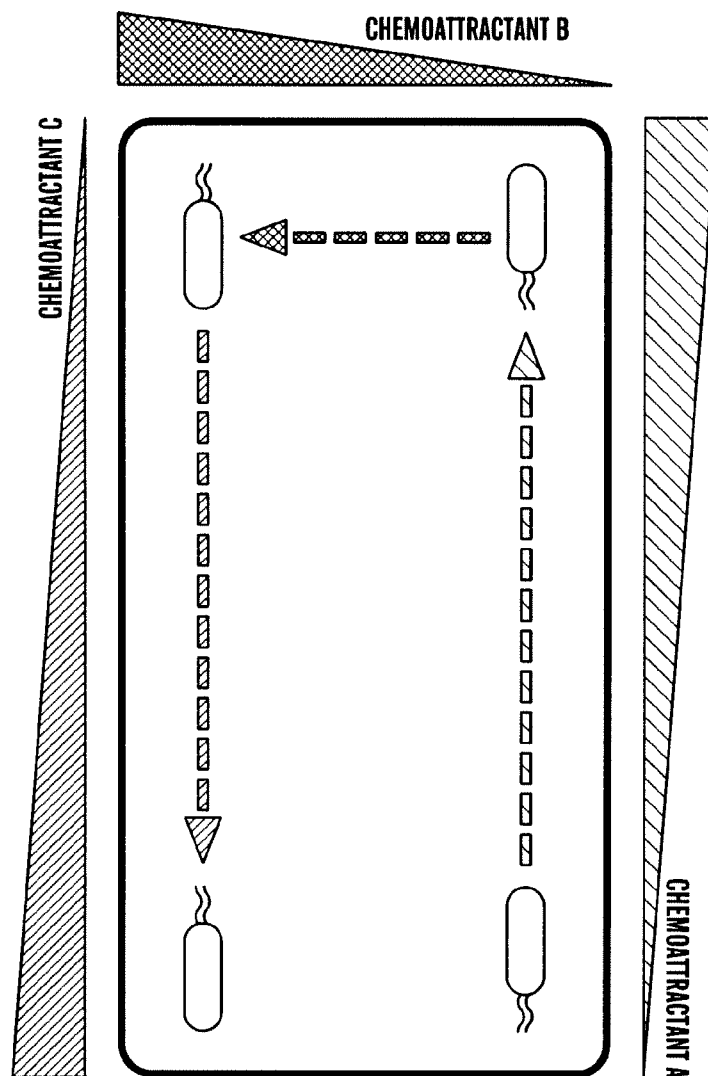
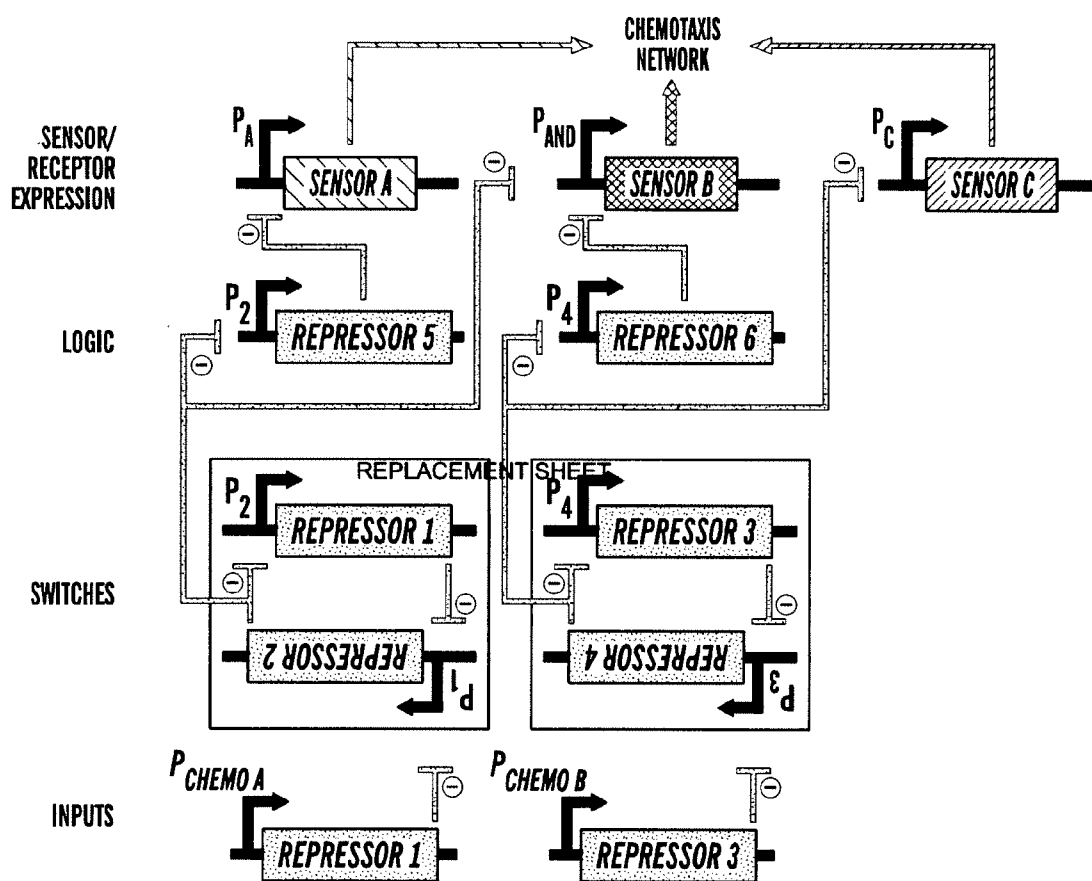


