Standardization of inducer-activated broad host range expression modules:

Debugging and refactoring an alkane-responsive AlkS/*P_{alkB}* device

by

Alejandro Arce-Rodríguez[✉], Ilaria Benedetti, Rafael Silva-Rocha[✉] and Víctor de Lorenzo*

Systems Biology Department, Centro Nacional de Biotecnología-CSIC, Campus de Cantoblanco,
Madrid 28049, Spain

Running title: Standardization of heterologous expression devices

Keywords: SEVA, AlkS, *Pseudomonas*, Crc, Hfq, cytometry, noise

* Correspondence to: Víctor de Lorenzo
Centro Nacional de Biotecnología-CSIC
Campus de Cantoblanco, Madrid 28049, Spain
Tel.: 34- 91 585 45 36; Fax: 34-91 585 45 06
E-mail: vdlorenzo@cnb.csic.es

Current addresses: [✉]Helmholtz-Zentrum für Infektionsforschung, Molecular Bacteriology
Department 38124 Braunschweig Germany, [✉]Ribeirão Preto Medical
School (FMRP), University of São Paulo Ribeirão Preto 14049-900 Brasil

ABSTRACT

Although inducible heterologous expression systems have been available since the birth of recombinant DNA technology, the diversity of devices and genetic architectures of the corresponding vectors have often resulted in a lack of reproducibility and interoperability. In an effort to increase predictability of expression of genes of interest in a variety of possible bacterial hosts we propose a composition standard for debugging and reassembling all regulatory parts that participate in the performance of such devices. As a case study we address the *n*-octane and dicyclopropyl ketone (DCPK)-inducible P_{alkB} promoter of the alkane biodegradation pOCT plasmid of *Pseudomonas putida*. The standardized expression module consisted of an edited *alkS* regulatory gene that is divergently expressed and separated of P_{alkB} by a synthetic DNA buffer sequence. The native DNA sequence of the structural *alkS* gene was modified to alleviate the catabolite repression exerted by some carbon and nitrogen sources through the Crc/Hfq complex of some hosts. The P_{alkB} promoter along with the *alkS* variants were then formatted as SEVA (Standard European Vector Architecture) cargoes and their activity parameters in *P. putida* determined with GFP and luminiscent reporters. The thereby refactored system showed improvements in various features desirable in conditional expression modules: inducibility, capacity, noise reduction and on/off ratio. When applied to other promoter/regulator pairs, the compositional standard thereby implemented in the *AlkS/P_{alkB}* module will enable more complex genetic programming in non-model bacteria.

INTRODUCTION

Heterologous expression of genes in hosts (e.g. bacteria) different from their native origin and triggered by an external inducer is one of the bases of modern Biotechnology (1). A large number of genetic devices to this end have been developed over the years following the pioneering use of the IPTG-inducible *lac* promoter for expression in *E. coli* (2). Typical modules involve one promoter that is activated or repressed by a cognate transcription factor (TF; either an activator or a repressor), which binds the target sequence or changes its activity in a fashion dependent on addition to the medium of a chemical inducer. This basic scheme is the template for a large

number of popular expression vectors based on a suite of regulator/promoter pairs (1, 3, 4). While they have been useful when the issue was to express one or few genes at time in one host, the onset of Synthetic Biology in recent years has multiplied the need of multiple, regulatable promoters endowed with specific parameters and as independent as possible of the physiological state of the host (4). One step in that direction was the creation in 2013 of the so-called Standard European Vector Architecture (5), which compiles a large number of standardized antibiotic marker genes, broad host range origins or replication and functional cargoes, aimed at simplifying genetic programming of a wide variety of bacteria of industrial and environmental interest (6). One type of such cargoes includes expression modules. While the boundaries of the corresponding DNA segments within the plasmid vector or transposon vector frame are well defined in the SEVA format, the organization of the regulatory elements inside the cargo have not been standardized yet. In this work we propose a specific arrangement for such inducible modules that attempts to preserve the inducibility of the TF/promoter pairs usable for heterologous expression while de-compressing and simplifying the native regulation of the cognate systems.

