Improved biosensor-based detection system

Rugbjerg, Peter; Genee, Hans Jasper; Sommer, Morten Otto Alexander

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 April 2015 (02.04.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/044456 A1

(51) International Patent Classification:
C12Q 1/25 (2006.01) *C12N 15/11* (2006.01)
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/EP2014/070954

(22) International Filing Date:
30 September 2014 (30.09.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
13186673.3 30 September 2013 (30.09.2013) EP

(71) Applicant: **DANMARKS TEKNISKE UNIVERSITET** [DK/DK]; Anker Engelunds Vej 101, DK-2800 Kgs. Lyngby (DK).

(72) Inventors: **RUGBJERG, Peter**; Borups Allé 2, 2. th., DK-2200 Copenhagen N (DK). **GENEE, Hans Jasper**; Baldersgade 63, 4. tv., DK-2200 Copenhagen N (DK). **SOMMER, Morten**; Parsbergsvej 40, DK-2830 Virum (DK).

(74) Agents: **LÖVQVIST, Anna** et al.; Inspicos A/S, P.O. Box 45, Kogle Allé 2, DK-2970 Hørsholm (DK).

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

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

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2015/044456 A1

(54) **Title**: IMPROVED BIOSENSOR-BASED DETECTION SYSTEM

(57) **Abstract**: Described is a new biosensor-based detection system for effector compounds, useful for in vivo applications in *e.g.* screening and selecting of cells which produce a small molecule effector compound or which take up a small molecule effector compound from its environment. The detection system comprises a protein or RNA-based biosensor for the effector compound which indirectly regulates the expression of a reporter gene via two hybrid proteins, providing for fewer false signals or less 'noise', tuning of sensitivity or other advantages over conventional systems where the biosensor directly regulates the reporter gene.

IMPROVED BIOSENSOR-BASED DETECTION SYSTEM

FIELD OF THE INVENTION

The present invention relates to methods of improving detection systems for effector compounds and to detection systems so obtained, as well as to related DNA molecules and host cells and to methods for using such a detection system in host cell screening or selection procedures.

BACKGROUND OF THE INVENTION

Biosensing is emerging as a powerful concept, allowing high-throughput *in vivo* screening for enzymatic or other intracellular production of specific compounds of interest (Michener *et al.*, 2012). A variety of small molecule-sensing elements exist and have been adapted for this and other regulation purposes, particularly including those that directly regulate gene expression, for example, RNA-based and transcription factor (TF)-based biosensors. Several different TF-based biosensors exist, typically based on the ability to bind to DNA in response to the binding of a ligand, for which it is a sensor. The binding to DNA then confers either activation or repression of transcription (Galvão and de Lorenzo, 2006).

RNA-based switches are biosensors of major interest for *in vivo* detection of molecules. These regulatory elements, also known as riboswitches, are able to modulate transcription or translation of a gene in response to ligand-binding (Serganov and Nudler, 2013). Riboswitches are found in nature, but the ligand-recognizing part of a riboswitch - the aptamer - can be tailored for a specific effector compound using SELEX (Stoltenburg *et al.*, 2007), and further advanced into a gene-regulating riboswitch by several different approaches (Muranaka *et al.*, 2009; Wittman and Suess, 2012).

In their normal environment, natural riboswitch biosensors are typically regulated metabolically, where modest fold changes of expression may be sufficient to adjust metabolic fluxes and the output signal from a sensor may be graded rather than binary. For applications in synthetic circuitry and screening of *e.g.* cellular expression libraries, however, this is less than ideal. The biosensor and the molecule to be 'sensed' are often present in few and/or variable numbers in the cells and the encounter between the biosensor and the molecule are constantly affected by varying dynamics such as cell cycle stage, stress level and age. All in all, this increases false, unintended signals or "noise", thereby decreasing the reliability of the detection system. For RNA-based biosensors, the effect may be particularly prominent because of their short half-life compared to protein-based biosensors.
Various strategies to improve the signal-to-noise ratio have been proposed. For example, it was possible to standardize for gene expression noise from a riboswitch regulating the expression of a fluorescence reporter gene by constitutively expressing a second fluorescent protein (Michener and Smolke, 2012). Another strategy employed to obtain signal stability

was to engineer feedback loops so that a signal-correlated TF-based biosensor was able to regulate its own expression as well as the output (Becskei and Serrano, 2000).

Other modified RNA-based detection systems have also been described. For example, WO2004/113495 describes the development of a ligand-dependent RNA-based transcriptional regulator reportedly mimicking the activation domain of a protein transcriptional regulator, and US20100197006 relates to exogenous molecular circuits in eukaryotic cells based on combinations of a plurality of mediators and two or more molecular switches, proposing noise-reduction strategies for, *e.g.*, circuit components receiving inputs from small RNAs. Further RNA and riboswitch-based detection systems are also described or discussed in WO 2011/088076, WO 2004/027035, WO 2007/100412, Dambach *et al.*, 2009; and Park et al., 2013. However, as concluded by Michener and Smolke, 2012, the future applications of RNA-switches in, e.g., screening technologies may be more effective if they incorporate biosensors with a higher signal to noise ratio. Also, to modify the sensitivity of a particular RNA-switch or TF, engineering of the actual RNA or TF sequence may be the only current option.

The so-called yeast-two-hybrid system was developed to sense protein-protein interactions *in vivo*, utilizing the DNA-binding domain (DBD) and activation domain (AD) of a transcription factor, respectively fused to a 'bait' and 'prey' domain so that a functional TF is only formed when the 'bait' and 'prey' domains interact (Fields and Song, 1989). The detected interaction signal is thus converted into a change in gene expression. The yeast-two-hybrid system has been further developed into more specialized formats for detection of other interaction types, such as, *e.g.*, the yeast three-hybrid (Baker *et al.*, 2002; Martin et al., 2012), reverse two-hybrid (Vidal *et al.*, 1996; WO9632503), controlled-release hybrid-systems (WO 2011/119956) and other variants (*e.g.*, WO0170816; Xu *et al.*, 1997; Stynen *et al.*, 2012, WO 2006/026712). In these systems, the use of split transactivating elements provides the regulation of a reporter gene, unlike TF-based and riboswitch-based biosensors, which in themselves are capable of regulating gene expression. In *Drosophila,* the so-called GAL4/UAS system was developed for detection of tissue-specific or developmental stage-specific gene expression (Potter *et al.*, 2010), assaying for specific promoter activation of GAL4 (or other transactivators) in certain tissues or developmental stages by operably linking the promoter in question to a heterologous *GAL4* leading to conditional reporter gene activation. However, the generation of protein-based biosensors *de novo* can be challenging, which in turn limits the number of reactions that can be detected or screened with this approach.

Thus, despite these and other advances in the art, there is still a need for biosensor detection systems, particularly those based on riboswitches or TF-based biosensors, where the biosensor 'switch' signals can be improved, *e.g.*, with respect to noise reduction, sensitivity level and/or inversion of signal.

SUMMARY OF THE INVENTION

It has been found by the present inventors that the addition of a protein-based "buffer device" can modulate the signal from a riboswitch or other biosensor which is in itself capable of regulating gene expression so as to reduce the noise and/or tune the detection sensitivity of the biosensor. Biosensor 'switch' signals can then be better separated from noise so that the system senses and responds to an input target molecule, i.e., an effector molecule, with improved discrimination in the assay and/or at a modified effector molecule concentration of interest. The invention thereby provides for novel detection systems suitable for use in high-throughput cell screening assays and other applications.

So, in a first aspect the present invention relates to an effector compound detection system comprising

(a) a first DNA molecule comprising a sequence encoding a first hybrid protein comprising a first interaction domain and an activation domain (AD), wherein the expression of the first DNA molecule in the host cell is under the control of a biosensor regulated by an effector compound of interest, optionally wherein the biosensor is a riboswitch or a transcription factor; and

(b) a second DNA molecule comprising a sequence encoding a second hybrid protein comprising a second interaction domain which binds to the first interaction domain and a DNA-binding domain (DBD) which binds a response element,

wherein the simultaneous binding of the second hybrid protein to the first hybrid protein and the response element triggers transcription of a reporter gene operably linked to the response element.

In one embodiment, the biosensor is a riboswitch. In this embodiment, the first DNA molecule can comprise a sequence encoding an RNA molecule comprising a riboswitch responsive to an effector compound, the riboswitch being operably linked to a coding region encoding a first hybrid protein comprising an activation domain (AD) and a first interaction domain, wherein binding of the effector compound to the riboswitch regulates expression of the first hybrid protein. In one embodiment, the riboswitch is an OFF riboswitch, so that the binding of the effector compound to the riboswitch decreases expression of the first hybrid protein. In one embodiment, each of the first and second DNA molecules is operably linked to a first and second promoter, respectively. In a particular embodiment, the strength of the first promoter is the same as or higher than that of the second promoter. In one embodiment, the effector compound detection system comprises a third DNA molecule encoding a third protein, wherein the third protein comprises the second interaction domain but does not bind the response element. This DNA molecule may also be operably linked to a promoter. In one additional or alternative embodiment, the effector compound detection system comprises a fourth DNA molecule comprising the reporter gene operably linked to a promoter region comprising the response element.

In a second aspect, the invention relates to a kit comprising one or more plasmids comprising the effector compound detection system of any preceding aspect or embodiment, wherein the kit optionally further comprises instructions for use.

In a third aspect, the present invention relates to a host cell comprising the effector compound detection system of any preceding aspect or embodiment. In one embodiment, the host cell further comprises a reporter gene operably linked to the response element, optionally integrated into a chromosome. In one embodiment, the host cell is a microbial cell or a mammalian cell.

The host cell can optionally be characterized by one or more of the following specific embodiments: (a) the host cell does not comprise an endogenous protein which binds to the first interaction domain, the second interaction domain or a binding site for the DBD in the response element; (b) in the absence of effector compound, the ratio of the expression level of the second interaction domain over the first interaction domain is in the range of about 0.03 to about 40 when the riboswitch is one where binding of the effector compound to the riboswitch decreases expression of the first hybrid protein; (c) the response element comprises at least 1, preferably at least 2 binding sites for the DBD; and/or (d) the ratio of the number of molecules comprising the second interaction domain to the number of binding sites for the DBD in the response element is at least 2. Host cells where (a) and (b); (a) and (c); (a) and (d); (b) and (c); (b) and (d); (c) and (d); (a), (b) and (c); (a), (c) and (d); (b), (c) and (d); or all of (a) through (d) are combined are also envisioned. In one embodiment, the change in expression of the reporter gene when the effector compound binds the riboswitch permits proliferation of the host cell under a predetermined condition under which the host cell would otherwise not proliferate, or proliferate at at least a 100-fold, such as at least a 500-fold, such as at least a 1000-fold lower frequency. In another embodiment, the change in expression of the reporter gene when the effector compound binds the riboswitch prevents proliferation of the host cell under a predetermined condition under which the host cell would otherwise proliferate, or proliferate at at least a 100-fold, such as at least a 500-fold, such as at least a 1000-fold higher frequency. Exemplary reporter genes of the preceding aspects and embodiments also include, but are not limited to those encoding, fluorescent, chromogenic or luminescent protein.

In a fourth aspect, the invention relates to an expression library of host cells according to the preceding aspect or any of its embodiments. The expression library typically comprises a plurality of expression constructs, each host cell comprising a vector encoding at least one potential modulator compound. Exemplary, non-limiting modulator compounds include (a) one or more enzymes of a biosynthetic pathway effecting a chemical transformation of a substrate into the effector compound in the host cell; or (b) one or more proteins transporting the effector compound into or out of the host cell.

In a fifth aspect, the invention relates to a method for identifying a host cell producing an effector compound of interest, comprising providing a plurality of host cells, each comprising

(i) a first DNA molecule comprising a sequence encoding a first hybrid protein comprising a first interaction domain and an activation domain (AD), wherein the expression of the first DNA molecule in the host cell is under the control of a biosensor regulated by an effector compound of interest, optionally wherein the biosensor is a riboswitch or a transcription factor; and

(ii) a second DNA molecule comprising a sequence encoding a second hybrid protein comprising a second interaction domain which binds to the first interaction domain and a DNA-binding domain (DBD) which binds a response element, and

(iii) a reporter gene operably linked to the response element, wherein the simultaneous binding of the second hybrid protein to the first hybrid protein and the response element triggers transcription of the reporter gene, and modulation of the expression of the reporter gene provides for a detectable signal; and

selecting at least one host cell where the expression of the reporter gene is modulated.

In a sixth aspect, the invention provides a method of producing an improved effector compound detection system from a first effector compound detection system, said first effector compound detection system comprising a DNA sequence encoding an RNA molecule comprising the transcript of a reporter gene operably linked to a riboswitch responsive to an effector compound, the method comprising the steps of:

(a) preparing a first DNA molecule comprising a sequence encoding an RNA molecule comprising the riboswitch operably linked to a coding region encoding a first hybrid protein comprising an AD and a first interaction domain, wherein binding of the effector compound to the riboswitch regulates expression of the first hybrid protein;

(b) preparing a second DNA molecule comprising a sequence encoding a second hybrid protein comprising (i) a second interaction domain which binds to the first interaction domain and (ii) a DBD which binds a response element;

thereby obtaining an improved effector compound detection system. In one embodiment, the binding of the effector compound to the riboswitch decreases expression of the first hybrid protein. In this embodiment, the first and second DNA molecules may comprise regulatory elements which, in the absence of the effector compound, provide for a higher expression level for the first than the second hybrid protein in a host cell, thereby obtaining a modified effector compound detection system having a reduced sensitivity for the effector compound. Alternatively, the first and second DNA molecules may comprise regulatory elements which, in the absence of the effector compound, provide for a lower expression level for the first than the second hybrid protein in a host cell, thereby obtaining a detection system having an increased sensitivity for the effector compound.

In a seventh aspect, the invention relates to an effector compound detection system comprising

(a) a first DNA molecule comprising a sequence encoding transcription repressor which binds a transcriptional activation domain (AD), wherein the expression of the first DNA molecule in the host cell is under the control of a biosensor regulated by an

effector compound of interest, optionally wherein the biosensor is a riboswitch or a transcription factor; and

(b) a second DNA molecule comprising a sequence encoding a first hybrid protein comprising a first interaction domain and the activation domain (AD), wherein the expression of the second DNA molecule in the host cell is optionally under the control of a constitutive promoter; and

(c) a third DNA molecule comprising a sequence encoding a second hybrid protein comprising a second interaction domain which binds to the first interaction domain and a DNA-binding domain (DBD) which binds a response element, and

(d) a reporter gene operably linked to the response element,

wherein the simultaneous binding of the second hybrid protein to the first hybrid protein and the response element triggers transcription of the reporter gene, and wherein the binding of the transcription repressor to the AD decreases the transcriptional activation of the reporter gene.

In one embodiment of this aspect, the AD and DBD are instead in the form of a single transcription factor (TF) protein. The second and third DNA molecules of this aspect can thereby be replaced by a single DNA molecule encoding a TF comprising an AD and a DBD. The detection system is then instead characterized by the binding of the TF to response element triggering transcription of the reporter gene, wherein the binding of the transcription repressor to the AD decreases the transcriptional activation of the reporter gene while surplus TF molecules act as noise-reducing buffer for the signal-correlated repressor.

In a specific embodiment of any of the aforementioned aspects, the biosensor is a riboswitch. In another specific embodiment of any of the aforementioned aspects, the biosensor is an OFF-riboswitch.

In other separate and specific aspects, the invention relates to kits or host cells comprising such effector compound detection systems, expression libraries of such host cells, and related detection or screening methods using such kits or host cells. These and other aspects and embodiments are discussed in further detail below.

LEGENDS TO THE FIGURES

Fig. 1 illustrates the general principles of noise-reduction in detecting an effector compound using exemplary buffer devices, exemplified by a system in which the biosensor is an 'OFF'-riboswitch and the reporter gene is *URA3*. (A) No effector compound ("molecule") is sensed by the detection system, resulting in cell death in the presence of the counter selection agent 5'-fluoroorotic acid (5'-FOA). (B) Effector compound ("molecule") is sensed by the system, leading to cell survival in the presence of 5'-FOA. A falsely switched signal, here an AD molecule expressed because of a 'leaky' riboswitch, is buffered by free DBD molecules.

Fig. 2 shows the design and components of exemplary plasmids comprising DNA molecules according to the invention. (A) Plasmid pEXP22.1. (B) Plasmid pEXP32.4. See Example 1 for details.

Fig. 3 A and B show the ability of different *S. cerevisiae* strains based on different combinations of riboswitch, AD/DBD systems and reporter genes to form colonies in presence or absence of the effector compound, here the riboswitch ligand tetracycline (250 µM), illustrating the noise-reduction effect. Important genetic construct features of each strain family is shown in order to relate the different combinations of riboswitch and reporter gene to the layout of activation domain hybrid protein (AD) and DNA-binding domain hybrid protein (DBD). Unless otherwise stated, a 5-FOA concentration of 0.1 % (w/v) was used. See Examples 2 and 6 for details.

Fig. 4 shows how the ability of the different *S. cerevisiae* strains to form colonies in the presence or absence of the riboswitch ligand can be shifted to higher concentrations (from 50 to 150 µM tetracycline) by changing the relative promoter strengths for the expression of the AD hybrid protein and the DBD hybrid protein. See Example 3 for details.