FIG. 1A

2/15

**FIG. 1B**

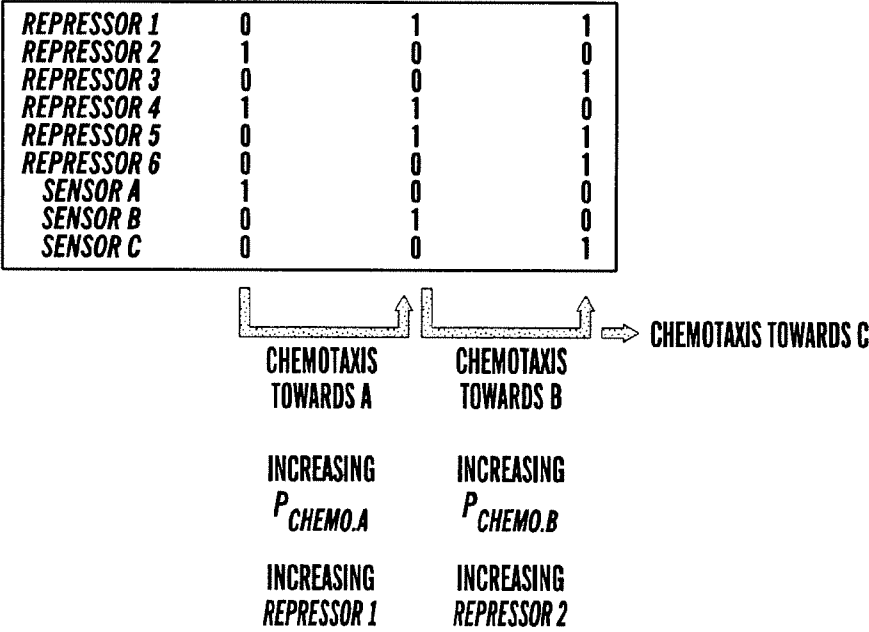


FIG. 1C

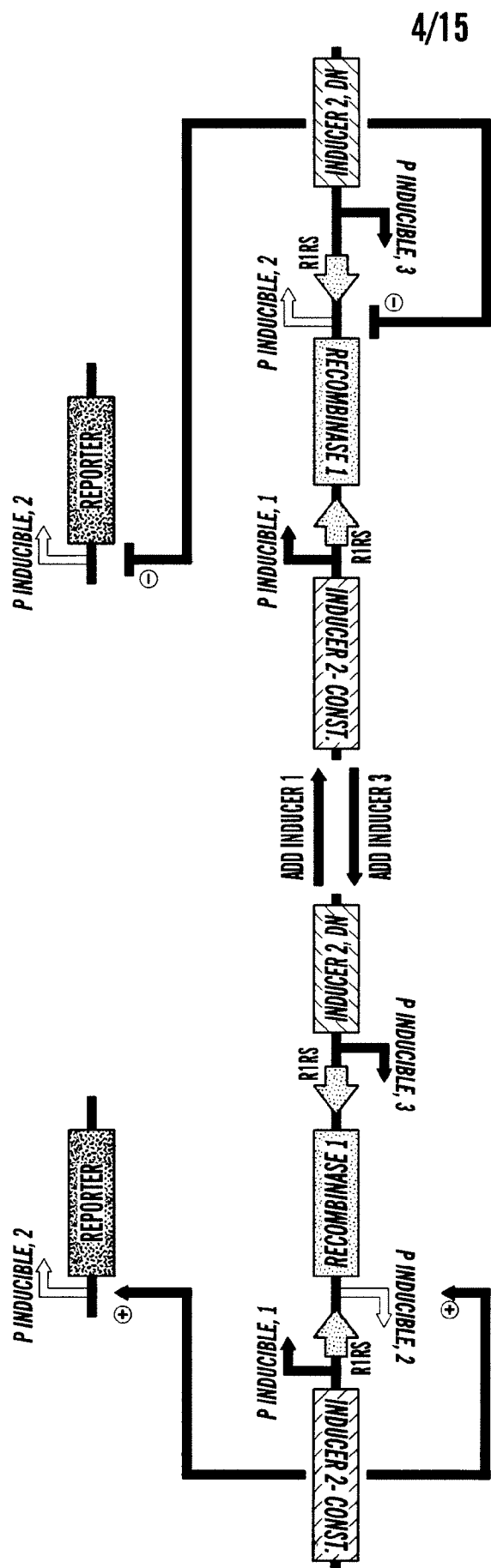