As a case study for such standardization effort we have chosen the regulatory node that controls expression of the *alk* genes for biodegradation of octane borne by the so-called OCT plasmid of soil bacterium *Pseudomonas putida* GPo1 (7). In its native arrangement, two gene clusters are involved in the process. *alkST* encodes the *n*-octane-inducible transcriptional regulator of the pathway (AlkS) and AlkT (a component of alkane hydroxylase). The second *alkBFGHJKL* cluster determines the rest of the activities, which are expressed from the upstream AlkS-dependent promoter P_{alkB} (8, 9). Once excised from its native context and assembled adjacent to each other in a single DNA segment, the *alkS*/ P_{alkB} pair has been used to develop a number of biosensors for alkanes as well as heterologous expression vectors (10-13). The last is facilitated by the use of the gratuitous, soluble inducer dicyclopropyl ketone (DCPK). However, this simple rearrangement of functional segments with *alkS* and P_{alkB} does not eliminate the regulatory complexity embodied in them. *alkS* is transcribed through two promoters, P_{alkS1} and P_{alkS2} , which are negatively and positively regulated, respectively by AlkS. In addition, translation of AlkS is subject to the post-transcriptional control of the Crc/Hfq complex, which enters a catabolite repression signal in the system (14). Finally, the activity of AlkS seems to be influenced also by

the cytochrome terminal oxidase Cyo (15). While such a regulatory density allows the extant system to compute many physiological signals other than the mere presence of pathway substrates, such intricacy is also a nuisance for the predictability of the gene expression module. In the work presented below we have constructed a refactored $AlkS/P_{alkB}$ device in which any known regulatory control—other than induction by DCPK— has been erased and replaced by non-regulatory DNA sequences, following a defined composition standard. As shown below, the resulting inducible expression module displays an improved performance as compared to its counterparts with the wild-type sequences. On this basis, we advocate the general application of the compositional standard used to assemble this device for increasing the reproducibility and interoperability of a large number of other devices made with regulatory parts mined from the genomes of environmental bacteria.

MATERIALS AND METHODS

Strains, plasmids, and growth conditions

The bacteria and the plasmids used in this work are listed in Table 1. All *Pseudomonas putida* strains were derived from the reference variant KT2440 or its rifampicin resistant derivative KT2442. The *E. coli* strains CC118 and DH5 α were used as hosts for maintenance of plasmids. Unless indicated otherwise, cells were grown at 30°C (*P. putida*) or 37°C (*E. coli*) in rich LB medium (16) amended, where necessary, with 100 μ g/ml ampicillin (Ap) or 50 μ g/ml streptomycin (Sm) to retain plasmids. For solid media preparation, LB medium was supplemented with 1.5% (w/v) Bacto Agar (Pronadisa). Where indicated, the expression of P_{alkB} promoter was induced by the addition of dicyclopropylketone (DCPK) in solid and liquid media at the concentrations indicated.

Recombinant DNA techniques

General methods for DNA manipulation were performed with standard protocols described elsewhere (16). The amplification of DNA fragments by polymerase chain reaction (PCR) was implemented in 50 μ l reactions containing approximately 100 ng of genomic DNA or 10 ng of plasmid as template, 0.25 mM dNTPs, 25 pmol of each primer and 1 U of GoTaq DNA

polymerase (Promega). Reactions were run by an initial denaturalization (5 min, 94°C) followed by 30 cycles of denaturalization (1 min, 94°C), annealing (1 min, 58°-64°C), extension (1-3 min at 72°C) and final extension (10 min, 72°C). PCR products were purified with the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel) and, when required, digested with restriction enzymes purchased from New England Biolabs. Plasmid DNA was isolated by means of the Wizard® Plus SV Minipreps DNA Purification system (Promega). *E. coli* cells were transformed with plasmids with the CaCl₂ method (16). In the case of *P. putida*, plasmids were incorporated by conjugative triparental mating using the *E. coli* HB101 (pRK600) as helper strain (17) or by electroporation of cells previously washed and concentrated with 300 mM sucrose as in Choi *et al.* (18).