Figure 5 shows quantification of change of effector response curve by perturbation of the transcriptional ratio of AD to DBD hybrid molecules. The output was determined as background-subtracted relative fluorescence units (RFU) from the two strains PRa78 (AD1) and PRa79 (AD2) after cultivation in medium supplemented with the indicated concentration of biosensor ligand (tetracycline in this case). In the upper, small window the same measurements are represented with a logarithmic x-axis. See Example 5 for details.

DETAILED DISCLOSURE OF THE INVENTION

The detection system of the invention is a modification of conventional systems where the biosensor, e.g., the riboswitch or other RNA-based biosensor, directly regulates the transcription of a reporter gene. By indirectly regulating the expression of the reporter gene via two hybrid proteins, fewer false signals or less 'noise', tuning of sensitivity or other advantages can be achieved over the original biosensor. Specifically, to provide signal robustness and to control and stabilize the two response states (OFF or ON) of a biosensor in an intracellular environment, the effector compound detection system of the invention, herein also referred to as "noise-reduction system", provides a surplus pool of unreactive proteins for false signal-correlated proteins to bind to during the inactive state rather than relaying the signal. Exemplary noise includes false positive cells in which the reporter gene is or has been incorrectly active or inactive, leading *e.g.* to cells surviving a selection condition in absence of the condition selected for. As indicated above, noise can arise from random biological fluctuations in molecule numbers due to their stochastic formation, regulation and degradation.

As used herein, the term "buffer device" refers to the components of the effector compound detection system after their introduction into a host cell, where the first and

second DNA molecules are transcribed and, at least the RNA encoding the second DNA molecule, translated into the second hybrid protein to form the surplus "buffer" pool. The basic system layout is depicted in Figure 1: The presence of effector compound in the host cell regulates the expression of a hybrid activation domain using a biosensor of choice, here illustrated by a riboswitch. To reduce noise in this regulation, represented by falsely present hybrid activation domain (Fig. 1B), the signal is not transmitted directly to a reporter gene, but is instead dependent on interaction between the activation domain and a hybrid DNA-binding domain, which is constitutively expressed in the host cell. Only when both domains are bound and occupying a DNA-binding site of the reporter gene promoter, the signal is transmitted, leading to reporter gene activation. Noise is thus reduced, since the number of hybrid DNA-binding domain exceeds the number of DNA-binding sites in the reporter gene promoter, which reduces the probability that a falsely present hybrid activation domain leads to reporter gene activation.

As a model system, the inventors chose a synthetic tetracycline-responsive but noise-prone riboswitch (see Example 2, PRd3-5 and PRa24) with limited control of reporter genes, first and second hybrid proteins based on Gal4 AD and DBD with RalGDS and Krev1 interaction domains. As shown in the Examples, using a colony formation assay with the classical yeast reporter gene *URA3,* it was demonstrated that the buffer device considerably improved the response of the riboswitch to the presence of effector compound. By implementing the buffer device, the selection window (i.e., the difference in cell colony formation in the absence as compared to the presence of effector compound) was expanded from approximately 1- to more than 10^4 –fold at 50 µM of effector compound and from approximately 10- to more than 10^4 –fold at 500 µM of effector compound, which is compatible with library screening. It was further demonstrated that the buffer device could be used as instrument to genetically program the detection sensitivity and thus the selection threshold for the effector compound (see Example 3). Additional advantages of the buffer device are described below.

Definitions

"Effector compound" as used herein refers to any compound whose binding to a specific RNA- or protein-based biosensor, such as a riboswitch or transcription factor, effects a conformational change in the biosensor which, in turn, directly regulates or changes the expression of a reporter gene. For example, the binding of an effector compound to a riboswitch biosensor regulates the expression of the coding sequence it is operably linked to, whereas the binding of an effector compound to a transcription factor biosensor regulates the transcription of a gene operably linked to a promoter which is regulated by the binding of the transcription factor. The effector molecule can be a known (*i.e.*, cognate) ligand for the biosensor or another molecule not formerly known to bind the biosensor, optionally structurally similar to a known ligand for the biosensor. Typically, the effector compounds are

small molecules comprising several carbon-carbon bonds, whose production in or transport into a microbial cell can be catalyzed or effected by a modulator compound. Optionally, the effector compound has a molecular weight of 2 kD or less, such as 1.5 kD or less, such as 1 kD or less, such as between about 0.05 kD to about 2 kD, such as between 0.1 kD to 1.5 kD.

"Response element" is a DNA sequence within or upstream of a promoter region of a gene which comprises one or more binding sites for the DNA-binding region of a specific transcription factor (a.k.a. "DBD-binding sites"), thus regulating the transcription of the gene. Native (*i.e.*, naturally occurring) response elements can comprise as little as only one DNA-binding site. However, a synthetic response element can, for example, be designed to comprise one, two, three, four, five, six, seven, eight, nine, ten or more such binding sites, or as described further below.

"Reporter gene" as used herein refers to a gene whose transcription or expression can be directly or indirectly detected, optionally under a predetermined condition. For example, a reporter gene can encode, or regulate the transcription or expression of, a fluorescent, chromogenic or luminescent protein; an enzyme which converts an artificial substrate to detectable compound; a protein which is essential to host cell proliferation or survival in a particular medium; or a protein which prevents host cell proliferation or survival in a particular medium. Many reporter genes and their selection conditions are known and used in the art, see, *e.g.*, Table 2 of WO 2007/073163, hereby incorporated by reference, and the present Examples.

"Modulator compound" as used herein refers to a compound which modulates the action of an effector compound, wherein said primary modulator compound either effects a chemical transformation of a substrate into an effector compound or transports said effector compound into a microbial or eukaryotic cell. Enzymes or other biocatalysts are particularly contemplated as modulator compounds promoting (*i.e.*, catalyzing) the conversion of a substrate into an effector compound.

A "plasmid" as used herein refers to a nucleic acid molecule such as a DNA molecule that can be used to introduce a specific coding nucleic sequence into a host cell. A particular type of plasmid is an "expression vector", which, after its introduction into a host cell can express the coding nucleic acid sequence. Once the expression vector is inside the cell, the RNA molecule or protein that is encoded by the nucleic acid sequence can be produced by the endogenous cellular transcription and transcription/translation machinery of the cell, respectively. Generally, expression vectors include regulatory elements which are operably linked to the coding sequence or gene and control its transcription and/or translation.

"Transcription" refers to the synthesis of RNA on a DNA or RNA template. "Translation" refers to the synthesis of a polypeptide or protein on a messenger RNA template.

A nucleic acid sequence is "operably linked" when it is placed into a regulatory relationship with another nucleic acid sequence. For example, a riboswitch is operably linked to a coding sequence if it regulates the translation of said sequence, and a response element is operably linked to a coding sequence if it regulates transcription of said sequence.

Standard recombinant DNA and molecular cloning techniques useful for preparing and manipulating the nucleic acid sequences and host cells of the present invention are described, *e.g.*, in Sambrook, J., Fritsch, E. F. and Maniatis, T., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989); and by Silhavy, T. J., Bennan, M. L. and Enquist, L. W., Experiments with Gene Fusions, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1984); Ausubel *et al.*, Short Protocols in Molecular Biology, 3rd ed., Wiley & Sons, 1995; and by Ausubel, F. M. *et al.*, Current Protocols in Molecular Biology, published by Greene Publishing Assoc. and Wiley-Interscience (1987). Additional methods used here are in Methods in Enzymology, Volume 194, Guide to Yeast Genetics and Molecular and Cell Biology (Part A, 2004, Christine Guthrie and Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, Calif.).

Specific embodiments of the invention

The basic components of the buffer device are described in further detail below.

Biosensor

Both riboswitches, which are in themselves capable of regulating the expression of a (reporter) gene, and transcription factors that regulate gene expression in response to the availability of an effector compound are contemplated as biosensors and can be used in the effector compound detection systems of the invention.

Riboswitches are RNA-based expression control elements regulating the expression of an (mRNA) coding sequence, and which change conformational state when bound by an effector compound. Riboswitches are typically comprised of two domains: the aptamer domain that selectively binds the effector molecule, and the expression platform domain that influences genetic control. The dynamic interplay between these two domains controls expression of the coding sequence, typically via controlling ribosomal access and thereby translation of the coding sequence, though riboswitches controlling expression at the level of transcription or mRNA splicing are known. Riboswitches generally function in two distinct modes; OFF and ON. OFF riboswitches ('negative selection' as used herein) down-regulate gene expression upon binding an effector compound to the aptamer domain. Typically, the conformational change conceals the ribosome binding site, which disallows expression of the associated gene, or a rigid RNA-structure is formed, preventing read-through by the ribosome. ON riboswitches ('positive selection' as used herein) allow gene expression upon binding an effector compound to the aptamer domain. For example, the conformational change exposes a ribosome binding site, which allows expression of the associated gene. 'Switching' of a riboswitch herein refers to the change in the riboswitch upon binding of the effector compound.

In the context of the present invention, the role of a riboswitch biosensor is to regulate the expression, preferably the translation, of the RNA coding sequence for the first hybrid

protein. Accordingly, in the case of an OFF riboswitch, the first hybrid protein should be expressed in the absence of effector compound, but 'switched off' when the effector compound binds the aptamer-portion of the riboswitch. In the case of an ON riboswitch, the first hybrid protein should only be expressed when the effector compound binds to the riboswitch which is then 'switched on.'

Table 1 lists preferred riboswitches for use in the context of the invention, the most preferred being tc3. However, other suitable riboswitches can be selected from any source, including naturally occurring riboswitches, chimeric riboswitches, engineered riboswitches, and recombinant riboswitches. Exemplary riboswitches are described in US 2005/0053951, US6831171, WO 2006/055351, US 60/625,864, US 2009/0305253, EP2322535 A1, WO 2011/088076, WO 2004/113495, US 2012/0244601, US 2013/0004980, US 8313901, Wittman and Seuss (2012), Michener and Smolke (2012), Ogawa and Maeda (2008), Yang *et al.* (2013) and Desai and Gallivan (2004), each of which is hereby incorporated by reference in its entirety, including the description and nucleic acid sequence of each riboswitch described or referred to therein. Typically, the riboswitch is selected based on the desired effector compound being a cognate ligand for the aptamer domain of the riboswitch. Methods for screening and selecting riboswitches and/or compounds which activate riboswitches are known in the art, see, *e.g.*, US 2009/0305253, US 2013/0004980, EP 2322535 A1, and WO 2011/088076, each of which is hereby incorporated by reference in its entirety, specifically for any method of screening and selecting riboswitches.

Typical riboswitches suitable for improvement according to the invention include, e.g., those that have an undesirable (i.e., too high or too low) sensitivity for its ligand, or, upon regulation of a reporter gene, show poor discrimination between the states with and without the ligand (i.e., noise). Other riboswitches may have undesired regulation directionality, meaning that they e.g. are OFF switches, but the required regulatory response for the reporter gene is an ON signal (switching ON gene expression upon presence of ligand). The invention now provides a "buffer device" for such riboswitches by which the riboswitch now regulates a specified component of the detection system of the invention which then ultimately transfers the signal of the riboswitch to the reporter gene.

TABLE 1

Riboswitch (ON/OFF)	**Known cognate ligand(s) of the aptamer**	**Reference**
tc3 (OFF)	tetracycline, 7-chloro-tetracycline, oxytetracycline, doxycycline, anhydrotetracycline	Kötter *et al.*, 2009 Müller *et al.*, 2006
M4 (OFF)	Neomycin	Weigand *et al.*, 2008
ThiMN15#8 (ON)	Thiamine pyrophosphate	Muranaka *et al.*, 2009

ThiMN15#19 (ON)		
B and B_2 (OFF)	Biotin	Harvey *et al.*, 2002
106 folT (OFF)	Tetrahydrofolate	Ames *et al* 2010
btuB leader (OFF)	5'-deoxy-5'- adenosylcobalamin	Nahvi *et al.*, 2002

Transcription factors are also contemplated as biosensors for the buffer devices, regulation transcription of the first hybrid protein according to the same 'ON/OFF' principles as the riboswitch biosensors described above. Suitable transcription factors for application as biosensors in the present invention are those that, upon binding by an effector molecule such as its cognate small molecule ligand, change binding affinity for the cognate DNA-binding site so that the transcription activity of operably linked gene is switched. Most preferred are noise-prone transcription factors which can cause undesired fluctuations in the expression of the gene they regulate. Suitable transcription factors functioning as biosensors for small molecules have been described in the art, including the mutant NahR from *P. putida*, which has benzoic acids as cognate ligands (Sint Fiet *et al.* (2006), WO2007/073163). See, in particular, Table 1 of WO 2007/073163, describing suitable combinations of effector compounds, transcriptional regulators and promoter combinations, which Table is hereby incorporated by reference in its entirety.

First and second hybrid proteins

In the buffer device of the invention, the first hybrid protein functions as a 'signal' transmitted by the biosensor, reflecting a change or 'switch' in the conformational state of the biosensor caused by the binding of the effector molecule. The second hybrid molecule functions, in turn, both as a buffer and as mediator of the signal in the actuator platform.

The first hybrid protein comprises an activation domain (AD) and a first interaction domain. The second hybrid protein comprises a DNA-binding domain (DBD) and a second interaction domain can associate with, typically bind to, the first interaction domain. When first and second hybrid proteins interact or bind to each other via their interaction domains, they reconstitute a functional transcription factor capable of activating transcription of the reporter gene (see Figure 1). Suitably, AD and DBD components from any yeast-two-hybrid or yeast three-hybrid system known in the art which comprises AD and DBD hybrid proteins can be adapted for use in the present invention. Exemplary systems are described in, *e.g.*, Vidal *et al.* (1996), WO9632503, WO9813502, US 20090005253 A1, Xu *et al.* (1997), Baker *et al.* (2002), Stynen *et al.* (2012), each of which incorporated by reference in its entirety, include the encoded amino acid sequence of each AD and DBD domain described therein.

Generally, the AD is capable of interacting with endogenous transcription co-regulators or other factors regulating transcription factor activity in the host cell contemplated. For example, ADs suitable for yeast cells can be derived from *S. cerevisiae* proteins such as Gal4, Oaf1, Leu3, Rtg3, Pho4, Gln3, Gcn4 and VP16 (*Herpes simplex*), whereas ADs suitable for

mammalian cells include p53, NFAT, NF-κB and VP16. All of these comprise a nine-amino-acid transactivation domain motif common to the transactivation domains of a large number of yeast and mammalian transcription factors (Piskacek *et al.* (2007), Genomics 89, 756-68). On-line tools are available at, *e.g.*, ExPASy and EMBnet Spain for identifying additional ADs based on such motifs. For bacterial host cells, exemplary ADs include *E. coli* B42 and the RNA polymerase α subunit (Dove and Hochschild, 2004). If the host organism has endogenous gene(s) encoding proteins that specifically obstruct the activating function of the AD, these gene(s) must be deleted first. An example is *GAL80*, which was deleted from the genome in all strains constructed from MaV203 in the examples.

Apart from being able to form a functional transcription factor when associated with the AD via the interaction domains, the DBD is generally selected based on its ability to bind to a specific DNA-binding site(s) of a response element operatively linked to the reporter gene, but activating transcription of the reporter gene only when the DBD is simultaneously associated with the first hybrid protein. Suitable DBDs include, for example, those derived from Gal4, LexA, cI (Serebriiskii *et al.*, 1999), LexA and TetR (Xu *et al.*,1997). Although, *e.g.*, Gal4 is a yeast protein, it is functional also in other organisms, such as *Drosophila* cells and human cells.

Additional suitable AD or DBD components for use in the detection systems of the present invention can be identified using the following exemplary assays.

To identify an alternative AD, an effector compound detection system described in the Examples can be modified to exchange the GAL4 AD for a candidate AD. Specifically, a tc3 riboswitch, a URA3 reporter gene, a SPAL10 promoter comprising an URS1 repressor-binding site and response elements as described in Example 2 can be used. Using the URA3 reporter gene in a *S. cerevisiae* host, colony formation should then typically not be permitted at a frequency higher than 10^{-4} plated cells, more preferably not higher than 10^{-5} plated cells, and most preferably not at a frequency higher than 10^{-6} plated cells, when plating on test synthetic complete medium lacking leucine and tryptophan (for plasmid maintenance) and containing 1 % 5-FOA as described in Example 2. Note that if other DBDs or IDs have been chosen then that could change the expression strength required for the AD, and the candidate AD should then be tested together with that particular combination of DBD and IDs.

To identify an alternative DBD, an effector compound detection system described in the Examples can be modified to exchange the GAL4 DBD for a candidate DBD. Specifically, a tc3 riboswitch, a GAL4 AD, the SPAL10 promoter comprising an URS1 repressor-binding site as described in Example 2 can be used. The candidate DBD then replaces the GAL4 DBD, while candidate DNA-binding sites are inserted in exchange of DNA-binding sites for GAL4 in the SPAL10 promoter (marked as UAS in disclosed sequence annotations). Using the URA3 reporter gene in a *S. cerevisiae* host, colony formation should then typically not be permitted at a frequency higher than 10^{-4} plated cells, more preferably not higher than 10^{-5} plated cells, and most preferably not at a frequency higher than 10^{-6} plated cells, when plating on test

synthetic complete medium lacking leucine and tryptophan (for plasmid maintenance) and containing 1 % 5-FOA as described in Example 2. Note that if other ADs or IDs have been chosen then that could change the expression strength required for the DBD, and the candidate DBD should then be tested together with that particular combination of AD and IDs.

In eukaryotic host cells, the first and second hybrid proteins should preferably be localized to the cell nucleus. If no native nuclear localization sequence (NLS) is present within the first, second and third hybrid protein sequences, one can be translationally fused. One example of an NLS is the SV40 NLS sequence.