FIG. 2A

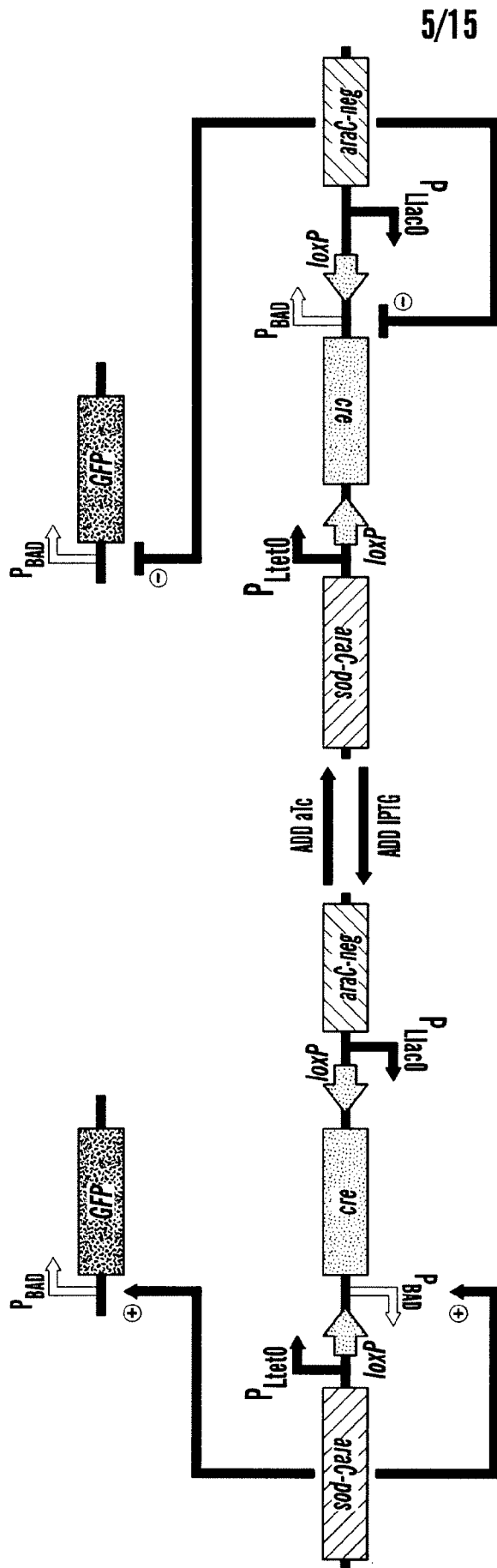


FIG. 2B

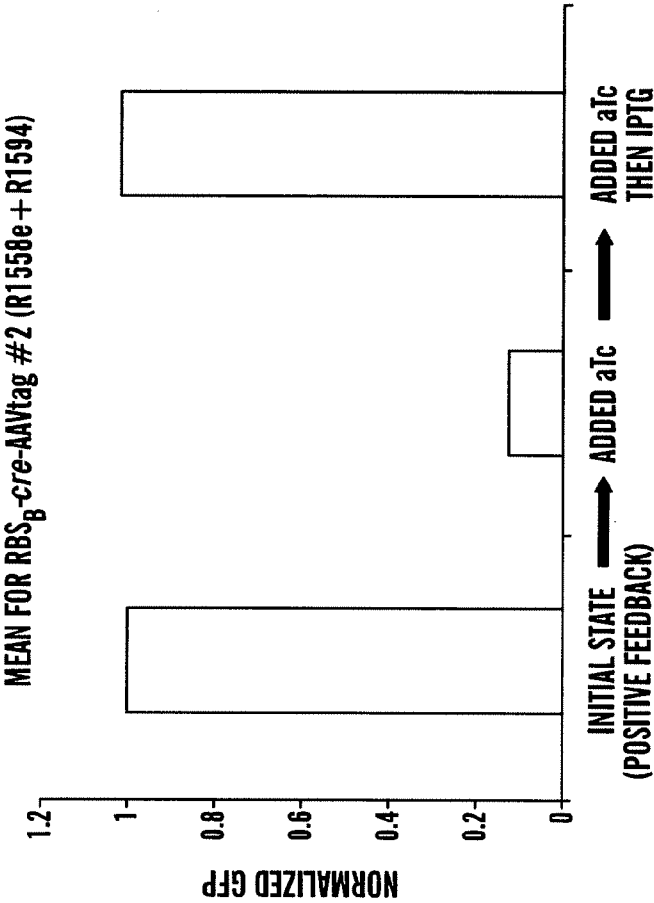