Construction of the *P_{alkB}*/AlkS expression modules compatible with the Standard European Vector Architecture (SEVA)

In order to create an expression system that could fit the SEVA plasmid platform (5, 6), the gene *alkS* from *Pseudomonas putida* GPo1 was edited to remove incompatible restriction sites (*alkS^{ED}*). The modified sequence (Supplementary Fig. S1) was fully synthesized by GeneArt/ThermoFisher (Waltham, Massachusetts) and delivered as an insert in the pMA vector that was called pAlkS3. Both the AlkS regulatory protein and the *P_{alkB}* promoter were then arrayed with the other DNA segments indicated in Fig. 1. The *P_{alkB}* promoter was PCR-amplified with primers aaPalkB1-F (5'**AGCGGATAACAATTTACACAGGAC**GTGTTTTTCCAGCAGACGAC3') and aaPalkB1-R (5'ATGACCTAGGCTCTCGACATCTTAAACCTGAGC3'), using as template genomic DNA from *P. putida* PBS4 (19). The *P_{neo}* promoter was also amplified by PCR with oligonucleotides aaPKm-NcoI1-F (5'TAGAA**CCATGG**TTTTTCCTCCTTATAAAG3') and aaM13-R24-rev (5' TCCTGTGTGAAATTGTTATCCGCT 3') from pBAM1 (Table 1). The sequence of primer aaM13-R24-rev is complementary to the 5' end of primer aaPalkB1-F (bold characters), allowing the assembly of *P_{neo}* with *P_{alkB}* by SOEing PCR (20). Insertion of the resulting product into the NcoI/AvrII restriction sites of a pSEVA frame gave rise to expression vector pSEVA429 *crc*⁺ (Supplementary Fig. S2)

Removal of the Crc binding site in the *alkS* gene

The *Crc* binding site in the 5' end of the *alkS* gene in pSEVA429 *crc*⁺ was -modified with PCR-based site-directed mutagenesis. Briefly, the first 570 bp of *alkS* were amplified with the oligonucleotides 5-*alkS*mut-NcoI (5'GCGCCCATGGGCATGAAGATCAAGATCATCAACAATGATTTCCCGGTTGCCAAGATCG3') and 3-*alkS*mut-XhoI (5'AGCGCCTGCAAGTTTAAGCC3') using the pAlkS3 as template. The forward primer contains the recognition sequence of NcoI enzyme (underlined) and also 6 single-nucleotide mutations introduced deliberately to eliminate the *Crc* site from *alkS* (in bold characters). The PCR fragment was then restricted with the enzymes NcoI and XhoI (the last coded in between the *alkS* sequence) and the resulting 406 bp fragment was recloned into the same sites of pSEVA429 *crc*⁺. In order to follow the SEVA nomenclature, this final expression vector was named pSEVA429 (i.e. Sm resistance, RK2 origin of replication and AlkS/*P*_{alkB} expression cargo; Supplementary Fig. S2)

Parameterization of the AlkS/*P*_{alkB} module

For generating constructs that report transcriptional activity as a fluorescent emission, the *GFP tir* gene of pSEVA427 was cloned into plasmids pSEVA429 *crc*⁺ and pSEVA429 as a HindIII/Spel insert to generate plasmids pSEVA429 *crc*⁺ → GFP and pSEVA429 → GFP respectively. As a control, we excised the NotI fragment from pJAMA30 containing the original *alkST* genes of pOCT transcribed by their own *P*_{alkST} promoter and, in divergent orientation, the *P*_{alkB} driving the expression of the *GFP tir* gene (11; Supplementary Fig. S3). This ~6.7 kb fragment was cloned into the NotI site of pSEVA421 to generate the control vector pARalkS. The three *GFP tir* reporter vectors described above, were transferred into wildtype *P. putida* KT2440 cells and into the *crc::tet* derivative of *P. putida* KT2442. Plasmid-bearing strains were then grown at 30°C until mid-exponential phase, the cultures added with 0.05% v/v DCPK and fluorescent emission for the next 3 hours followed as described in detail in (21) with a Gallios (Perkin Elmer) flow cytometer. GFP was excited at 488 nm, and the fluorescence signal was recovered with a 525/40 BP filter. The resulting data was processed using FlowJo v. 9.6.2 software (FlowJo LLC, Ashland, OR, USA). For monitoring activity of the AlkS/*P*_{alkB} module at a population level, plasmids pSEVA429 *crc*⁺ and pSEVA429 were inserted with the promoterless luminescent reporter *luxCDABE* operon excised from pSEVA426 as a HindIII/Spel fragment. This originated plasmids pSEVA429 *crc*⁺ → *luxCDABE* and pSEVA429 → *luxCDABE* respectively. As with the GFP counterparts before, these plasmids were passed to wild type and *crc*-minus *P. putida* hosts. For measuring light

emission under various DCPK concentrations, cells grown overnight in LB were diluted in same medium, placed in a Microtest™ 96-well assay plate (BD Falcon), regrown to mid-exponential phase, added with the inducer and luminescence recorded after 4 h.