Preferred choices of AD and DBDs for the first and second hybrid proteins are shown in Table 2, together with the response element(s) to which the DBD binds. Suitable AD and DBD pairs can then be combined by engineering via the interaction domain pairs. The most preferred systems are the Gal4 DBD together with one of the Gal4 AD and VP16.

TABLE 2

AD	**Reference**
Gal4 AD	Fields and Song, 1989
VP16 AD	Berger et al., 1990
B42	Gorska *et al.*, 2006
GCN4	Hope and Struhl, 1986
DBD (response element)	**Reference**
Gal4 DBD (GAL4 UAS)	Fields and Song, 1989
Zif268 (Zif268 binding site)	Bae *et al.*, 2003, Joung *et al.*, 2000
Estrogen receptor DBD (Estrogen receptor elements)	Stynen *et al.*, 2012
LexA (lexA operator)	Gorska *et al.*, 2006
QA-1F (QUAS) (mammalian cells)	Potter *et al.*, 2010

The first and second hybrid proteins associate with each other via the first and second interaction domains or "IDs". These associate (herein referred to as 'bind') to each other via the first and second interaction domains. The IDs can, for example, be selected so as to have a lower mutual binding affinity than the affinity of the DBD to the DNA-binding site(s) of the response element. The binding affinity should further be retained upon fusion to the AD and DBD molecules. The IDs can be selected from proteins, which are known to interact, for

example as identified using the yeast-two-hybrid technology, or could for example be heterodimerizing leucine zippers. Further, while the classic yeast two-hybrid format was designed to screen for 'bait' and 'prey' proteins binding to each other when fused to AD and DBD domains, respectively, suitable interacting domains for use in the context of the present invention can of course be selected from any such 'bait and 'prey' combination found in such an assay to allow the AD and DBD domains to form a functional transcription factor. Suitable assays to identify additional ID pairs thus include those described in the references cited herein for yeast two-hybrid and related systems.

Additional suitable ID pairs for use in the detection systems of the present invention can be identified using the following exemplary assays. To identify an alternative ID pair, an effector compound detection system described in the Examples can be modified to exchange the ID pair for a candidate ID pair. Specifically, a tc3 riboswitch, a GAL4 AD, a GAL4 DBD and SPAL10 promoter can be used. Using the URA3 reporter gene in *S. cerevisiae*, colony formation should then typically not be permitted at a frequency higher than 10^{-4} plated cells, more preferably not higher than 10^{-5} plated cells, and most preferably not at a frequency higher than 10^{-6} plated cells, when plating on test synthetic complete medium lacking leucine and tryptophan and containing 1 % 5-FOA as described in Example 2 Furthermore, a test strain should be constructed devoid of the ID2-AD molecule and using the same test conditions that strain should fully form colonies under the same test conditions.

Table 3 lists preferred IDs for use in the present invention, including the most preferred; the rat/human proteins RalGDS and Krev1.

TABLE 3

Interaction domain pairs	**Reference**
Krev1 + RalGDS	Serebriiskii *et al.*, 1999
Heterodimerizing leucine zippers	Moll *et al.*, 2001
GlnB + NtrB	Pawlowski *et al.*, 2003

In a particular embodiment, the effector compound detection system comprises a third DNA molecule encoding a "third" protein which can bind to the first hybrid protein ('AD-ID1') via the first interaction domain but which cannot bind to the response element. This protein suitably comprises the second interaction domain or a variant or fragment thereof which can bind to ID1, but does not comprise the DBD or any DBD variant capable of binding the response element. In a host cell, the effect of this third protein is to bind to false first hybrid proteins (*e.g.* in presence of the effector compound), thereby reducing noise, since these complexes cannot lead to reporter gene activation at any time. For example, in a host cell where the first hybrid protein is DBD-Krev1 and the second hybrid protein is AD-RalGDS, the

third protein can be RalGDS or a hybrid protein of RalGDS and an irrelevant protein, *i.e.*, one not interacting with any other component of the buffer device.

In case the buffer device is intended for use in a particular type of host cell, certain considerations should be made in selecting the AD, DBD and interaction domains. For example, as indicated above, the chosen AD should be capable of interacting with endogenous transcription co-regulators or other factors regulating transcription factor activity in the selected host cell. The AD should also be selected such that the intended host cell does not contain endogenous repressor elements significantly reducing the function of this particular AD. The DBD, in turn, should not be endogenously expressed or present in the host cell of interest, so as to avoid the second hybrid protein competing with an endogenous protein in binding to the DBD-binding site in the response element. Likewise, the host cell should be free of endogenously expressed proteins binding the first or second interaction domains. In such cases, however, the genes encoding any such endogenous protein(s) could alternatively be deleted from the genome or its/their transcription or expression blocked, as described below. The skilled person is aware of these considerations, since they equally apply to the well-known yeast-two-hybrid systems.

Signal inversion

Provided by the present invention is also a riboswitch-based detection system by which the signal from the riboswitch can be inverted, e.g., so that an OFF signal from an OFF-riboswitch is changed to an ON output from the detection system, i.e., increased expression of the reporter gene. Specifically, this is made possible by the repressor protein inverting the signal while surplus AD molecules function as a signal buffer. This is exemplified in Example 6, wherein the riboswitch of the detection system controls the expression of a transcriptional repressor protein, which functions by masking of the effect of the AD. Typically, in this aspect, the riboswitch is an OFF riboswitch, directly linked to the transcriptional repressor protein instead of the reporter gene. Also, the AD and DBA hybrid proteins are typically expressed at same strength as the transcriptional repressor protein, or alternatively AD and DBD are expressed within a single protein (e.g. *S. cerevisiae* GAL4) with the same strength as the transcriptional repressor. Exemplary riboswitches, ADs and DBDs, and ID pairs for this aspect of the invention are described in Tables 1-3. Exemplary repressor proteins for this aspect include, but are not limited to the GAL80 repressor protein from *S. cerevisiae*. Exemplary single proteins containing both AD and DBD include, but are not limited to GAL4 from *S. cerevisiae*.

Actuator platform

As shown in Figure 1, the typical layout of the actuator platform comprises a response element which comprises one or more binding sites for the DBD and which is operably linked to the reporter gene, typically via a promoter. So, when a second hybrid protein ("DBD-ID2") is simultaneously bound to a first hybrid protein ("AD-ID1") and a DBD-binding site in the

response element, transcription of the reporter gene is induced. Generally, the response element is located upstream of the minimal promoter region and upstream of optional upstream repressor binding sites.

The number of DBD-binding sites influences how strongly the presence of DBD activates transcription of the reporter gene. Accordingly, the number of DBD-binding sites within the response element is at least one, optionally between 1 and about 100, such as between 1 and about 50, 1 and about 30, or 1 and about 10. In one embodiment, the number of DBD-binding sites within the response element is at least 2, such as between 2 and about 100, 2 and about 50, or 2 and about 20. In another embodiment, the number of DBD-binding sites within the response element is at least 2, at least 3, a least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, 11, 12, 13, 14, 15 or more, optionally at most about 100, or at most about 20.

Exemplary promoters for the reporter gene include, but are not limited to bacterial P_{spc}, P_L, and P_{bla}, and, for yeast, *GAL1*, *ADH1, CYC1* and the synthetic promoter *SPAL10* (yeast) (Vidal *et al.*, 1996). Optionally, the promoter comprises an upstream repressing sequence, for the purpose of ensuring no or minimal basal expression of the reporter gene as described, *e.g.*, in WO 96/32503. In a preferred embodiment, the DBD is the Gal4 DBD and the reporter gene cassette is located downstream of the specific Gal4 DNA-binding sites within a minimal promoter such as, *e.g.*, *SPO13*, comprising a repressor binding site to prevent basal expression, as previously engineered for a classical yeast two-hybrid strain (Vidal et al (1996). Other promoters can likewise be engineered to comprise upstream repressor sequences, using conventional cloning techniques.

The reporter gene is a gene whose transcription or expression can be directly or indirectly detected, for example under a predetermined condition. Typically, the host cell does not contain an endogenous copy of the reporter gene, or a gene providing a similar signal. A wide variety of genes known in the art are suitable for use as reporter genes in the present invention. For example, in embodiments where it is desirable to identify, select and/or isolate cells where the reporter gene is expressed ("positive selection"), the reporter gene can be a gene whose expression provides antibiotic resistance, complement an auxotrophy in the host cell, allow growth under specific conditions that do not permit growth of host cells that do not contain or express this same gene, or a gene encoding a protein which leads to a chromogenic, luminescent, fluorescent signal, either in itself (such as, e.g,, GFP) or an enzyme capable of converting a substrate to a directly detectable product (*e.g.*, lacZ). In embodiments, where it is desirable to identify, select and/or isolate cells where the reporter gene is not expressed ("negative selection"), the reporter gene can be a counter-selection gene, including, but not limited to, *URA3* (expression of which is lethal in the presence of 5-FOA) and SacB (expression of which is lethal in the presence of sucrose).

The actuator platform can be part of the effector molecule detection system, for example, comprised in a fourth DNA molecule. In such an embodiment, the DNA molecules of the actuator platform and the other components of the effector molecule detection system

can be part of the same kit and/or jointly transfected or otherwise introduced into the host cell. In another embodiment, the actuator platform and the other parts of the effector molecule detection system are separately introduced into the host cell. The DNA molecule comprising the actuator platform may as such be part of a plasmid, episome or a chromosome in the host cell, the latter achieved, *e.g.*, by homologous recombination. For example, in the case of an *URA3* reporter gene and an *S. cerevisiae* host cell, the actuator platform can be comprised in a yeast integration vector containing two fragments homologous to the upstream and downstream regions of the open reading frame of the endogenous *URA3* gene and allowing integration by homologous recombination at the level of the locus *URA3* of a sequence inserted between these two fragments.

DNA molecules

The DNA molecules of the effector molecule detection system of the invention can be comprised in, *e.g.*, plasmids, episomal DNA or chromosomal DNA. In one embodiment, the first and second DNA molecules, encoding the "biosensor and AD-ID1" constructs and the "DBD-ID2" protein, respectively are comprised in the same plasmid, episome or are both incorporated into chromosomal DNA. In another embodiment, the first and second DNA molecules are comprised in separate plasmids or episomes or located so that one of the DNA molecules is on a plasmid or episome while the other is incorporated in chromosomal DNA.

In a particular embodiment, the first and second DNA molecules are both comprised in the same or in two separate expression plasmids, each operably linked to a promoter, which can be an inducible or constitutive promoter. In another particular embodiment, the first and second DNA molecules are both comprised in the chromosome, each operably linked to a promoter, which can be an inducible or constitutive promoter. In one embodiment, the first and second DNA molecules are both linked to inducible promoters, optionally induced by the same molecule. Inducible promoters suitable for this purpose include, but are not limited to, those of Cu^{2+}-inducible CUP1, and arabinose-inducible araBAD. In one embodiment, the first and second DNA molecules are both operably linked to constitutive promoters. Suitable constitutive promoters include, but are not limited to, the ADH1, TPI, TEF1, TDH3, KEX2 and ACT1 promoters for yeast cells; the CMV, EF-1 a, SV40, UB, RSV LTR promoters for use in mammalian cells, and in prokaryotic cells P_{spc}, P_L and P_{bla}.

The expression vector may also comprise additional regulatory elements suitable for regulating the transcription of the DNA molecules in the host cell in question. Apart from promoters, regulatory elements suitable for a prokaryotic cell include, for example, operator sequences and ribosome binding sites. In eukaryotic cells, apart from promoters, polyadenylation signals, promoter-proximal elements, UASs, and enhancers can be used. In general, regulatory elements can include promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. Although, in some cases, particular regulatory control sequences are preferred for specific types of host cells, expression vectors can also be

designed so that they can propagate in two different host species, in so-called shuttle vectors. For example, in one embodiment, one or both of the first and second DNA molecules, optionally together with the third and fourth DNA molecules, are inserted into a shuttle vector suitable for both prokaryotic (*e.g.*, bacterial) and eukaryotic (*e.g.*, yeast) cells. Examples include a yeast shuttle vector including an origin of replication, a selectable marker (*e.g.* for antibiotic resistance), an autonomously replicating sequence (ARS), a yeast centromere (CEN), and a yeast selectable marker) and adenovirus shuttle vectors, which can propagate in both *E. coli* and mammalian cells. Further, for integration into a chromosome, the DNA molecules can further comprise sequences enabling homologous recombination into a chosen site of a chromosome, preferably one selected for minimal interference from surrounding elements.

In one embodiment, the regulatory elements for DNA molecules encoding a buffer device comprising an OFF-riboswitch are selected so that, in the absence of the effector compound, the ratio of the expression level of the second interaction domain (whether comprised in the DBD-ID2 second hybrid protein or the optional third 'ID2' protein) over the first interaction domain (which is part of the AD-ID1 first hybrid protein) in the host cell is estimated or predicted to be in the range of between about 0.01 to about 100, such as about 0.03 to about 40, such as about 0.1 and about 9, such as about 0.2 and about 5, such as about 0.3 and about 3, such as about 0.5 and about 2, such as about 0.75 and about 1.5, such as about 0.9 and about 1.1, or in the range of about 0.01 and about 5, such as about 0.05 and about 3, such as about 0.1 and about 2, such as about 0.5 and about 1, such as about 1. In one preferred embodiment, the ratio is between 0.05 and 20. In another preferred embodiment, the ratio is between 0.1 and 9.

In one embodiment, the regulatory elements for DNA molecules encoding a buffer device comprising an ON-riboswitch are selected so that, in the presence of the effector compound at a predetermined level, e.g., the concentration to be detected, the ratio of the expression level of the second interaction domain (whether comprised in the DBD-ID2 second hybrid protein or the optional third 'ID2' protein) over the first interaction domain (which is part of the AD-ID1 first hybrid protein) in the host cell is estimated or predicted to be in the range of between about 0.01 to about 100, such as about 0.03 to about 40, such as about 0.1 and about 9, such as about 0.2 and about 5, such as about 0.3 and about 3, such as about 0.5 and about 2, such as about 0.75 and about 1.5, such as about 0.9 and about 1.1, or in the range of about 0.01 and about 5, such as about 0.05 and about 3, such as about 0.1 and about 2, such as about 0.5 and about 1, such as about 1. In one preferred embodiment, the ratio is between 0.05 and 20. In another preferred embodiment, the ratio is between 0.1 and 9.

In another embodiment, the ratio of the number of molecules comprising the second interaction domain (whether comprised in the DBD-ID2 or the optional third 'ID2' protein) to the number of binding sites for the DBD in the response element is at least about 2, such as about 2 to about 5000, such as about 2 to about 500, such as about 2 to about 100, such as

about 2 to about 50, such as about 2 to about 25. Without being bound to theory, under these conditions, in the presence of the effector compound, the probability of interaction between a falsely present first protein and the second protein bound to a response element is considerably below the probability of interaction between that first protein and a free second protein, thus providing reduction of noise, since binding of first protein to free second protein does not confer reporter gene activation.

In some cases, it is possible that the desired expression level of the reporter gene differs considerably from the optimal expression level of the riboswitch regulator for a given application. The response of a pool of riboswitches may *e.g.* be less variable if the number of riboswitches in the pool is high, compared to a low number. This is a general biochemical concept for noise (McAdams and Arkin, 1999). However, since riboswitches are normally operably linked directly to the reporter gene, any changes in the expression level of the riboswitch must also directly affect the reporter gene expression level. Using the invented buffer device, it is possible to express both the first and the second hybrid protein from a high or low expression level, respectively using weak or strong promoters as exemplified below, while engineering the response elements into a weak promoter (constructing *e.g. SPAL10*) for the reporter gene. Without being limited to theory, a system with the riboswitch being operably linked to a nucleic acid sequence encoding a single, full-length transcription factor (as opposed to the effector compound detection system of the invention) which in turn regulates the reporter gene would be less advantageous, since in this case it would only be possible to regulate the single transcription factor within a particular range defined by its binding affinity to the response element of the reporter gene. Outside of that constrained range, the transcription factor would constitutively saturate or not saturate the DNA binding site(s) of the response element, leading to no control of reporter gene regulation.

In one embodiment, the promoter for the second DNA molecule is weaker than that for the first DNA molecule. Under conditions in which the biosensor is 'switched on', permitting transcription or translation of the coding sequence for the first (AD-ID1) hybrid protein, the ratio of the level of AD-ID1 to DBD-D2 (and any third ID2' protein) molecules in the host cell is then increased as compared to having promoters of the same strength controlling the first and second DNA molecules. For example, in a yeast host cell, the promoter for the first DNA molecule could be a strong promoter, optionally selected from, *e.g.*, the *ADH1*, *ENO2*, *FBA*, *PYK1* and *TEF1* promoters, while the promoter for the second DNA molecule could be a weak promoter relative to the first promoter, optionally selected from, *e.g.*, the CUP1 promoter (in the absence of copper ions beyond the standard supplement in synthetic complete medium), *MET25*, *STE5* and the minimal promoter of *CYC1*. As shown in Example 3, increasing the relative strength of the first over the second promoter increases the detection level of the effector molecule detection system, *i.e.*, so that cell proliferation under selection conditions is exclusively allowed in presence of a higher concentration of effector compound. Without being bound by theory, this effect can be used to design the effector molecule detection system to not only screen for the mere presence or production of an effector compound but

to actually optimize the system to screen for a particular concentration, *e.g.*, production level, of the effector molecule in a library of host cells.