FIG. 2C

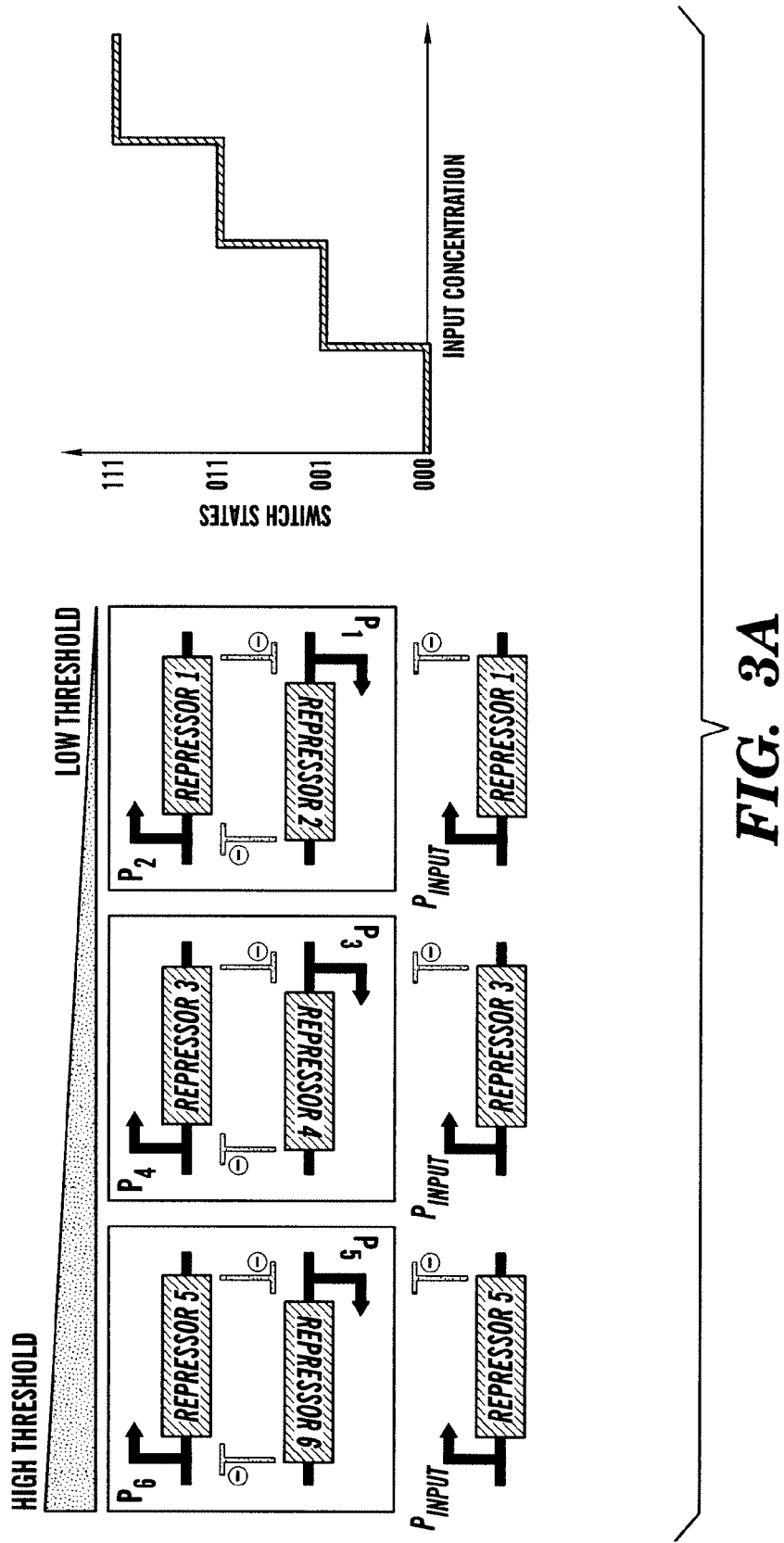


FIG. 3A

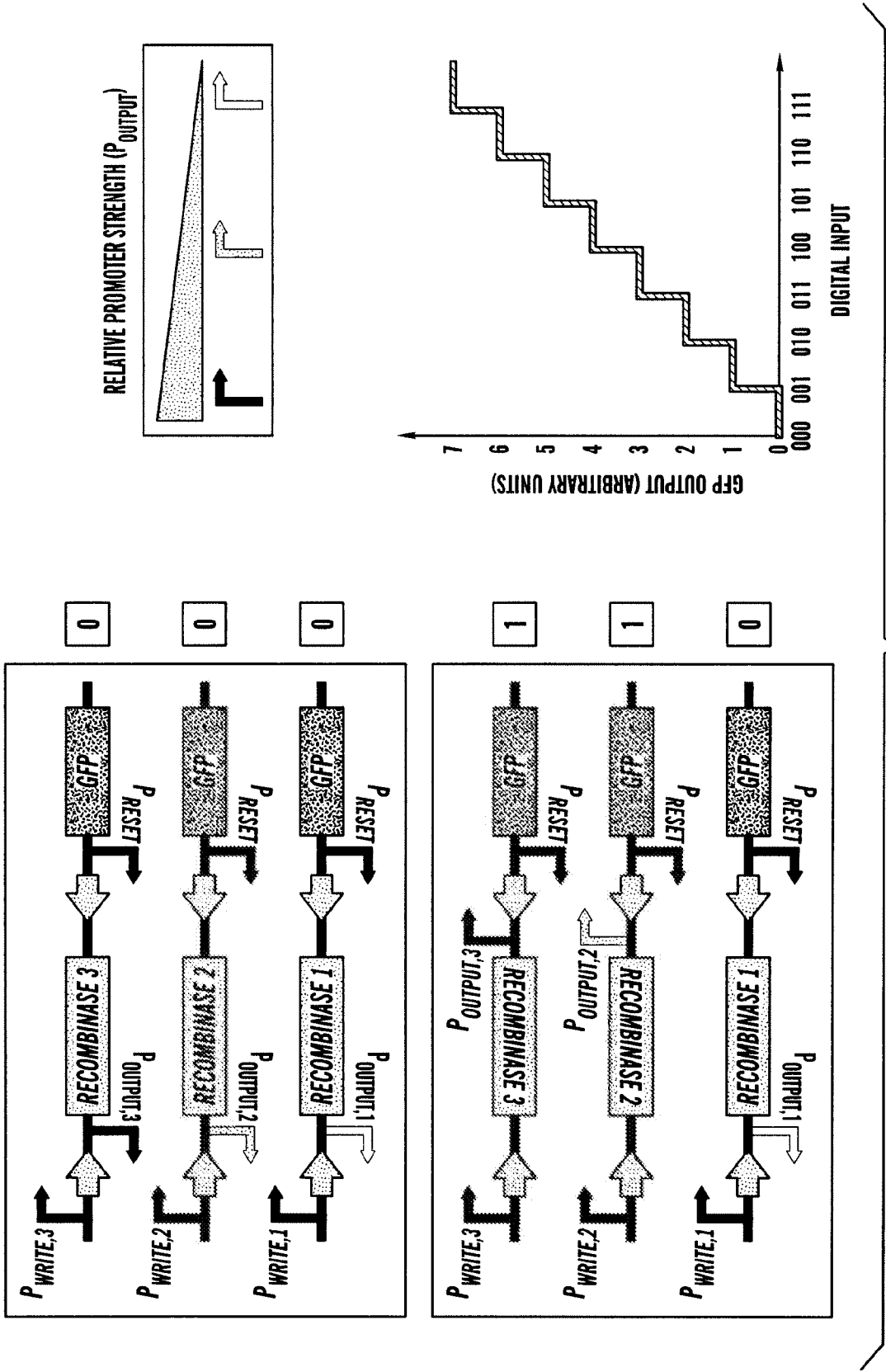


FIG. 3B