RESULTS AND DISCUSSION

A compositional standard for engineering inducer-dependent heterologous gene expression

Fig. 1A sketches the organization of the inducible expression module proposed in this work for SEVA cargoes (5) aimed at heterologous expression of genes of interest in a variety of bacterial hosts. First, the standard asks for constitutive expression of the gene(s) encoding the effector-responsive regulator. In their natural context, TFs are often subject to a degree of self-regulation, either positive or negative (22) that enters a layer of unnecessary complexity that the arrangement shown in Fig. 1A mitigates if not entirely eliminates. The standardized sequence that connects the various functional parts of the device (Fig. 1B) includes the following segments. The source of transcription of the signal-responsive TF gene is the 106 bp minimal promoter P_3/P_{neo} that drives expression of the kanamycin resistance gene of pBAM1 (23). Following this promoter, default translation efficiency is also fixed by means of a 24 bp translation initiation region (TIR) retrieved from the GFP variant borne by pGreenTIR plasmid (24). This is an unusual ribosome binding sequence known to act as a translational enhancer that is expected to curb expression noise that could originate in a poor translation. The segment for constitutive expression/translation of the TF is followed upstream by a 150 bp neutral segment of DNA retrieved from the $lacI^q$ - P_{trc} expression system of plasmid pTrcA (25). This sequence has no known regulatory elements and it functions as a buffer region to ease mutually negative supercoiling that could stem from transcription of divergent promoters (26). As shown in Fig. 1A, the adjacent piece of DNA is the one that bears the promoter targeted by the inducible regulator and orientated opposite in respect to the sequences for expression of the TF gene. The specific DNA sequence of this promoter obviously changes from case to case but it should by default be accommodated within a segment of not more than 100 pb. The 3' of this promoter sequence is bound by an AvrII site, which links this segment to the start of the SEVA polylinker (5). Note that two sites at the boundary buffer sequence/promoter and at the end of the MCS have target

sequences for oligonucleotides R24 and F24 (5) for easy amplification and diagnose of possible inserts. Once the gene encoding the inducer-responsive TF is placed in this arrangement as a NcoI (overlapping the leading ATG)-PacI DNA fragment, the whole expression module becomes inserted in the SEVA frame as a PacI-AvrII addition (Fig. 1C), shielded both upstream and downstream by transcriptional terminators contributed by the vector structure—and ready to be inserted with any gene of interest (GOI) cloned in the corresponding polylinker.

Reshaping the *alkS* and *P_{alkB}* pair as an inducible expression device

As a case study of formatting a naturally-occurring inducible promoter into a standardized expression cargo we picked the regulatory elements that control transcription of the *alk* genes of the OCT plasmid of *Pseudomonas putida* GPo1 (7, 8). The choice was motivated by the exemplary regulatory density of the extant system, that includes transcriptional and post-transcriptional control layers checking expression of AlkS (Fig. 2A). This provided an archetypal case to inspect the impact of the simplified formatting explained above on the behaviour of the resulting expression module. To this end, we first edited the wild-type DNA sequence of the regulator to eliminate restriction sites incompatible with the SEVA rules while preserving the primary amino acid sequence. The resulting DNA segment was then produced as a 5' → 3' 2668 bp NcoI-PacI DNA fragment. The complete list of changes entered in this *alkS* variant (that we term *alkS^{ED}* for edited) is compiled in Supplementary Fig. S1. The *alkS^{ED}* variant was then coupled to the spacer shown in Fig. 1B, which was added with the wild-type 90 bp sequence of the target *P_{alkB}* promoter (see Materials and Methods). The resulting cargo was subsequently placed in the frame of plasmid pSEVA421 (6) as a PacI-AvrII insert, resulting in expression vector pSEVA429 *crc+* (Table 1; Supplementary Fig. S2). For parameterization of the activity of the thereby refactored expression device, the construct was added with the promoterless GFP *tir* gene of pSEVA427 (6) downstream of *P_{alkB}* (Materials and Methods) resulting in reporter plasmid pSEVA429 *crc+* → GFP. To have a control of the wild-type expression device endowed of all native regulatory parts present in plasmid pOCT we excised the ~6.7 kb NotI fragment of plasmid pJAMA30 (11) containing *alkST* and a divergent *P_{alkB}* promoter upstream of a GFP reporter (Supplementary Fig. S3). This DNA was then inserted into the same plasmid frame of pSEVA421 used for the edited devices, thereby originating control plasmid pAR*alkS* (Table 1; note that the *alkS* variant in this case is the very original wild type *alkS^{WT}* as shown in Supplementary Fig. S1).