So, in one embodiment, the regulatory elements for DNA molecules encoding a buffer device comprising an OFF-riboswitch are selected so that, in the absence of the effector compound, the ratio of the expression level of the second interaction domain (whether comprised in the DBD-ID2 or the optional third 'ID2' protein) over the first interaction domain (which is part of the AD-ID1) in the host cell is estimated or predicted to be in the range of between about 0.01 and about 10, such as about 0.02 and about 2, such as about 0.05 and about 1.5, such as about 0.1 and about 1.5, such as about 0.1 to about 1, such as about 0.5 and about 1, such as about 1. In one preferred embodiment, the ratio is between 0.1 and about 1. In another preferred embodiment, the ratio is about 1.

In one embodiment, the regulatory elements for DNA molecules encoding a buffer device comprising an ON-riboswitch are selected so that, in the presence of the effector compound at a predetermined concentration, *e.g.*, the concentration to be detected, the ratio of the expression level of the second interaction domain (whether comprised in the DBD-ID2 or the optional third 'ID2' protein) over the first interaction domain (which is part of the AD-D1) in the host cell is estimated or predicted to be in the range of between about 0.05 and about 10, such as about 0.1 and about 2, such as about 0.1 and about 1.5, such as about 0.5 and about 1, such as about 1. In one preferred embodiment, the ratio is between 0.1 and about 1. In another preferred embodiment, the ratio is about 1.

As described herein, the effector compound detection system of the invention can be modulated to as to increase or decrease the sensitivity of the detection system, i.e., change the level of effector compound to which the detection system responds. To increase the concentration of effector compound necessary for the detection system to respond, the expression level of the first hybrid protein must be increased relative to the expression level of the second hybrid protein. This can e.g. be achieved using a stronger promoter for the first hybrid protein than for the second hybrid protein. Using this method, more effector compound molecules are needed in the cell to convey a given response. The range within which this change of sensitivity is possible can be changed by varying the expression strengths of first to second hybrid protein and measuring the activity of the reporter gene with and without the effector compound. A maximum limit for changing the detection level in this fashion will exist and this can be determined experimentally for each detection system. So, in one embodiment, the regulatory elements for DNA molecules encoding a buffer device comprising an OFF riboswitch are selected so that, in the absence of the effector compound, the ratio of the expression level of the second interaction domain (whether comprised in the DBD-ID2 or the optional third 'ID2' protein) over the first interaction domain (which is part of the AD-ID1) in the host cell is estimated or predicted to be in the range of between about 0.01 and about 1, such as about 0.01 to about 0.9, such as about 0.01 and about 0.7, such as about 0.01 and about 0.5, or between about 0.02 and about 0.9, such as about 0.05 and

about 0.9, such as about 0.1 and about 0.9, such as about 0.5 to about 0.9. In one preferred embodiment, the ratio is between 0.1 and about 0.9.

Using the same principle, to decrease the sensitivity for the ligand, the relative expression level of first hybrid protein to second hybrid protein must be decreased. Likewise, the range within which this change of sensitivity is possible can be changed by varying the expression strengths of first to second hybrid protein and measuring the activity of the reporter gene with and without the effector compound. So, in one embodiment, the regulatory elements for DNA molecules encoding a buffer device comprising an OFF-riboswitch are selected so that, in the absence of the effector compound, the ratio of the expression level of the first interaction domain (whether comprised in the DBD-ID2 or the optional third 'ID2' protein) over the second interaction domain (which is part of the AD-ID1) in the host cell is estimated or predicted to be in the range of between about 0.01 and about 1, such as about 0.01 to about 0.9, such as about 0.01 and about 0.7, such as about 0.01 and about 0.5, or between about 0.02 and about 0.9, such as about 0.05 and about 0.9, such as about 0.1 and about 0.9, such as about 0.5 to about 0.9. In one preferred embodiment, the ratio is between 0.1 and about 0.9.

Kits comprising the first and second, optionally also the third and/or fourth, DNA molecules described herein, typically in the form of plasmids or expression vectors, are also envisioned. Such a kit may further comprise instructions for use, *e.g.*, in transfecting one or more host cell populations with the plasmids or vectors.

Specific, exemplary combinations of regulatory elements and designs of expression vectors are provided in Figure 2 and Table 6, as well as in SEQ ID NOS:1-18, representing *e.g.* vectors used in the host cell systems reported in Table 5.

Host cells

The host cell of the present invention may be any host cell capable of providing the cellular transcription and translation machinery required for transcription and/or expression of the effector molecule detection system. The host cell can, for example, be a microorganism or microbial cell, such as a prokaryotic or eukaryotic microbial cell, or a plant cell, animal cell, or the like. Suitable prokaryotic host cells include, for example, *archaea* and bacterial host cells, including, but not limited to, *Bacillus*, *Pseudomonas* and *Escherichia*, including *E. coli* host cells. Suitable eukaryotic microbial host cells include yeast cells, fungi or protist cells, including, but not limited to, *Pichia*, *Schizosaccharomyces* and *Saccharomyces*, including *S. cerevisiae* host cells. Suitable animal host cells include cells or cell lines from insects, birds or reptiles, as well as mammalian cells or cell lines, including human cells or cell lines.

As indicated above, the host cell preferably does not endogenously express any proteins comprising the DBD or first or second interaction domains, nor any proteins interfering with the binding between the first and second domains or the DBD and the DBD-binding site in the response element, nor any AD-specific repressors. The suitability of the combination of the specific buffer device and host cell can be tested, e.g., by constructing a

host cell which expresses the intended combination of first and second hybrid proteins and actuator platform (though not under the control of a biosensor) and measuring activity of the reporter gene (e.g. *URA3* in a colony formation assay as in the Example 2). This activity should then typically be higher than the activity of a control host cell where the sole difference from the first host is the obstruction of interaction between the first and second hybrid proteins (e.g. by removal of an interaction domain). If the difference in reporter gene activity is suitable (e.g. above 10^4 fold difference in colony formation with the *URA3* reporter gene), the host cell can typically be used without further modification. However, if the difference in activity of the reporter gene is not allowing suitable discrimination, the host cell can, for example, be exchanged for another host cell or modified. For example, if any of the chosen components, such as the DBD, ID1, ID2 or any AD-specific repressors are endogenously present in the host cell, the gene(s) encoding the component(s) can be deleted, inactivated or its/their transcription or expression blocked prior to or in conjunction with the introduction of the effector compound detection system, using standard techniques known in the art. For example, the endogenous *GAL4* and *GAL80* genes had been deleted in all *S. cerevisiae* strains constructed from MaV203 in the Examples.

In one aspect, the invention relates to a host cell comprising the effector compound detection system, including the actuator platform (*i.e.*, the reporter gene operably linked to the response element), and at least one candidate modulator compound which either (i) promotes (*e.g.*, catalyzes) the transformation of a substrate into the effector compound, (ii) promotes the degradation of the effector compound into another (product) compound; (iii) transports the effector compound into the host cell, or (iv) transports the effector compound out of the cell. In embodiments (i) and (ii) the candidate modulator compound is typically a biocatalyst such as, but not limited to, an enzyme, a complex of enzymes, an enzymatic pathway, a combination of an enzyme and a co-factor, a biocatalytic RNA molecule, or the like. In embodiments (iii) and (iv), the candidate modulator compound is typically an influx or efflux pump, which may be able to transport the effector compound from the extracellular space to the intracellular space, or vice versa.

In all of (i) to (iv), the effector molecule detection system can then be applied to detect whether the candidate modulator compound is an actual modulator compound. As will be apparent to the skilled person after reading the present disclosure, the various possible combinations of biosensor (ON or OFF), reporter gene (chosen for positive or negative selection) and relative expression levels of the components (described in the preceding section) provide for a range of possibilities to optimize the detection or identification of a host cell or host cell population which either comprise or do not comprise a modulator compound at the level the buffer device is optimized for.

However, just by way of example, one possible assay to detect or identify a host cell comprising a modulator compound capable of converting a substrate into the effector compound using the noise-reducing system of the invention, is to select an OFF riboswitch as the biosensor and a reporter gene providing negative selection (*e.g.*, *URA3* in the presence of

5-FOA). In the presence of an active modulator compound, the effector compound is produced and binds the riboswitch. The riboswitch, in turn, switches off translation of the first (AD-D1) protein, and any 'leaked' AD-D1 is buffered by the surplus, free DBD-D2. Then, transcription of the reporter gene is not activated and the cell cannot convert the added 5-FOA into the toxic 5-fluorouracil, and the host cell therefore proliferates and can be identified. In a host cell where the effector compound is not produced, the riboswitch permits translation of the AD-D1, allowing a functional transcription factor to be formed at the response element, whereupon the *URA3* gene expression product converts the added 5-FOA to 5-fluorouracil and the cell then dies or can at least not proliferate. Switching the reporter gene to one suitable for positive selection (*e.g.* also *URA3* with a selection medium devoid of uracil) but maintaining an OFF riboswitch would be suitable for detecting a cell where a modulator compound degrades the effector compound, since, in such a cell, the riboswitch then permits translation of the AD-D1, allowing transcription of the reporter gene and detection of the host cell by survival on selection medium lacking uracil.

Screening methods

According to the same principles, a library or another heterogenous population of host cells comprising candidate modulator compounds can be screened for the presence or absence of actual modulator compounds, and one or more host cells or clones identified as providing the desired reporter gene output can be identified, selected and/or isolated.

A library of host cells can be produced by digesting, *e.g.*, all DNA from a suitable source organism with a restriction enzyme or by sonication and cloning the fragments into a vector, which will result in the production of many different recombinant vectors, each with a different fragment of DNA cloned into it. The collection of many different recombinant vectors together forms a vector library. The library may be produced by any method available to the skilled person, for instance by shotgun cloning. "Shotgun cloning" refers to using the whole genome of an organism as the starting point for cloning. Alternatively, DNA/RNA can be isolated from the environment, resulting in thousands of different genes, which may be derived from many different organisms; such a collection of nucleic acid material is referred to as a metagenome. At the other end of the spectrum, specific libraries may be constructed that contain only genes expected to encode the candidate modulator compounds, based for instance on DNA homologies, previous screening experiments, or commercially available libraries. A vector library constructed by shotgun cloning techniques may contain from a few thousand (microorganisms) to hundreds of thousands (higher eukaryotes or microbial metagenomes) of different recombinant plasmids while only one or a few comprise a sequence of interest.

In order to reduce the size of libraries derived from eukaryotic organisms, it is sometimes advantageous to produce cDNA libraries. Since protein coding regions generally account for only a small percentage of the total genome size of eukaryotic organisms and since eukaryotic cells generally express only a subset of their genes, the production of cDNA

copies of the cellular mRNA using reverse transcriptase and the cloning thereof into a cDNA library results in a library of reduced size comprising only the part of the organism's genome that is of most interest (the protein coding sequences that are expressed). cDNA library construction is well known to the skilled person. Other library types include, but are not limited to, a library obtained by (i) mutation (directed or random) of a single parent gene, (ii) recombination of several parent genes (such as libraries produced by gene shuffling), and (iii) a combination thereof.

For a host cell library, the host or host cell population may already have the desired biocatalytic activity, and the aim of applying the invention is to develop an improved biocatalyst. The population of potential biocatalysts may, for instance, be provided in the form of a mixed population of potentially biocatalytically active bacterial cells, *i.e.* cells of different strains, for instance isolated from a soil or water sample, to which cells the effector compound detection system is added, in order to select the best strain, selected for the optimum production or degradation of the effector compound, whichever is desired.

The individual genes that together form the library or population of potential biocatalysts may be encoded on a plasmid, or chromosomal or episomal DNA. The genes that together form the library or population of potential biocatalysts may be expressed constitutively; or they may be on an inducible expression vector, in which case they are expressed as a result of activation through one of various possible operably linked regulatory elements. In fact, it is not limiting how and if the biocatalyst expression is induced. Expression may be constitutive, it might be under control of a native regulator, either known or not yet known, or it might be part of a library under control of a specifically designed induction system. Regardless of the way in which the biocatalyst gene(s) is/are expressed, it is preferred that expression is sufficient to allow proliferation of the host cell. In a further preferred embodiment of the above-described aspects at least one microbial host cell may comprise one (or more) nucleic acid molecule(s) encoding at least two or more candidate biocatalysts, each capable of catalyzing at least one reaction of a multi-step chemical conversion reaction. Preferably, the product of the reaction catalyzed by one potential biocatalyst is the substrate for the reaction catalyzed by another potential biocatalyst, wherein at least one of said products is an effector compound capable of switching the riboswitch. In this way, complete biochemical pathways can be uncovered and specific biocatalysts selected for each reaction step.

In another embodiment, nucleic acid molecules encoding the various potential biocatalysts may be provided in separate expression systems, each expression system facilitating the expression of separate candidate biocatalysts. Also in this manner, multi-step chemical conversion reactions can be provided for. The host organism may or may not itself comprise enzymes and optional co-factors for catalysing a multi-step reaction in which the substrate is converted into a final intermediate compound, which said final intermediate compound is then converted into the desired product by a potential biocatalyst. The final intermediate compound may be a natural intermediate in the host cell that is produced by an

available enzyme or a set of enzymes which is/are native to the host cell. Alternatively, the potential biocatalyst may produce an intermediate compound, which may or may not be an effector compound capable of switching the riboswitch, and which can be converted into a product by an enzyme endogenous to the host organism under the conditions provided. Thus, the potential biocatalyst need not necessarily catalyze the final reaction in a multi-step reaction process. It should be noted that in such instances, the enzyme system present in the host cell that contributes towards forming the effector compound for switching the riboswitch in such a multi-step system, may be encoded on host cell chromosomal or episomal DNA, including possibly plasmid DNA that also encodes the effector compound detection system itself. Thus, apart from potentially catalysing the final reaction towards formation of the product, the potential biocatalyst as defined herein may also catalyze the critical reaction in a possible chain of reactions of a multi-step reaction process, which critical reaction is the sought after reaction that ultimately leads (possibly via additional reaction steps) to the production of the sought after product. Thus, the host organism may or may not comprise one or more nucleic acid molecules encoding at least two biocatalysts each capable of catalyzing at least one reaction of a multistep chemical conversion reaction.

In yet another preferred embodiment of any of the above-described aspects, any host cell identified as comprising a biocatalyst may or may not gain an increase in biocatalytic activity by genome engineering of the host genome. Methods for genome engineering are known to persons skilled in the art and include, but are not limited to, multiplex genome engineering (MAGE), mutator strains, chemical mutagenesis, genome shuffling, and genome synthesis. In yet another preferred embodiment of the above-described aspects, the selection method may be used to identify a host cell or cells which survive in higher concentrations of the desired product, such that increased levels of production of the product will result in the escape of premature death of the host cell. The chemical conversion reaction may be any chemical reaction which benefits from activation or acceleration by a biocatalyst. Reaction types that can be catalyzed by the potential biocatalyst include, but are not limited to, reactions catalysed by each of the major enzyme classes and subclasses: oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases.

The following examples are provided for illustration purposes only and are not intended to limit the invention in any way.

EXAMPLE 1

This Example describes the genetic construction of buffer devices and controls. Table 4 shows exemplary functional buffer strains and their genetic features.

The buffer device GAL4 DBD was fused to the first interaction domain Krev1 and the gene was transcribed from a promoter (P1), together with an AD fused to the second interaction domain RalGDS transcribed from a promoter (P2), as described in Table 4. These

two interaction domains are known to interact and are described in literature (Serebriiskii *et al.*, 1999). Both constructs propagated on plasmids or chromosome (O, see Table 4).

In buffer devices, the ligand-responsive OFF riboswitch controlled the expression of the AD-RalGDS hybrid protein. The riboswitch selected as a model system was the tetracycline-responsive "tc3" riboswitch (Kötter *et al.*, 2009). The actuator platform controlling the transcription of the *URA3* reporter gene was composed of the SPAL10 promoter originating from a previous study (Vidal *et al.*, 1996) driving expression of the *URA3* gene followed by the natural terminator sequence for *URA3*. The response elements for the GAL4 DBD and a URS1 repressor-binding site were thus contained in the SPAL10 promoter. For controls representing conventional riboswitch regulation, the sequence encoding the riboswitch was fused directly to the sequence encoding the reporter gene URA3 (PRd3, PRd4, PRd5). The control strain PRa24 also has the riboswitch fused directly to the URA3 sequence, but with expression driven by constitutive AD and DBD from the SPAL10 promoter.

Table 5 shows a full list of genotypes for all strains that propagated plasmids. Table 6 describes plasmids, including selected annotations of key components so as to identify the nucleic acid sequence segment encoding them in each SEQ ID number. The corresponding amino acid sequence of each protein component encoded can be obtained via on-line translation tools (available at, e.g., bioinformatics portals such as expasy.org/tools/) based on the standard genetic code. See also Figure 2.