9/15

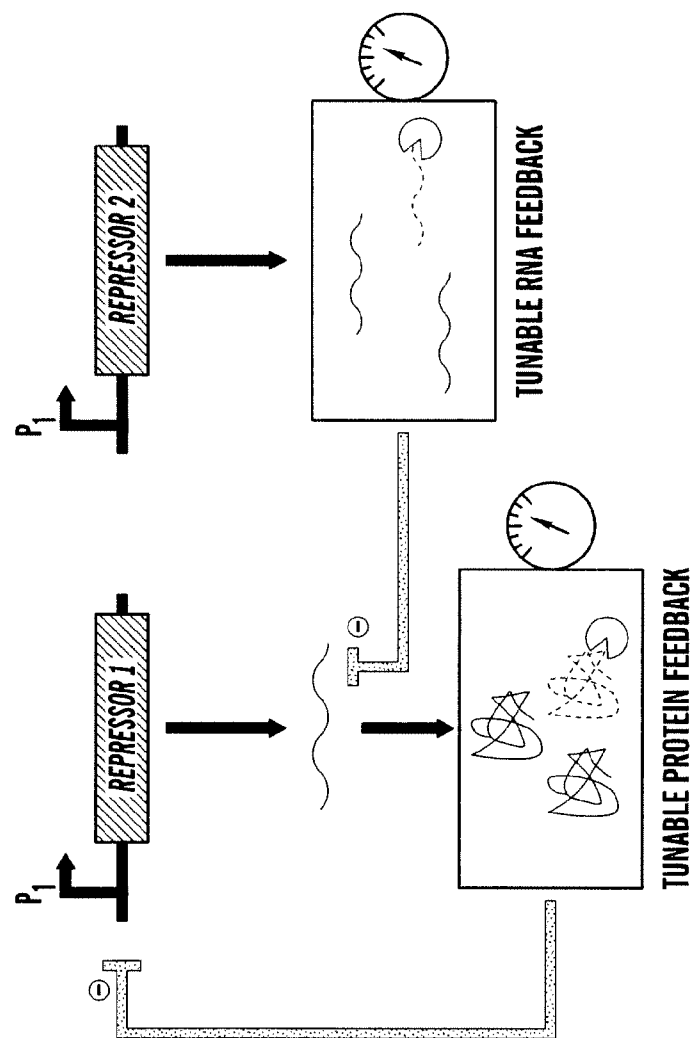


FIG. 4

10/15

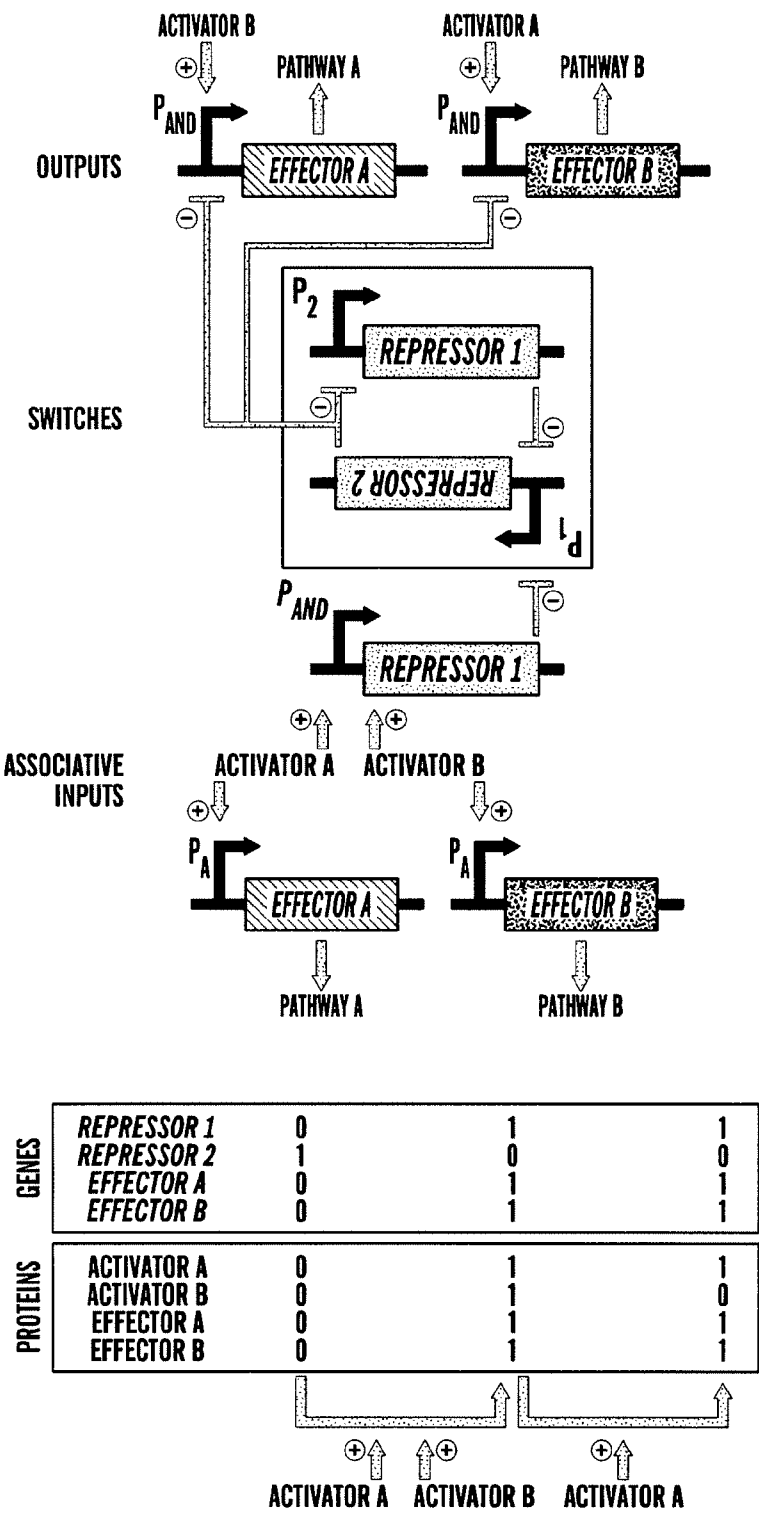
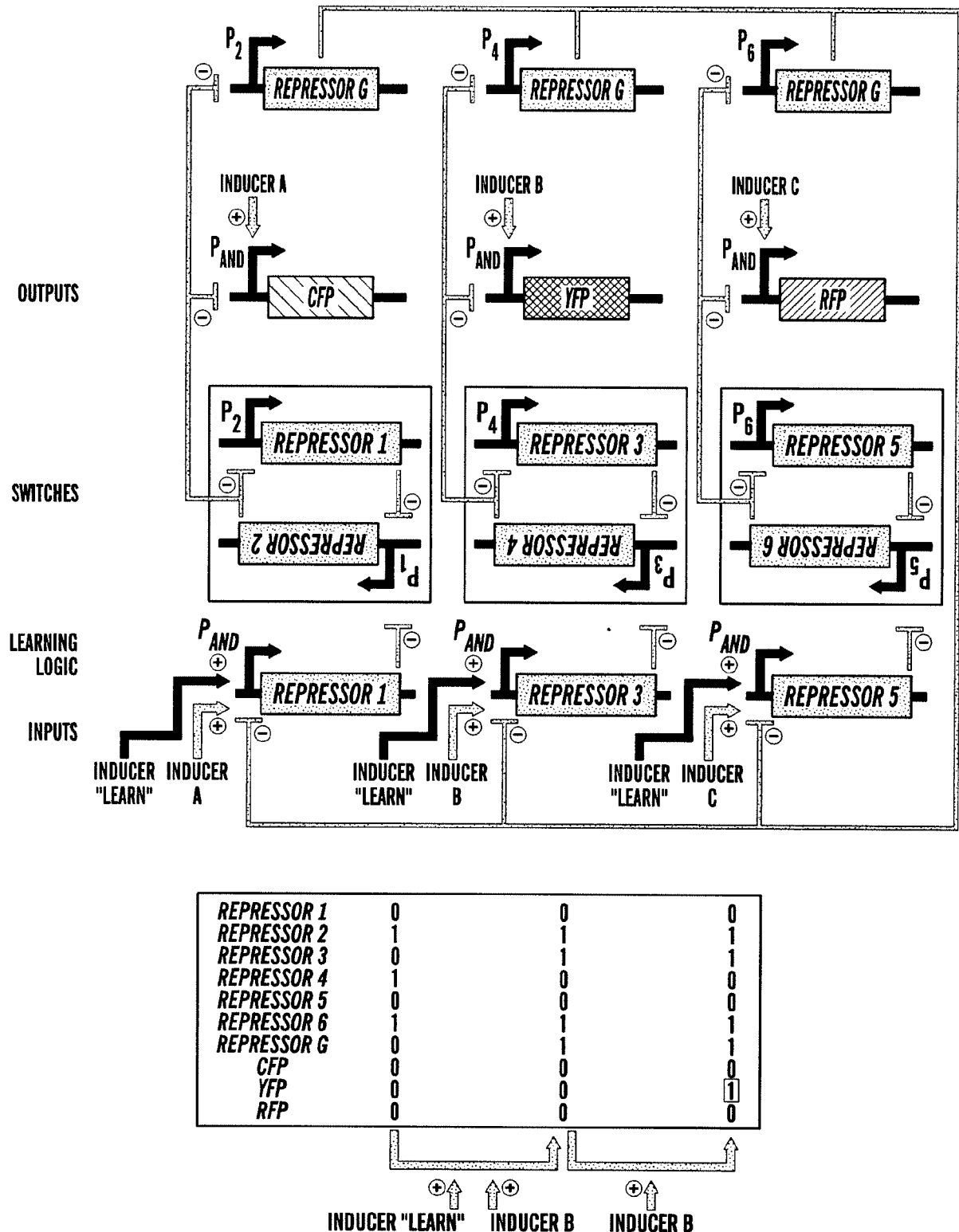
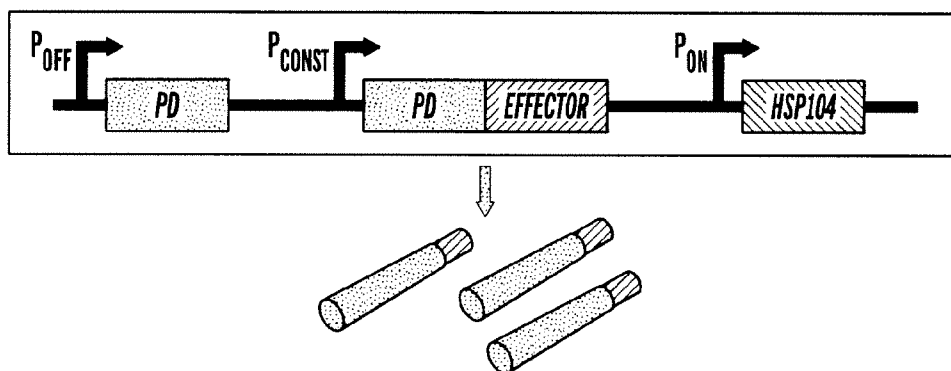
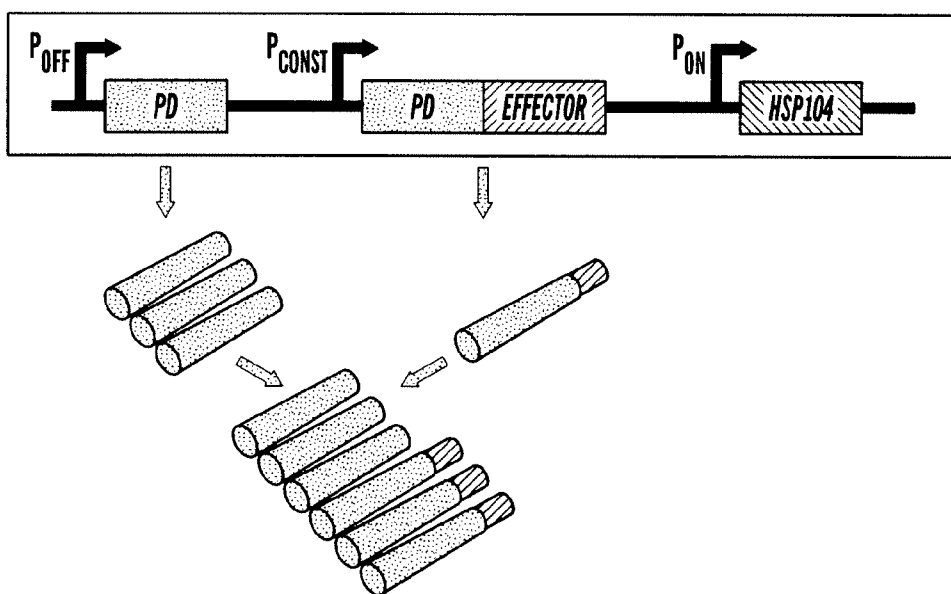