TABLE 4

Strain	**Promoter P1**	**Propagation O**	**AD**	**Reporter propagation**	**Promoter P2**
PRa4	ADH1	ARS/CEN plasmid	*S. cerevisiae* GAL4 AD	Chromosome	ADH1
PRa22	ADH1	ARS/CEN plasmid	*S. cerevisiae* GAL4 AD	ARS/CEN plasmid	ADH1
PRa46	CUP1	ARS/CEN plasmid	*S. cerevisiae* GAL4 AD	ARS/CEN plasmid	CUP1
PRa34	ADH1	Chromosome	*S. cerevisiae* GAL4 AD	Chromosome	ADH1
PRa41	ADH1	ARS/CEN plasmid	*Herpes simplex* VP16 AD	ARS/CEN plasmid	ADH1
PRa47	ADH1	ARS/CEN plasmid	*S. cerevisiae* GAL4 AD	ARS/CEN plasmid	CUP1
PRa74	ADH1	ARS/CEN plasmid	*S. cerevisiae* GAL4 AD	ARS/CEN plasmid	ADH1

PRa78	ADH1	ARS/CEN plasmid	*S. cerevisiae* GAL4 AD	ARS/CEN plasmid	ADH1
PRa79	ADH1	ARS/CEN plasmid	*S. cerevisiae* GAL4 AD	ARS/CEN plasmid	CUP1
PRa84	ADH1	ARS/CEN plasmid	*S. cerevisiae* GAL4 AD	ARS/CEN plasmid	ADH1

Strains:

Saccharomyces cerevisiae MaV203 (MATα, *leu*2-3,112, *trp*1-901, *his*3Δ200, *ade*2-101, *gal4*Δ, *gal80*Δ, *SPAL*10::*URA3*, *GAL1::lacZ*, $HIS3_{UAS\ GAL1}$::*HIS3@LYS2*, *can*1^R, *cyh*2^R)

Saccharomyces cerevisiae PRa18 (MATα, *leu*2-3,112, *trp*1-901, *ura3Δ, his3*Δ200, *ade*2-101, *gal4Δ*, *gal80*Δ, *GAL1::lacZ, HIS3*$_{UAS\ GAL1}$::*HIS3@LYS2, can*1^R, *cyh*2^R). Constructed from MaV203.

Saccharomyces cerevisiae PRa26 (MATα, *leu*2-3,112, *trp*1-901, *ura3Δ, his3*Δ200, *ade*2-101, *gal4Δ*, *gal80*Δ, *GAL1::lacZ, kanMX, can*1^R, *cyh*2^R). Constructed from PRa18.

Saccharomyces cerevisiae PRa34 (MATα, *leu*2-3,112, *trp*1-901, *ura3Δ, his3*Δ200, *ade*2-101, *gal4Δ*, *gal80*Δ, *GAL1::lacZ, can*1^R, *cyh*2^R) with chromosomal integration of AD part and DBD-reporter part. Constructed from MaV203.

Saccharomyces cerevisiae CEN.PK2-1C (MATa; *ura*3-52; *trp*1-289; *leu*2-3,112; *his3*Δ1; *MAL2-8*^{C}; *SUC2*)

Saccharomyces cerevisiae PRd5 (MATa*; ura*3-52*; trp*1-289*; leu*2-3,112*; his3*Δ1*; MAL2-8*^{C}*; SUC2*) with chromosomal integration of direct riboswitch regulation of *URA3*. Constructed from CEN.PK2-1C.

TABLE 5

Strain	**Plasmid #1**	**Plasmid #2**	**Plasmid #3**	**Parent strain**
PRa4	pEXP22.1	pEXP32-Krev1	-	MaV203
PRa6	pEXP22RalGDSwt	pEXP32.1	-	MaV203
PRa21	pEXP22RalGDSm2	pEXP32.4	-	PRa18
PRa22	pEXP22.1	pEXP32.4	-	PRa18
PRa24	pEXP22RalGDSwt	pEXP32.3	-	PRa18
PRa28	pEXP22.2	pRS415-empty	-	PRa18
PRa41	pEXP22.3	pEXP32.4	-	PRa18
PRa45	pEXP22.1CUP	pEXP32.4	-	PRa18

PRa46	pEXP22.1CUP	pEXP32.4CUP	-	PRa18
PRa47	pEXP22.1	pEXP32.4CUP	-	PRa18
PRa56	pEXP22.2	pEXP32.4	-	PRa18
PRa74	pEXP22.1	pEXP32.4	*HIS3*-based plasmid without GFP	PRa26
PRa78	pEXP22.1CUP	pEXP32.4CUP	pEXP32.5XX	PRa26
PRa79	pEXP22.1	pEXP32.4CUP	pEXP32.5XX	PRa26
PRa84	pEXP42.2	pEXP22.1	pEXP32.4	PRa26
PRd3	pRP2-URA3	pRS414-empty	-	CEN.PK2-1C
PRd4	pRP4-URA3	pRS414-empty	-	CEN.PK2-1C

TABLE 6

The selected annotations list the starting and ending nucleic acid residue of selected key components of the plasmids in Table 5. Accordingly, the tc3 riboswitch has a sequence be encoded by residues 1724 to 1970 of SEQ ID NO:1, and so forth.

Selected annotations	SEQ ID NO:
LOCUS pEXP22.1 7718 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Promoter 272..1701 /label=pADH1 Riboswitch 1724..1970 /label=tc3 CDS 1981..2004 /label=NLS CDS 2011..2355 /label="GAL4 AD" CDS 2416..2712 /label=RalGDSwt Terminator 2991..3148 /label=tADH1 Replication_ori 3475..3930 /label=f1 CDS complement(4033..4707) /label=TRP1 Replication_ori 5243..5760 /label=ARS/CEN CDS 5892..6752 /label=AmpR Replication_ori 6897..7579 /label=pUC	1
LOCUS pEXP32-Krev1 11062 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Promoter 103..1557	2

/label=ADH1p CDS 1581..2024 /label="GAL4 DBD" CDS 2082..2630 /label=Krev1 Terminator 2915..3072 /label=ADH1t Replication_ori 3399..3854 /label=f1 CDS 4563..5657 /label=LEU2 replication_ori 6386..6903 /label=ARS/CEN CDS complement(7248..7781) /label=GentamicinR replication_ori 8629..9311 /label=pUC Terminator complement(9726..9915) /label="CYC1 terminator" CDS complement(10241..10690) /label=CYH2	
LOCUS pEXP32.1 9261 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Promoter 272..1701 /label=pADH1 Riboswitch 1724..1970 /label=tc3 CDS 1981..2424 /label="GAL4 DBD" CDS 2482..3033 /label=Krev1 Terminator 3315..3472 /label=tADH1 Replication_ori 3799..4254 /label=f1 CDS 4963..6057 /label=LEU2 Replication_ori 6786..7303 /label=ARS/CEN CDS 7435..8295 /label=AmpR Replication_ori 8440..9122 /label=pUC	3
LOCUS pEXP22RalGDSwt 7468 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Promoter 272..1726 /label=ADH1p CDS 1731..1754 /label=NLS CDS 1761..2105 /label="GAL4 AD" CDS 2166..2462 /label=RalGDSwt Terminator 2741..2898 /label=ADH1t Replication_ori 3225..3680 /label=f1 CDS complement(3783..4457) /label=TRP1 replication_ori 4993..5510 /label=ARS/CEN CDS 5642..6502 /label=AmpR replication_ori 6647..7329 /label=pUC	4

LOCUS pEXP32.3 10661 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Terminator complement(272..521) /label="URA3 terminator" CDS complement(522..1325) /label=URA3 Riboswitch complement(1336..1582) /label=tc3 Promoter complement(1592..1752) /label=pSPO13 URS complement(1679..1687) /label=URS1 Promoter complement(1776..1868) /label=5xUASgal UAS complement(1776..1792) /label=UASgal UAS complement(1795..1811) /label=UASgal UAS complement(1814..1830) /label=UASgal UAS complement(1833..1849) /label=UASgal UAS complement(1852..1868) /label=UASgal Promoter 1916..3370 /label=pADH1 CDS 3381..3824 /label="GAL4 DBD" CDS 3882..4433 /label=Krev1 Terminator 4715..4872 /label=ADH1t Replication_ori 5199..5654 /label=f1 CDS 6363..7457 /label=LEU2 Replication_ori 8186..8703 /label=ARS/CEN CDS 8835..9695 /label=AmpR Replication_ori 9840..10522 /label=pUC	5
LOCUS pEXP22.2 9587 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Promoter 272..1701 /label=pADH1 Riboswitch 1724..1970 /label=tc3 CDS 1981..4626 /label=GAL4 Terminator 4860..5017 /label=tADH1 Replication_ori 5344..5799 /label=f1 CDS complement(5902..6576) /label=TRP1 Replication_ori 7112..7629 /label=ARS/CEN CDS 7761..8621 /label=AmpR Replication_ori 8766..9448 /label=pUC	6
LOCUS pRS414-empty 4408 bp DNA circular UNA FEATURES Location/Qualifiers Replication_ori 123..1063	7

/label=PMB1 CDS complement(1064..1924) /label=AmpR Replication_ori complement(2056..2573) /label=ARS/CEN CDS 3106..3780 /label=TRP1	
LOCUS pRS415-empty 5641 bp DNA circular UNA FEATURES Location/Qualifiers Replication_ori 152..1092 /label=PMB1 /note="origin of replication" Misc._feature 2095..2211 /label=CEN6 /note="centromere VI" Misc._feature 2212..2586 /label=ARSH4 CDS complement(3330..4424) /label=LEU2 Replication_ori complement(5141..5597) /label=F1_ORI	8
LOCUS pEXP22.3 5924 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Promoter 272..1717 /label=pADH1 CDS 1731..1754 /label=NLS CDS 1761..2000 /label="VP16 AD" CDS 2055..2351 /label="RalGDSwt" Terminator 2630..2787 /label=tADH1 Replication_ori 3449..3966 /label=ARS/CEN CDS 4098..4958 /label=AmpR Replication_ori 5103..5785 /label=pUC	9
LOCUS pEXP32.4 10408 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Terminator complement(272..521) /label="URA3 terminator" CDS complement(522..1325) /label=URA3 Promoter complement(1326..1649) /label=SPAL10(MaV203) URS complement(1413..1421) /label=URS1 UAS 1510..1526 /label=UAS1 UAS 1529..1545 /label=UAS2 UAS 1548..1564 /label=UAS3 UAS 1567..1583 /label=UAS4 UAS 1586..1602 /label=UAS5 Promoter 1650..3104 /label=ADH1p CDS 3128..3571 /label="GAL4 DBD" CDS 3629..4180	10

/label=Krev1 Terminator 4462..4619 /label=ADH1t Replication_ori 4946..5401 /label=f1 CDS 6110..7204 /label=LEU2 Replication_ori 7933..8450 /label=ARS/CEN /label=P13 CDS 8582..9442 /label=AmpR Replication_ori 9587..10269 /label=pUC	
LOCUS pEXP22.1CUP 6735 bp DNA circular DNA circular DEFINITION FEATURES Location/Qualifiers Promoter 257..708 /label=pCUP1 /promoter_type="eukaryotic" /vntifkey="29" Riboswitch 741..987 /label=tc3 CDS 1001..1021 /label=NLS CDS 1028..1372 /label="GAL4 AD" CDS 1433..1729 /label=RalGDSwt Terminator 2008..2165 /label=tADH1 Replication_ori 2492..2947 /label=f1 CDS complement(3050..3724) /label=TRP1 Replication_ori 4260..4777 /label=ARS/CEN CDS 4909..5769 /label=AmpR Replication_ori 5914..6596 /label=pUC	11
LOCUS pEXP32.4CUP 9409 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Terminator complement(272..521) /label="URA3 terminator" CDS complement(522..1325) /label=URA3 Promoter complement(1326..1649) /label=SPAL10(MaV203) URS complement(1413..1421) /label=URS1 Promoter 1654..2105 /label=pCUP1 CDS 2129..2572 /label="GAL4 DBD" CDS 2630..3181 /label=Krev1 Terminator 3463..3620 /label=ADH1t Replication_ori 3947..4402 /label=f1 CDS 5111..6205 /label=LEU2 Replication_ori 6934..7451 /label=ARS/CEN	12

CDS 7583..8443 /label=AmpR Replication_ori 8588..9270 /label=pUC	
LOCUS AD_part 2955 bp DNA linear UNA	13
LOCUS DBD-reporter_part 4426 bp DNA linear UNA	14
LOCUS pRP2-URA3 8494 bp DNA circular UNA FEATURES Location/Qualifiers CDS complement(663..1766) /label=LEU2 Replication_ori complement(2606..2912) /label=f1 Terminator complement(3012..3332) CDS complement(3334..4137) /label=URA3 Riboswitch complement(4148..4394) /label=Tc3 Promoter complement(4404..4855) /label=CUP1 Terminator 4901..5090 /label="CYC1 terminator" Replication_ori 5277..5944 /label=pUC CDS complement(6095..6952) /label=AmpR Replication_ori 7086..8241 /vntifkey="33" /label="2 micron ori"	15
LOCUS pRP4-URA3 7666 bp DNA circular UNA FEATURES Location/Qualifiers Terminator complement(82..402) CDS complement(404..1207) /label=URA3 Riboswitch complement(1218..1464) /vntifkey="21" /label=Tc3 Promoter complement(1474..1925) /label=CUP1 Terminator 1971..2160 /vntifkey="43" /label="CYC1 terminator" replication_ori 2347..3014 /label=pUC CDS complement(3165..4022) /vntifkey="4" /label=AmpR replication_ori 4154..4671 /label=ARS/CEN CDS complement(5399..6502) /label=LEU2 Replication_ori complement(7342..7648) /label="f1 ori"	16
LOCUS pCUP1-tc3-URA3_integrated 1844 bp DNA linear UNA DEFINITION FEATURES Location/Qualifiers Promoter 1..452 /label=pCUP1 Riboswitch 462..708 /label="tc3 riboswitch" CDS 719..1522 /label=URA3 Terminator 1524..1844	17
LOCUS pEXP22RALGDSm2 7468 bp DNA circular UNA DEFINITION FEATURES Location/Qualifiers Promoter 272..1726	18

/label=ADH1p CDS 1731..1754 /label=NLS CDS 1761..2105 /label="GAL4 AD" CDS 2166..2462 /label=RalGDSm2 Terminator 2741..2898 /label=ADH1t Replication_ori 3225..3680 /label=f1 CDS complement(3783..4457) /label=TRP1 replication_ori 4993..5510 /label=ARS/CEN CDS 5642..6502 /label=AmpR replication_ori 6647..7329 /label=pUC	
LOCUS pEXP32.5XX 9992 bp DNA circular DEFINITION FEATURES Location/Qualifiers Terminator complement(272..521) /label="URA3 terminator" CDS complement(522..1244) /label=sfGFP CDS complement(1245..1967) /label=GFP CDS complement(1968..2681) /label=yeGFP CDS complement(2685..3404) /label=sfGFP CDS complement(3405..4127) /label=GFP CDS complement(4128..4841) /label=yeGFP Promoter complement(4842..5165) /label=SPAL(MaV203) URS complement(4929..4937) /label=URS1 Replication_ori 5591..5897 /label="f1 origin" CDS complement(6292..6948) /label=HIS3 Replication_ori 7517..8034 /label=ARS/CEN CDS 8166..9026 /label=AmpR Replication_ori 9171..9853 /label=pUC	19
LOCUS pEXP42.2 8448 bp DNA circular DEFINITION FEATURES Location/Qualifiers Promoter 272..1701 /label=pADH1 Riboswitch 1724..1970 /label=tc3 CDS 1981..2004 /label=NLS CDS 2005..3312 /label=GAL80 Terminator 3546..3703 /label=tADH1 Replication_ori 4030..4485 /label=f1 CDS complement(4748..5404) /label=HIS3	20

Replication_ori	6488..6490 /label=ARS/CEN	
CDS	6622..7482 /label=AmpR	
Replication_ori	7627..8309 /label=pUC	

EXAMPLE 2

This Example describes how buffer devices based on AD/DBD constructs reduced the noise, *i.e.*, increased the difference in survival between true and false positive cells, in reporter gene-regulation by a ligand-responsive riboswitch.

The buffer devices were applied to a riboswitch to down-regulate the classic yeast selection/reporter gene *URA3* for which population-wide down-regulation can be assayed by the ability of the cell population to form colonies in presence of the counter selection agent 5'-fluoroorotic acid (5-FOA). Tetracycline-dependent colony formation in *Saccharomyces cerevisiae* was obtained by introducing various combinations of genetic constructs in the parent strain MaV203, in which the wild-type regulatory genes *GAL4* and *GAL80* had previously been deleted, similar to the strain MaV103 (Vidal *et al.*, 1996; MacDonald, 2001). The functional buffer strains were constructed genetically by several methods, all enabling the riboswitch to down-regulate the reporter *URA3* robustly to allow a high degree of ligand-dependent colony formation unlike conventional application of the riboswitch to control the same reporter gene. Various control constructs were also tested.

The general method for obtaining the gene regulation that lead to small molecule-dependent colony formation is described below.

Protocol (master)

The following protocol was used for evaluating small molecule-dependent colony formation:

4 mL of synthetic complete (SC) medium lacking appropriate nutrients for auxotrophic plasmid maintenance (*e.g.* lacking leucine and tryptophan when culturing the strains in Table 5) was inoculated with a single colony of the yeast host strain. The culture was split into two halves and 250 µM tetracycline (as riboswitch-binding small molecule) was added to one from a 20 mM stock solution. Both cultures were cultured for > 18 hours at 30 degrees Celsius with 250 rpm horizontal shaking. Ten-fold serial dilutions were made for both culture conditions and 5 µL of each dilution was spotted to SC plates (2% agar) lacking the appropriate nutrients for plasmid maintenance and to 'selection plates', which were SC plates supplemented with 1.0 % (w/v) 5-FOA with and without the relevant concentration of tetracycline. Note: Some batch-to-batch variation can occur with 5-FOA, and testing the 5-FOA batch at +/- 40% of the concentration is recommended. The spot tests on plates without 5'-FOA were used to verify the number of colony-forming cells in each dilution.