FIG. 5A

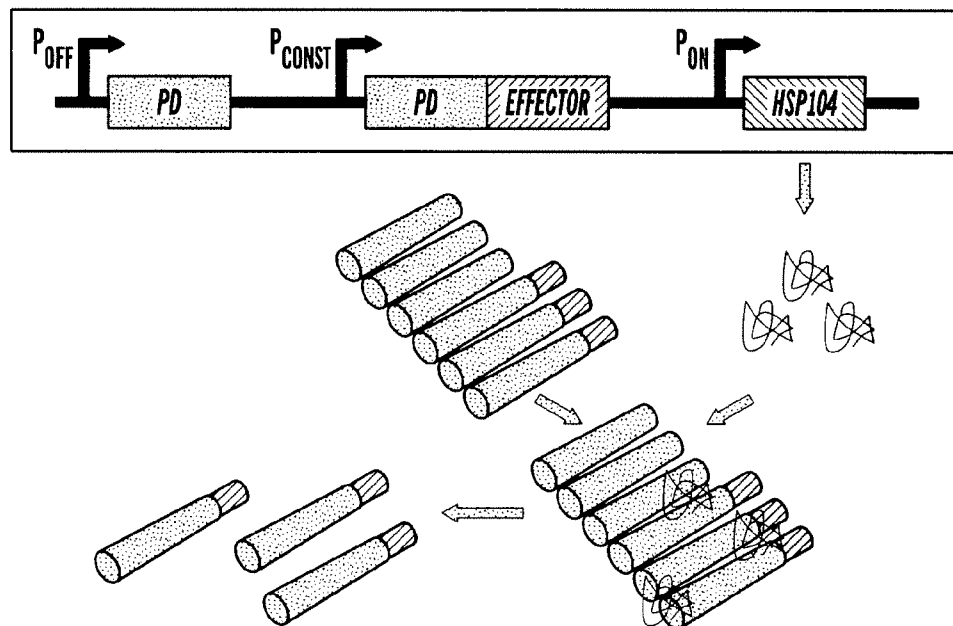
11/15

**FIG. 5B**

12/15

**FIG. 6A****FIG. 6B**

13/15

**FIG. 6C**

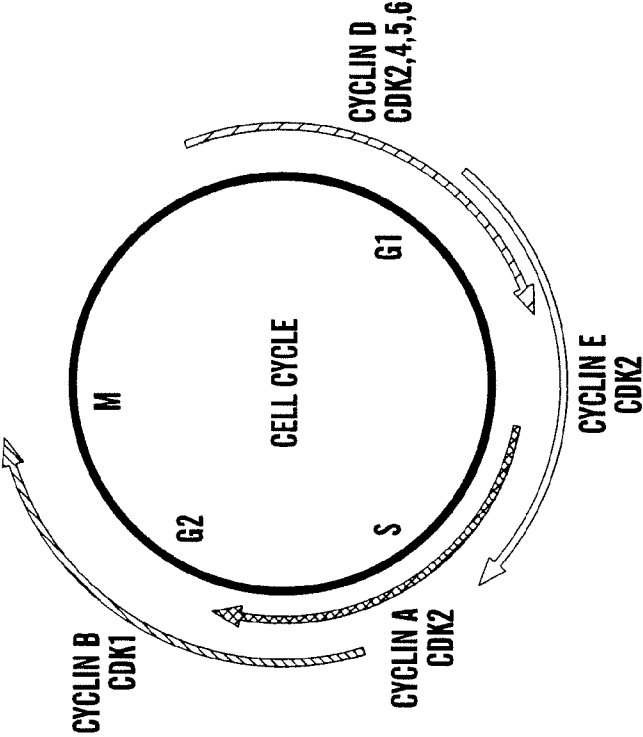
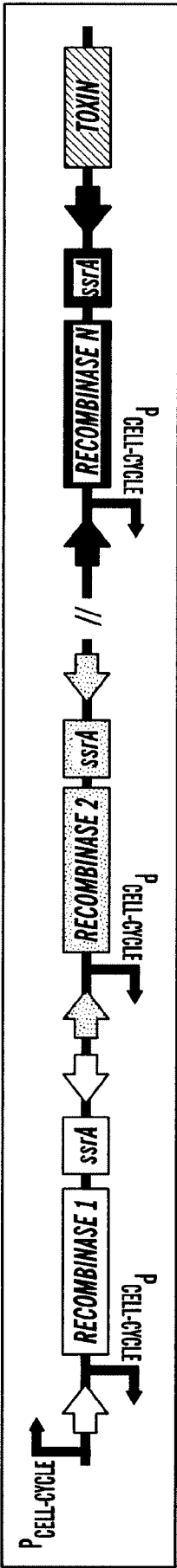


FIG. 7

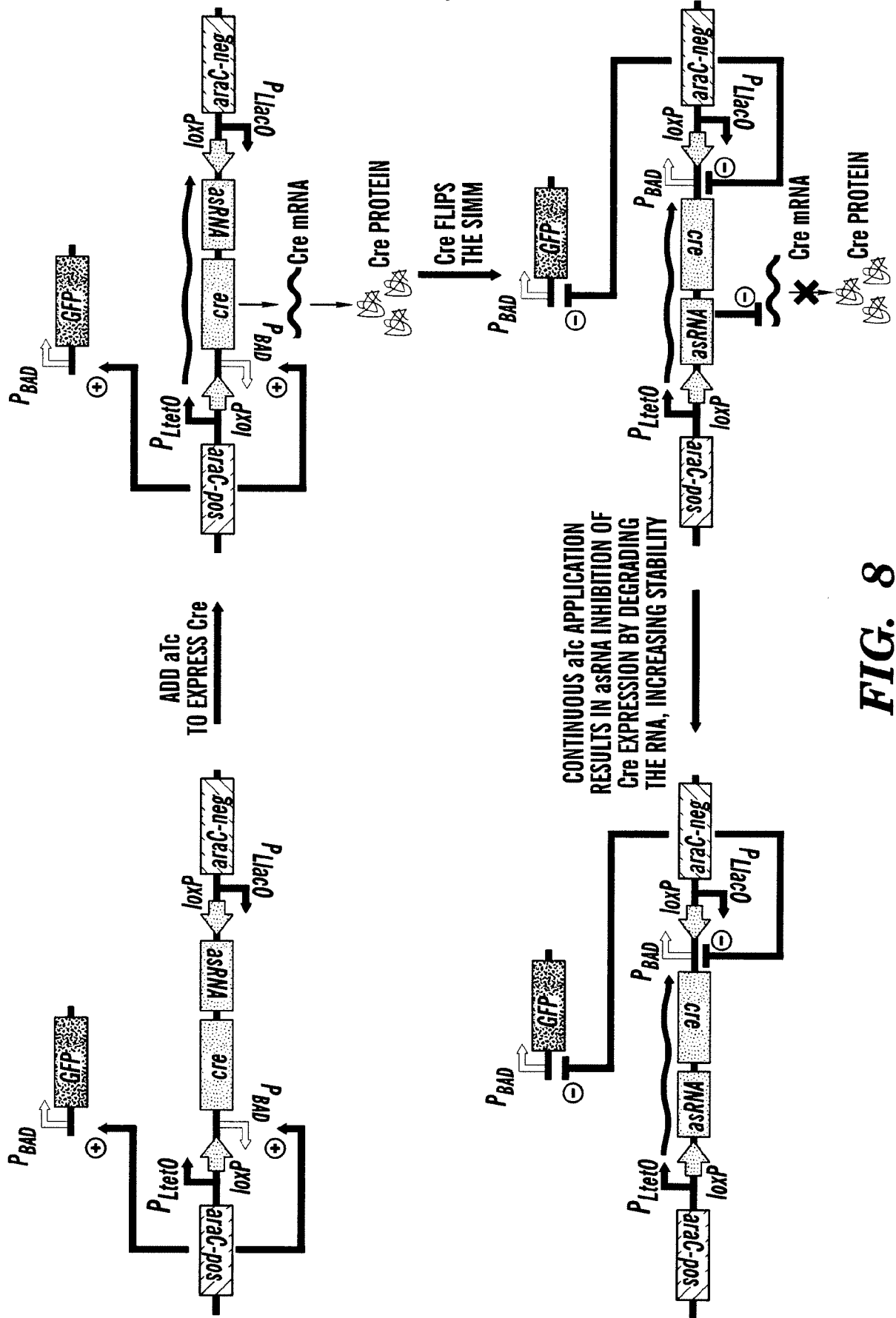


FIG. 8