Results

The results are shown in Figure 3. For each type of genetic construct, the colony formation response, reflecting noise-reduction capacity, is shown, along with a representation of the relation between AD and DBD molecules in the cell. A high noise-reduction capacity can be identified as a high or full improvement of colony formation in presence of the effector compound compared to the same strain tested in absence of the effector compound, whereas low/no noise-reduction capacity can be identified by no or low improvement of colony formation by presence of the effector compound, compared to use of the biosensor without using the system of the invention, i.e., before placing it in the context of a buffer device.

The strain family represented by PRa4, PRa22, PRa34 and PRa41, all based on a DBD and riboswitch-controlled translation of AD, displayed buffering functionality by requiring the addition of 50 µM tetracycline to form colonies on the selection plates. The buffering effect was not dependent on the particular AD, plasmids or reporter gene propagation method. In PRa4, the effector platform was integrated within the chromosome, while the AD and DBD hybrid proteins were expressed from plasmids. In PRa41, the GAL4 AD was replaced by the AD of *Herpes simplex* VP16, with the buffer function retained. To construct PRa34, we integrated the essential components (defined as 'AD part' and 'DBD + reporter part') of the buffer device into the chromosome of a yeast strain. PRa34 retained the observed tetracyclin-dependent ability to form colonies, showing that the buffering effect was independent of the particular plasmids pEXP32.4 and pEXP22.1. The nucleotide sequences of 'AD part' and 'DBD + reporter part' for PRa34 are shown in SEQ ID NOS: 13 and 14, respectively. Finally, in strains PRa22, PRa46, PRa47, the *URA3* gene and its GAL4-regulated promoter pSPAL10 (Vidal *et al.*, 1996) were deleted from the chromosome of MaV203 and introduced on an ARS/CEN plasmid. The buffering function was retained irrespective of the reporter gene propagation method.

Various control strains were studied. In the control strains PRd3, PRd4, and PRd5, the reporter gene was regulated directly by the riboswitch, *i.e.* the riboswitch was fused directly to the coding sequence. The observed lack of colony formation control by the riboswitch ligand demonstrated that this method was insufficient. For PRd3, PRd4 and PRd5, the plate concentration of tetracycline was also tested at 500 µM tetracycline, without changing the results shown in Figure 3. Since expression noise can arise from the method by which a gene is propagated in the cell, we tested propagation by respectively 2-micron plasmid (PRd3), ARS/CEN plasmid (PRd4) or chromosomal integration (PRd5). As seen in the control strain PRa24, only poor control of colony formation was obtained when regulating the reporter gene mRNA directly and activating the reporter gene transcription with interacting DBD and AD molecules, as shown by Figure 3A. In order to better observe any possible regulation of colony formation, the 5-FOA concentration for this strain was elevated to 0.15 % (w/v) 5-FOA, since the insertion of the riboswitch in the 5-untranslated region of the reporter gene, itself lowered the reporter gene expression too much for 0.1 % (w/v) 5-FOA to be toxic for the cells even in absence of riboswitch regulation. Since this did not result in good signal

regulation, it can be concluded that the buffering effect obtained with strains PRa4/PRa22/PRa34/PRa41 was not merely a general consequence of expressing the reporter gene using a DBD that interacts with DNA upstream activating sequences and an AD, or a general consequence of expressing the reporter gene near this expression level. In PRa28, the riboswitch controlled the expression of the wildtype GAL4 transcriptional activator, while the effector platform was integrated chromosomally. The lack of colony formation control confirmed that the regulation efficiency of the invention was not a direct consequence of the binding behaviour of GAL4, which is shared by both GAL4 (PRa28) and GAL4 DBD in the buffer device. Based on the behaviour of PRa28, a strain in which the riboswitch controls the expression of the GAL4 DBD rather than GAL4 AD should also show no control of colony formation, since the DBD molecules would be guided to the DNA-binding sites by the same high affinity as full-length GAL4. In PRa56, the riboswitch controlled the expression of the wildtype GAL4 transcriptional activator, while the GAL4 DBD was expressed constitutively from the same promoter. The lack of control of colony formation showed that competition for binding to the response elements between a functional (wildtype GAL4) and non-functional (GAL4 DBD) transcription factor could not account for the obtained effect in the invention. Finally, PRa21 was a control strain showing that when the interaction between the AD and DBD hybrid proteins was abolished by mutation, the reporter gene was not activated.

EXAMPLE 3

This Example describes how the detection limit for an effector compound was increased.

To increase the detection level for the small molecule, the promoter of DBD (P2) was exchanged for the weaker promoter of CUP1, which can be induced by copper ions, but was not in this case (PRa47) (see Tables 4 and 5). Therefore, the expression ratio of AD to DBD in the cell increases (provided no Cu^{2+} is added beyond the medium supplement). Using the protocol of Example 2, it was found that this lead to a change of tetracycline detection level from 50 µM tetracycline in the plates (with PRa45 and PRa46) to 150 µM tetracycline in the plates (with PRa47) before the detection system allowed the plated cells to form colonies (Figure 4).

EXAMPLE 4

This is an exemplary protocol for small molecule dependent fluorescence (see Example 5).

Protocol

4 mL of synthetic complete (SC) medium lacking appropriate nutrients for auxotrophic plasmid maintenance (*i.e.* lacking leucine and tryptophan) is inoculated with a single colony of the yeast host strain in which a nucleotide sequence encoding a fluorescent protein is under control of a promoter including a response element. The culture is split into two equal portions and 250 µM tetracycline (as riboswitch-binding small molecule) are added to one

from a 20 mM stock solution. Both cultures are cultured for > 18 hours at 30 degrees Celsius with 250 rpm horizontal shaking. The cultures are analyzed for fluorescence at single cell level on a flow cytometer to allow distinction of whether each cell has a fluorescence level that is different depending on whether the cell is cultured in presence of the small molecule effector compound.

EXAMPLE 5

This Example shows how a fluorescent reporter gene was applied to verify the observed ability of the system to modify the sensitivity for the effector molecule.

Protocol

Pre-cultures of the strains PRa74, PRa78 (AD_1) and PRa79 (AD_2) were inoculated from a single colony of the strain in SC medium (2 % glucose, pH = 5.6) lacking leu, trp and histidine. Following 18 hours of cultivation at 30 degrees C, 175 rpm horizontal shaking, 200 µL microtiter main cultures were inoculated from these in 75 % SC medium (diluted with milliQ water and back-standardized to 2 % glucose) with the relevant concentrations of tetracycline added. The cultures were sealed with a gas-permeable Breathseal (Greiner bio-one) and plastic lid and were cultured in a horizontal shaker (Innova) at 30 degrees C, 300 rpm shaking. Following 16 hours of cultivation, the cultures were measured by flow cytometry on a BD LSRFortessa Cell Analyzer using a FITC filter with collection limit set to 10,000 cells. The mean FITC intensity for each sample was reported. The measurements from the GFP-devoid PRa74 strain were used for background-subtraction.

Results

The results (Figure 5) showed how the output of the system changed, when the sensitivity was modified using the system. In strain PRa78 (marked as AD1), the AD and DBD hybrid molecules were expressed with same transcriptional strength, corresponding to strain PRa22 and PRa46. In another strain PRa79 (marked as AD2), the AD and DBD hybrid molecules were expressed respectively with the promoter of ADH1 and CUP1, to produce a strain that only permits colony formation at an increased concentration of the effector molecule, corresponding to strain PRa47. In both strains, an additional construct was introduced featuring a translational fusion of six GFPs under control of the same promoter and response elements as the *URA3* reporter gene. As seen from Figure 5, shifting the transcriptional ratio of the AD to DBD hybrid molecules quantitatively produced a shifting of the output signal. Note: While we chose to perturb the ratios of AD to DBD at the transcriptional level, the effect arises at the protein level, and thus perturbations that directly affect the protein level ratios would affect the output similarly.

EXAMPLE 6

This Example shows how the signal of the biosensor could be inverted to give the opposite direction of switching without changing the biosensor or the reporter gene. The GAL80 masking-repressor protein was expressed under control of the riboswitch-based biosensor tc3, while the GAL4 AD and DBD hybrid molecules were expressed constitutively. In principle, they could also be expressed as a fusion or simply as the wildtype GAL4. Using the same URA3 reporter gene and promoter with response elements, the signal from the riboswitch was inverted from OFF to ON. As used herein, masking-repressors function by binding to a transcriptional activation domain to hereby reduce the activation enough to convey a detectable signal with the particular reporter gene used. This implies that the masking-repressor and activation domain should interact for use in the context of the invention.

The protocol followed exactly the master protocol of Example 2.

Results

Strain PRa84 was constructed with the biosensor regulating the GAL80 protein from *S. cerevisiae* expressed from an ADH1 promoter. Meanwhile, the first and second hybrid proteins were expressed also from ADH1 promoters. As seen in FIG. 3B, colony formation was now only permitted in absence of the riboswitch ligand when plated on medium containing 0.1 % (w/v) 5-FOA, while addition of ligand caused the cells to strongly reduce the colony formation ability to less than 1 of 10^4 plated cells.

LIST OF REFERENCES

Ames *et al*. (2010). Chemistry & biology 17(7), 681–5.
Bae *et al.* (2003). Nature biotechnology 21, 275–80.
Baker *et al.* (2002). PNAS 99, 16357-16542.
Becskei A and Serrano L (2000). Nature 405:590–3.
Berger et al. (1990), Cell 61(7), 1199–208.
Dambach et al., Curr Opin Microbiol 2009;12:161-169
Desai and Gallivan (2004). J Am Chem Soc 126, 13247-54.
Dove and Hochschild (2004). Methods in Molecular Biology 261, 231–246.
Fields and Song (1989). Nature 340, 245–246.
Galvão and de Lorenzo (2006). Current opinion in biotechnology 17, 34–42.
Gorska *et al.* (2006). The Journal of biological chemistry 281, 14429–39.
Harvey *et al.*, (2002). RNA 8, 452–463.
Hope and Struhl (1986). Cell 46, 885–94.
Joung *et al.* (2000). PNAS USA 97, 7382–7.
Kötter *et al.* (2009). Nucleic acids research 37, e120.
Macdonald (2001). In: Methods in Molecular Biology. Humana Press, New Jersey.

Mcadams AND Arkin (1999). Trends in genetics : TIG 15, 65–9.
Michener and Smolke (2012). Metabolic Engineering 14:306–316.
Michener *et al.* (2012). Metabolic engineering 14:212–22.
Moll *et al.* (2001). Protein Science 649–655.
Müller *et al.* (2006). Nucleic acids research 34, 2607–17.
Martin et al., Methods 2012;58:367-375.
Muranaka *et al.*, (2009). Nucleic acids research 37, e39.
Nahvi *et al.* (2002). Chemistry & biology 9, 1043.
Ogawa and Maeda (2008). Chembiochem 9, 206-9.
Park et al., Sensors 2013;13:5777-5795.
Pawlowski *et al.* (2003). Journal of Bacteriology 185, 5240–5247.
Potter *et al.* (2010). Cell 141, 536–48.
Serebriiskii *et al.* (1999). The Journal of biological chemistry 274, 17080–7.
Serganov, A., Nudler, E., 2013. A decade of riboswitches. Cell 152, 17–24.
Sint Fiet *et al.* (2006). PNAS 103, 1693-8.
Stoltenburg *et al.* (2007). Biomolecular engineering 24:381–403
Stynen *et al.* (2012). Microbiology and molecular biology reviews 76, 331–82.
Vidal *et al.* (1996). PNAS USA 93, 10315–20.
Weigand *et al.* (2008). RNA (New York, N.Y.) 14, 89–97.
Wittman and Seuss (2012). FEBS Letters 586:2076-2083.
Xu *et al.* (1997). PNAS 94, 12473–12478.
Yang *et al.* (2012). Nature Communications 4:1413.
WO2004/113495
US20100197006
WO2001/119956
WO2011/088076
WO2004/027035
WO2007/100142

CLAIMS

1. An effector compound detection system comprising

(a) a first DNA molecule comprising a sequence encoding an RNA molecule comprising a riboswitch responsive to an effector compound, the riboswitch being operably linked to a coding region encoding a first hybrid protein comprising an activation domain (AD) and a first interaction domain, wherein binding of the effector compound to the riboswitch regulates expression of the first hybrid protein, and

(b) a second DNA molecule comprising a sequence encoding a second hybrid protein comprising (i) a second interaction domain which binds to the first interaction domain and (ii) a DNA-binding domain (DBD) which binds a response element,

wherein the simultaneous binding of the second hybrid protein to the first hybrid protein and the response element triggers transcription of a reporter gene operably linked to the response element.

2. The effector compound detection system of claim 1, wherein the binding of the effector compound to the riboswitch decreases expression of the first hybrid protein.

3. The effector compound detection system of any one of the preceding claims, wherein each of the first and second DNA molecules is operably linked to a first and second promoter, respectively, optionally wherein the strength of the first promoter is the same as or higher than that of the second promoter.

4. The effector compound detection system of any one of the preceding claims, comprising a third DNA molecule encoding a third protein, wherein the third protein comprises the second interaction domain but does not bind the response element.

5. A host cell comprising the effector compound detection system of any one of the preceding claims and a reporter gene operably linked to the response element.

6. The host cell of claim 5, which is a microbial cell or a mammalian cell.

7. The host cell of any one of claims 5 and 6, wherein

(a) the host cell does not comprise an endogenous protein which binds to the first interaction domain, the second interaction domain or a binding site for the DBD in the response element;

(b) in the absence of effector compound, the ratio of the expression level of the second interaction domain over the first interaction domain is in the range of about 0.03 to about 40 when the riboswitch is one where binding of the effector compound to the riboswitch decreases expression of the first hybrid protein;

(c) the response element comprises at least 2 binding sites for the DBD;

(d) the ratio of the number of molecules comprising the second interaction domain to the number of binding sites for the DBD in the response element is at least 2; or

(e) a combination of (a) and (b); or

(f) a combination of any of (a) to (d).

8. The host cell of any one of claims 5 to 7, wherein the change in expression of the reporter gene when the effector compound binds the riboswitch permits proliferation of the host cell under a predetermined condition under which the host cell would otherwise proliferate at at least a 100-fold lower frequency.

9. The host cell of any one of claims 5 to 8, wherein the reporter gene encodes a fluorescent, chromogenic or luminescent protein.

10. An expression library of host cells according to any one of claims 5 to 9, the expression library comprising a plurality of expression constructs, each host cell comprising a vector encoding at least one potential modulator compound.

11. The expression library claim 10, wherein the modulator compound is

(a) an enzyme of a biosynthetic pathway effecting a chemical transformation of a substrate into the effector compound in the host cell; or

(b) a protein transporting the effector compound into or out of the host cell.

12. A kit comprising one or more plasmids comprising the effector compound detection system of any one of claims 1 to 4, optionally further comprising a fourth DNA molecule comprising a reporter gene operably linked to the response element.

13. A method for identifying a host cell producing an effector compound of interest, comprising

providing a plurality of host cells, each comprising

(i) a first DNA molecule comprising a sequence encoding a first hybrid protein comprising a first interaction domain and an activation domain (AD), wherein the expression of the first DNA molecule in the host cell is under the control of a riboswitch or a transcription factor regulated by an effector compound of interest;

(ii) a second DNA molecule comprising a sequence encoding a second hybrid protein comprising a second interaction domain which binds to the first interaction domain and a DNA-binding domain (DBD) which binds a response element; and

(iii) a reporter gene operably linked to the response element, wherein the simultaneous binding of the second hybrid protein to the first hybrid protein and the response element triggers transcription of the reporter gene, and modulation of the expression of the reporter gene provides for a detectable signal; and

selecting at least one host cell where the expression of the reporter gene is modulated.

14. The method of claim 13, wherein each host cell is a host cell according to any one of claims 5 to 9.

15. The method of claim 13, wherein the plurality of host cells is an expression library of host cells according to any one of claims 10 to 11.

16. A method of producing an improved effector compound detection system from a first effector compound detection system, said first effector compound detection system comprising a DNA sequence encoding an RNA molecule comprising the transcript of a reporter gene operably linked to a riboswitch responsive to an effector compound, the method comprising the steps of:

(a) preparing a first DNA molecule comprising a sequence encoding an RNA molecule comprising the riboswitch operably linked to a coding region encoding a first hybrid protein comprising an AD and a first interaction domain, wherein binding of the effector compound to the riboswitch regulates expression of the first hybrid protein;

(b) preparing a second DNA molecule comprising a sequence encoding a second hybrid protein comprising (i) a second interaction domain which binds to the first interaction domain and (ii) a DBD which binds a response element;

thereby obtaining an improved effector compound detection system.

17. The method of claim 16, wherein the binding of the effector compound to the riboswitch decreases expression of the first hybrid protein, optionally wherein the modified effector compound detection system further comprises the features of any one of claims 3 to 15.

18. The method of claim 17, wherein the first and second DNA molecules comprise regulatory elements which, in the absence of the effector compound, provide for a higher expression level for the first than the second hybrid protein in a host cell, thereby obtaining a modified effector compound detection system having a reduced sensitivity for the effector compound.

19. The method of claim 17, wherein the first and second DNA molecules comprise regulatory elements which, in the absence of the effector compound, provide for a lower expression level for the first than the second hybrid protein in a host cell, thereby obtaining a modified effector compound detection system having an increased sensitivity for the effector compound.

Fig. 1A 1/7

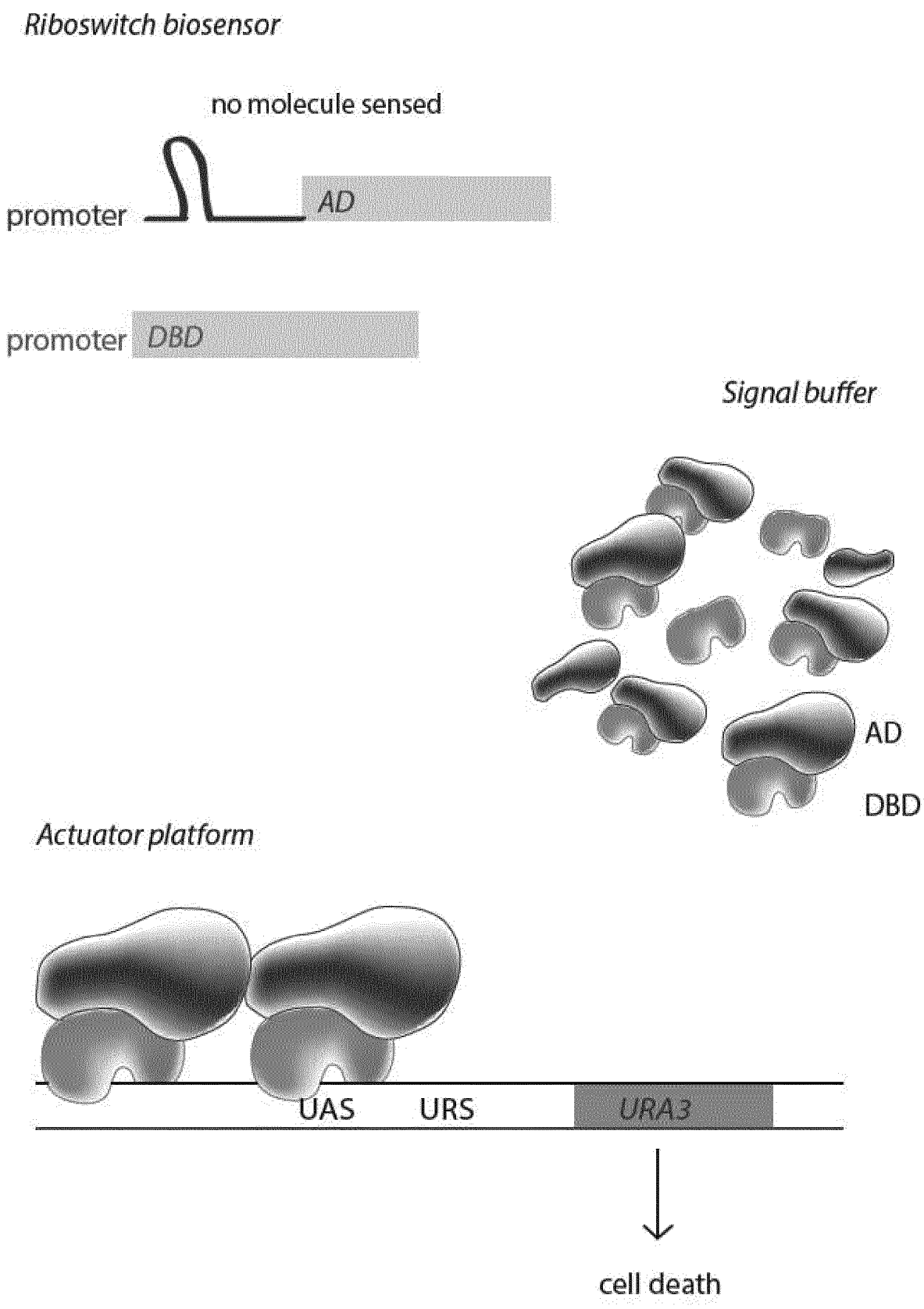

Riboswitch biosensor
no molecule sensed
promoter
AD
promoter
DBD
Signal buffer
AD
DBD
Actuator platform
UAS
URS
URA3
cell death

Fig. 1B

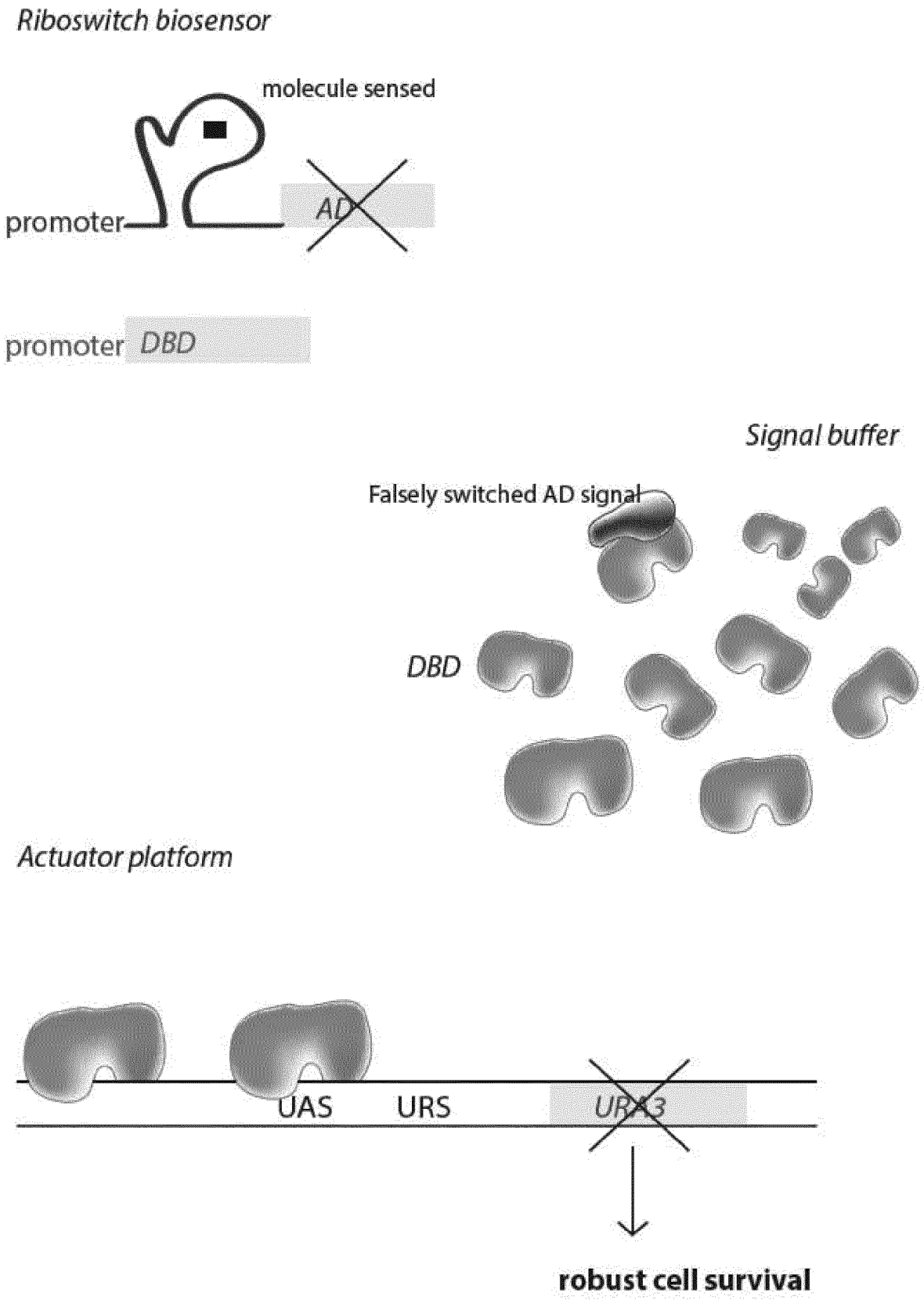
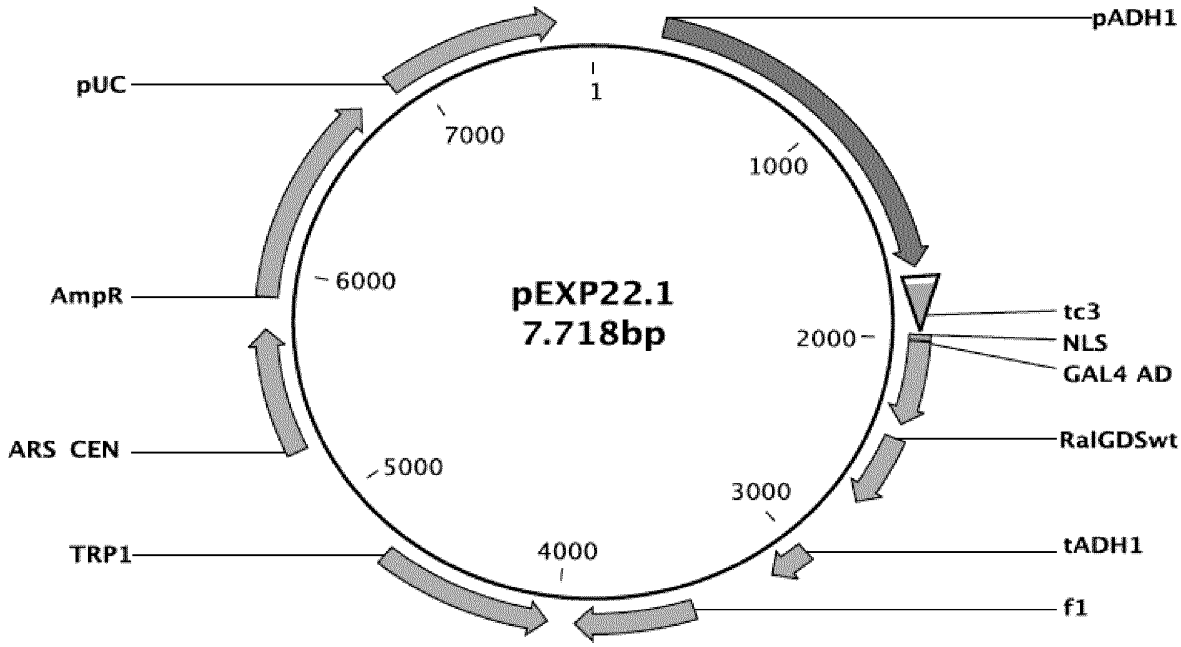

Riboswitch biosensor
molecule sensed
promoter
AD
promoter
DBD
Signal buffer
Falsely switched AD signal
DBD
Actuator platform
UAS
URS
URA3
robust cell survival

Fig. 2

A


pADH1
pUC
1
7000
1000
6000
pEXP22.1
7.718bp
AmpR
2000
tc3
NLS
GAL4 AD
RalGDSwt
ARS CEN
5000
3000
TRP1
4000
tADH1
f1

B

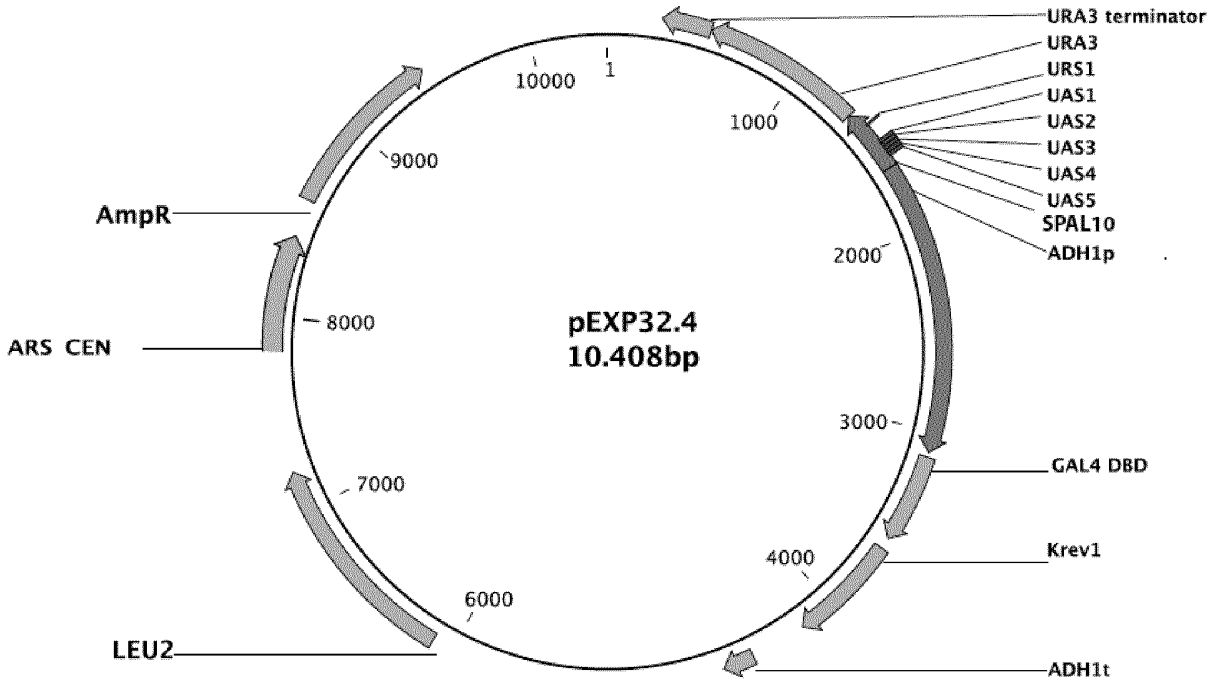

URA3 terminator
URA3
URS1
UAS1
UAS2
UAS3
UAS4
UAS5
SPAL10
ADH1p
10000
1
1000
9000
AmpR
2000
8000
pEXP32.4
10.408bp
ARS CEN
3000
GAL4 DBD
7000
4000
Krev1
6000
LEU2
ADH1t

Fig. 3A

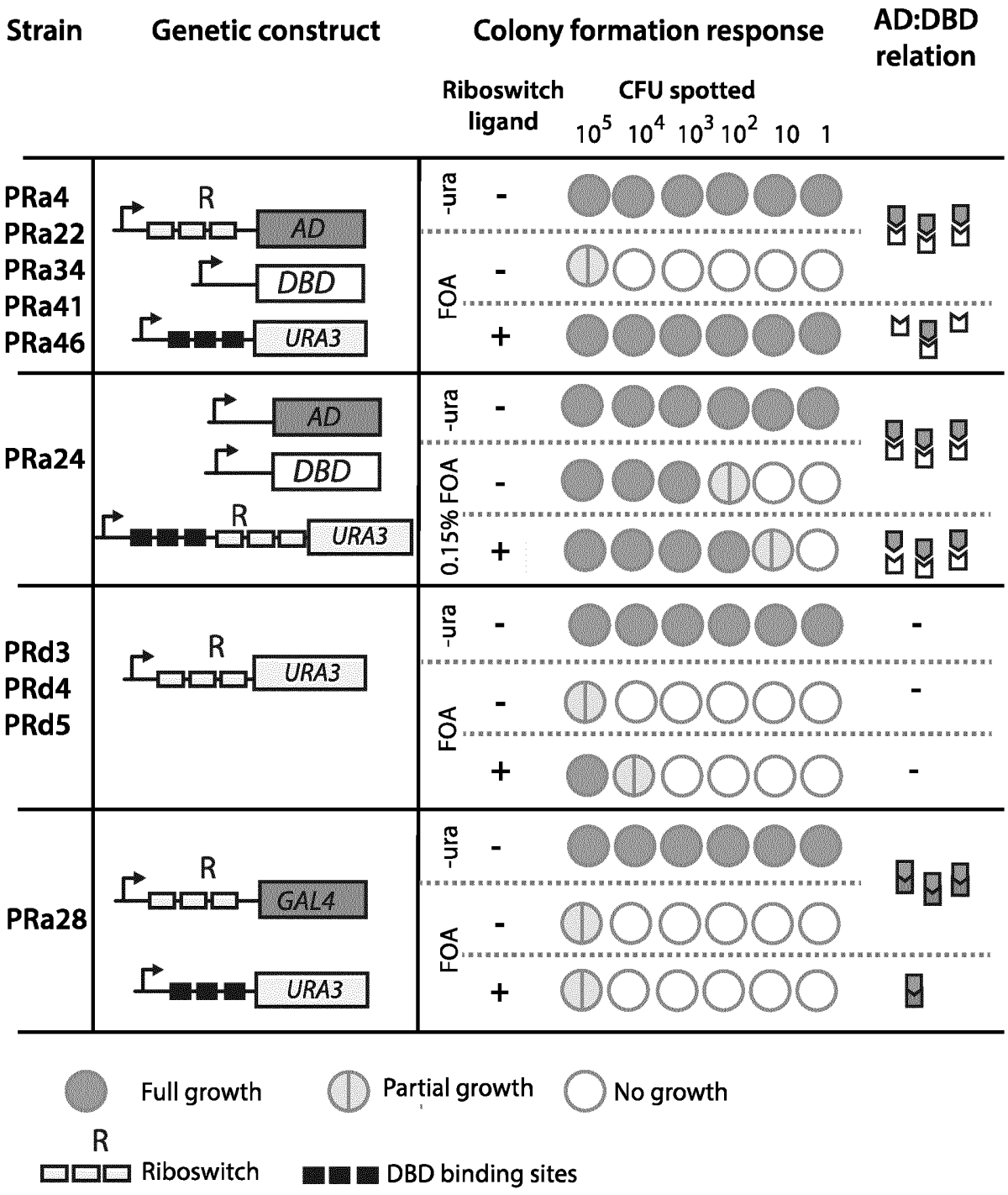

Strain
Genetic construct
Colony formation response
AD:DBD relation
Riboswitch ligand
CFU spotted
10^5 10^4 10^3 10^2 10 1
PRa4
PRa22
PRa34
PRa41
PRa46
R
AD
DBD
URA3
-ura
FOA
-
-
+
PRa24
AD
DBD
R
URA3
-ura
0.15% FOA
-
-
+
PRd3
PRd4
PRd5
R
URA3
-ura
FOA
-
-
+
PRa28
R
GAL4
URA3
-ura
FOA
-
-
+
Full growth
Partial growth
No growth
R
Riboswitch
DBD binding sites

Fig. 3B

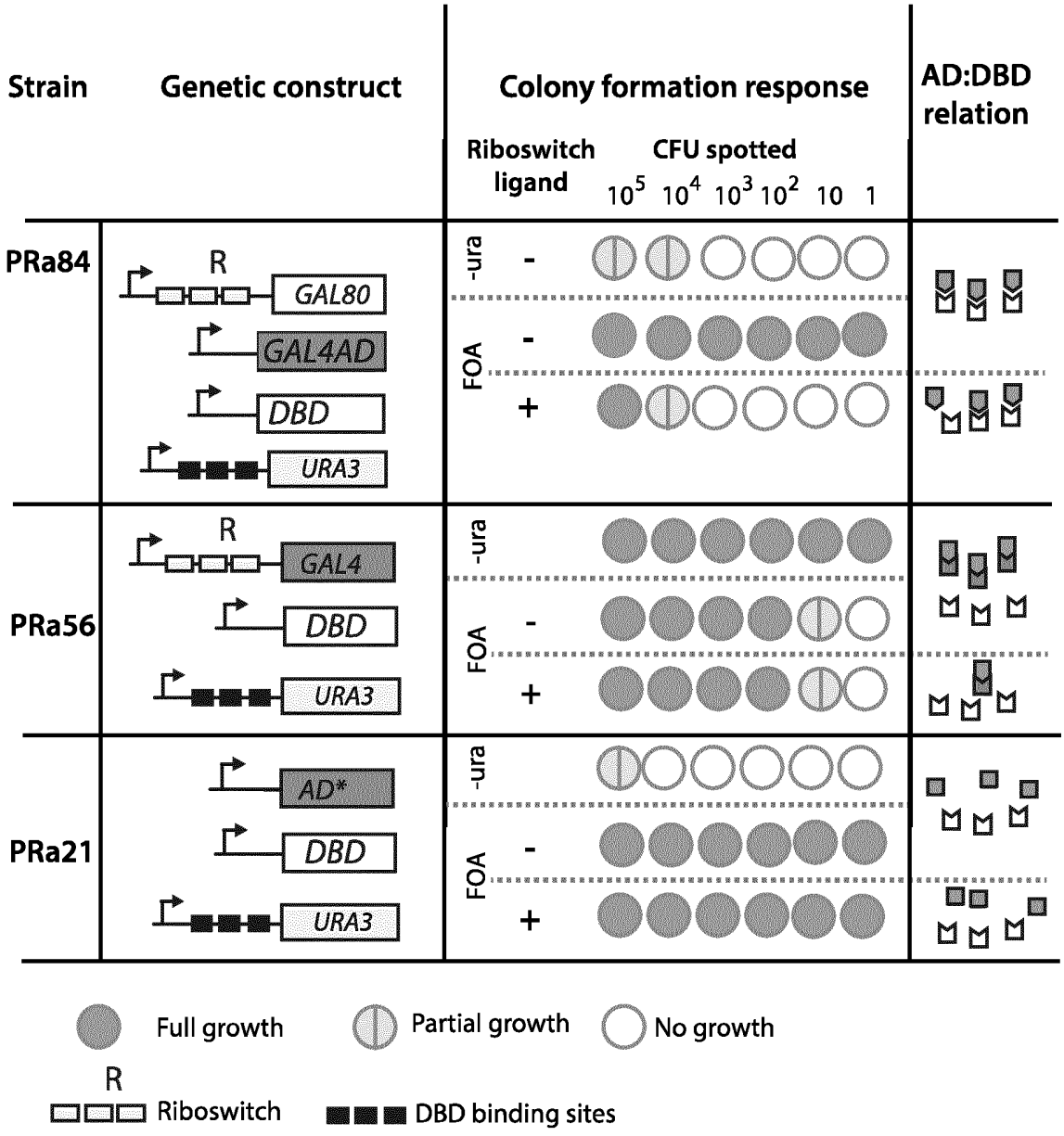

Strain
Genetic construct
Colony formation response
Riboswitch ligand
CFU spotted
10^5 10^4 10^3 10^2 10 1
AD:DBD relation
PRa84
R
GAL80
GAL4AD
DBD
URA3
-ura
FOA
-
-
+
PRa56
R
GAL4
DBD
URA3
-ura
FOA
-
+
PRa21
AD*
DBD
URA3
-ura
FOA
-
+
Full growth
Partial growth
No growth
R
Riboswitch
DBD binding sites

Fig. 4

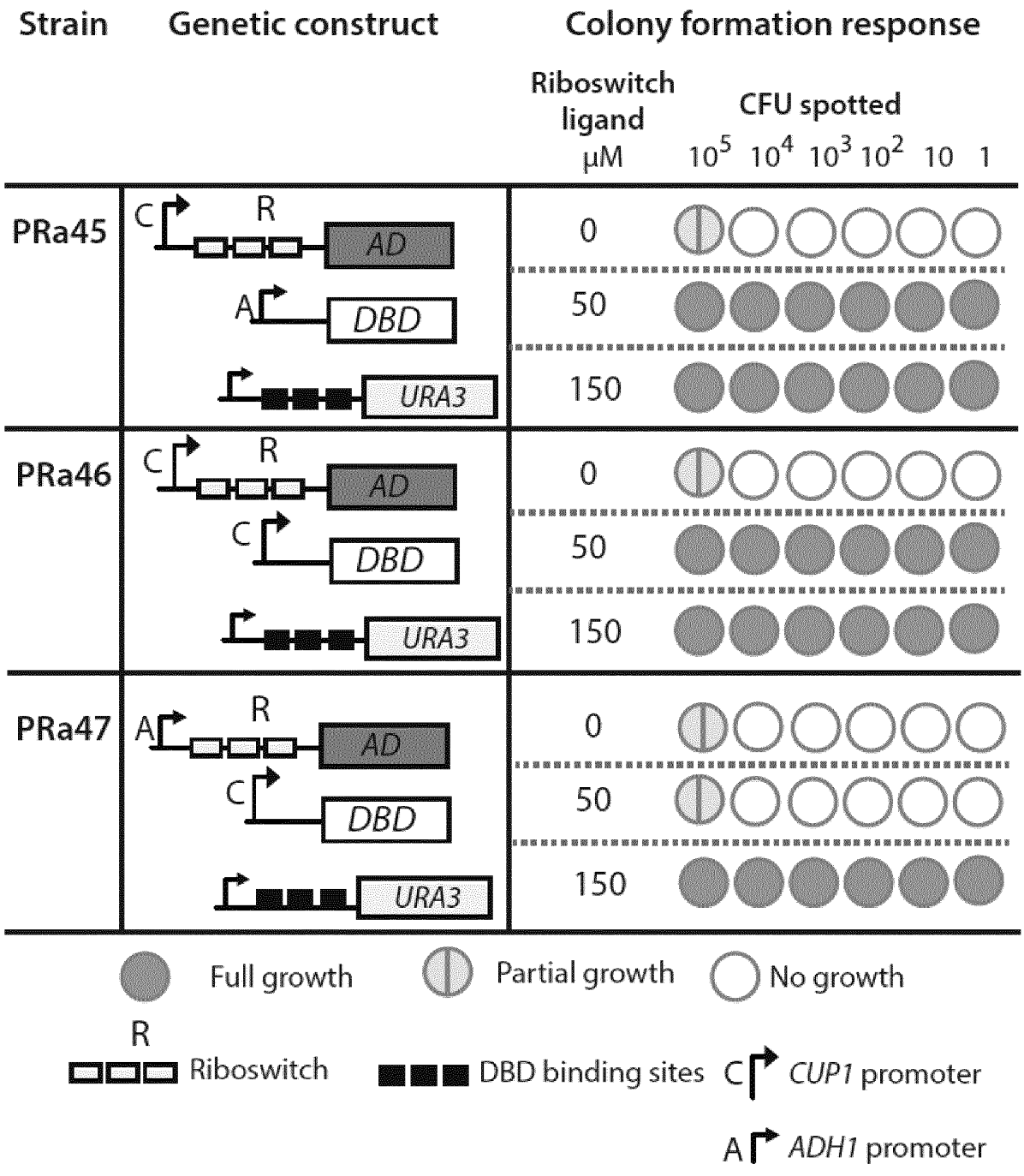
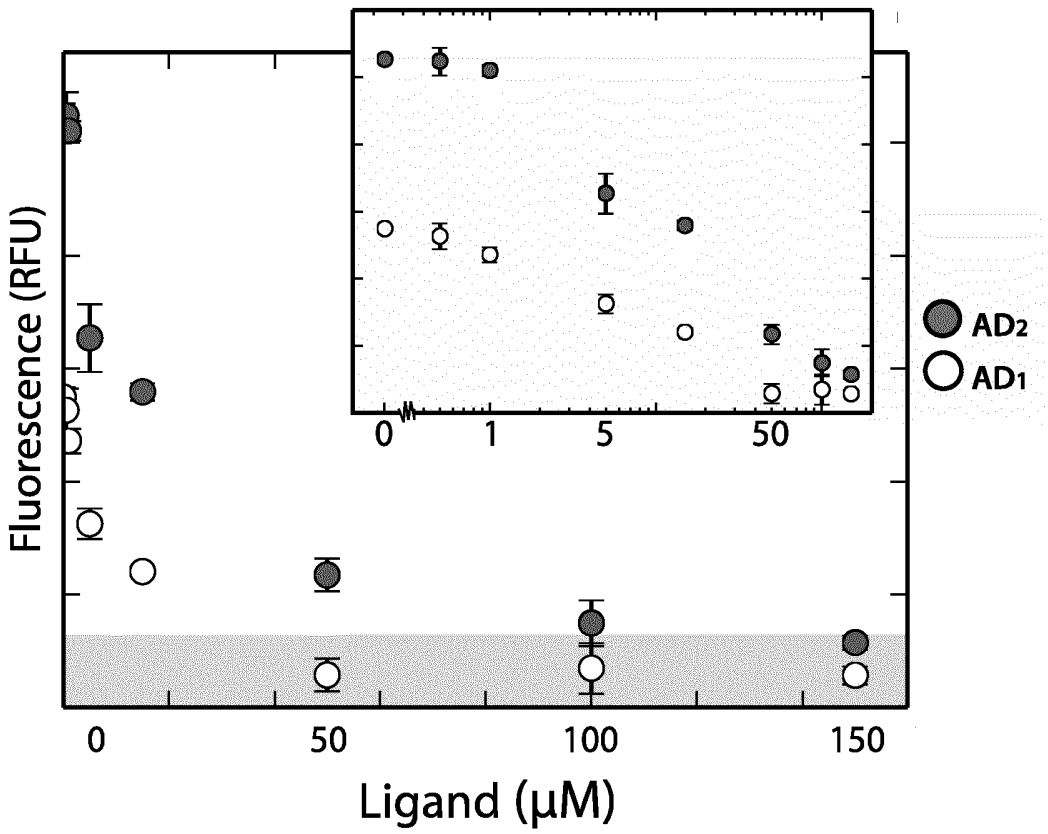

Strain
Genetic construct
Colony formation response
Riboswitch ligand µM
CFU spotted
10^5 10^4 10^3 10^2 10 1
PRa45
PRa46
PRa47
C
R
A
AD
DBD
URA3
0
50
150
Full growth
Partial growth
No growth
Riboswitch
DBD binding sites
CUP1 promoter
ADH1 promoter

Fig. 5



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/070954

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/25 C12Q1/68 C12N15/11
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2011/119956 A1 (INTEGRATECH PROTEOMICS LLC [US]; SANDROCK TANYA [US]) 29 September 2011 (2011-09-29) see claims 1-2, 14, 30,32-35 and page 17-27; the whole document -----	1-19
Y	WO 2011/088076 A2 (UNIV YALE [US]; BREAKER RONALD R [US]; WEINBERG ZASHA [US]) 21 July 2011 (2011-07-21) cited in the application in particular see claim 14,16-17; the whole document -----	1-19
Y	WO 2004/027035 A2 (UNIV YALE [US]) 1 April 2004 (2004-04-01) see claim 1 and pages 4-9; the whole document -----	1-19
	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 December 2014	23/12/2014

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Vix, Olivier

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2007/100412 A2 (UNIV YALE [US]; BREAKER RONALD R [US]; BLOUNT KENNETH F [US]; PUSKARZ) 7 September 2007 (2007-09-07) in particular see 3-5, 12-18; the whole document	1-19
Y	WO 2006/026712 A2 (FOX CHASE CANCER CT [US]; SEREBRIISKII ILYA [US]; GOLEMIS ERICA A [US]) 9 March 2006 (2006-03-09) detection method using the classic two-hybrid system, see claim 7, pages 2-3,6-7; the whole document	1-19
Y	DAMBACH M D ET AL: "Expanding roles for metabolite-sensing regulatory RNAs", CURRENT OPINION IN MICROBIOLOGY, CURRENT BIOLOGY LTD, GB, vol. 12, no. 2, 1 April 2009 (2009-04-01), pages 161-169, XP026036695, ISSN: 1369-5274, DOI: 10.1016/J.MIB.2009.01.012 [retrieved on 2009-02-26] incentive to use synthetic riboswitches for metabolite sensing - see page 161 and page 165 and conclusion page 168; the whole document	1-19
A	ALEXANDER SERGANOV ET AL: "A Decade of Riboswitches", CELL, vol. 152, no. 1-2, 1 January 2013 (2013-01-01), pages 17-24, XP055078909, ISSN: 0092-8674, DOI: 10.1016/j.cell.2012.12.024 the whole document	1-19
A	FRANCK MARTIN: "Fifteen years of the yeast three-hybrid system: RNA-protein interactions under investigation", METHODS, vol. 58, no. 4, 1 December 2012 (2012-12-01), pages 367-375, XP055096666, ISSN: 1046-2023, DOI: 10.1016/j.ymeth.2012.07.016 the whole document	1-19

-/--

1

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/070954

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MISO PARK ET AL: "Microbial Biosensors: Engineered Microorganisms as the Sensing Machinery", SENSORS, vol. 13, no. 5, 6 May 2013 (2013-05-06), pages 5777-5795, XP055096775, ISSN: 1424-8220, DOI: 10.3390/s130505777 discussion of systems using whole cell sensor for novel targets based on genetic/protein engineering; the whole document	1-19
Y	WO 2012/153142 A2 (UNIV MANCHESTER [GB]; DIXON NEIL [GB]; MICKLEFIELD JASON [GB]) 15 November 2012 (2012-11-15) in particular see claims 1-10 and pages 5-10; the whole document	1-19
A	US 2013/004980 A1 (HUANG JIAN-DONG [CN] ET AL) 3 January 2013 (2013-01-03) the whole document	1-19

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2014/070954

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. (means)

   ☐ on paper

   ☒ in electronic form

   b. (time)

   ☒ in the international application as filed

   ☐ together with the international application in electronic form

   ☐ subsequently to this Authority for the purpose of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2014/070954

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011119956 A1	29-09-2011	EP 2553097 A1	06-02-2013
		US 2013035255 A1	07-02-2013
		WO 2011119956 A1	29-09-2011
WO 2011088076 A2	21-07-2011	EP 2524042 A2	21-11-2012
		SG 182365 A1	30-08-2012
		US 2012321647 A1	20-12-2012
		WO 2011088076 A2	21-07-2011
WO 2004027035 A2	01-04-2004	AU 2003288906 A1	08-04-2004
		CA 2499770 A1	01-04-2004
		EP 1546170 A2	29-06-2005
		EP 2233494 A1	29-09-2010
		EP 2322535 A2	18-05-2011
		JP 2006500030 A	05-01-2006
		JP 2010213712 A	30-09-2010
		JP 2012183068 A	27-09-2012
		US 2005053951 A1	10-03-2005
		US 2010041742 A1	18-02-2010
		US 2010190244 A1	29-07-2010
		US 2011150854 A1	23-06-2011
		US 2011151471 A1	23-06-2011
		US 2011152213 A1	23-06-2011
		US 2011152215 A1	23-06-2011
		WO 2004027035 A2	01-04-2004
WO 2007100412 A2	07-09-2007	EP 1974052 A2	01-10-2008
		JP 2009524411 A	02-07-2009
		US 2009305253 A1	10-12-2009
		WO 2007100412 A2	07-09-2007
WO 2006026712 A2	09-03-2006	AT 546532 T	15-03-2012
		DK 1791958 T3	21-05-2012
		EP 1791958 A2	06-06-2007
		ES 2381384 T3	25-05-2012
		US 2008118919 A1	22-05-2008
		WO 2006026712 A2	09-03-2006
WO 2012153142 A2	15-11-2012	EP 2707487 A2	19-03-2014
		US 2014213779 A1	31-07-2014
		WO 2012153142 A2	15-11-2012
US 2013004980 A1	03-01-2013	CN 103797118 A	14-05-2014
		EP 2726608 A1	07-05-2014
		US 2013004980 A1	03-01-2013
		WO 2013000278 A1	03-01-2013