Once equivalent constructs with *alkS^{WT}* (pAR*alkS*) and *alkS^{ED}* (pSEVA429 *crc⁺* → GFP) were in hand, we were able to evaluate the effect of the standardization of the architecture of the regulatory node on transcriptional performance.

The effect of formatting *AlkS/P_{alkB}* in the performance of the expression system

In a first series of experiments we compared the behaviour of the *AlkS/P_{alkB}* pair assembled with all the native regulatory system borne by the pOCT plasmid (Fig. 2A) and recreated in plasmid pAR*alkS* versus that of the same regulatory system arranged with the composition standard of Fig. 1 as implemented in pSEVA429 *crc⁺* → GFP. Differences included a synthetic expression segment for transcription and translation of the regulator and an upstream buffer DNA sequence that was followed by the divergent *AlkS* target promoter *P_{alkB}* as sketched in Fig. 2B. Wild-type *P. putida* KT2440 was transformed with each of these two plasmids and transformants were grown in LB medium with Sm and fluorescent readout followed as explained in the Materials and Methods section. The cytometry results of these experiments are shown in Fig. 3. Inspection of the resulting graphs revealed some features of both the native and the formatted system that are worth to consider for handling of the expression device. First, formatted or not, the *P_{alkB}* has a noticeable non-induced basal transcription level that spontaneously increased with growth (Fig. 3A). Although this level is not too high (it remains within the same order of magnitude than the baseline expression), it has to be taken into account when expression of toxic proteins or coupling with other devices are pursued (27). A second aspect of the results shown in Fig. 3 is the noticeably stronger and faster induction of the standardized module of pSEVA429 *crc⁺* → GFP as compared to the original regulatory node of pAR*alkS*. Finally, both devices displayed a growingly sharp unimodal expression pattern (28) along the induction period with little or no significant noise or cell-to-cell variation. While these results were good enough to accredit the correct functioning of the expression device following reassembly of its DNA parts, we wondered about the less predictable effects of other physiological inputs that operate on the *AlkS/P_{alkB}* in its native context.

Effect of *Crc* on the formatted and not-formatted *AlkS/P_{alkB}* device

Ideally, for engineering reliable genetic devices regulatory parts should deliver their function in a context-independent manner (29). Such a context includes not only genomic locations (30) and availability of resources (31), but also physiological signals (32, 33) that orchestrate the induction hierarchy. One of these is catabolite repression (34), which in the case of *P. putida* operates through a complex interplay between proteins Crc and Hfq with small RNAs to inhibit translation of mRNAs of target genes (14). AlkS happens to be subject to such post-transcriptional regulation when placed in *P. putida*, but not in *E. coli* (35). In order to calibrate the effect of such a control layer and whether it was kept or not in the standardized construct, we run the experiments shown in Fig. 4. To this end pARalkS and pSEVA429 *crc*⁺ → GFP were placed in a *crc::tet* mutant of *P. putida* (Table 1) known to be blind to catabolite repression caused by many of the components of LB medium (36, 37). Although not as intense as reported in the native configuration an effect of lacking Crc became evident, specifically on the net accumulation of reporter signal in both constructs as well as an increase on the noise of basal promoter activity at longer growth periods. But otherwise, the overall behaviour of the standardized and non-standardized expression devices remained basically identical.

The data above were welcome results, as refactoring of an existing regulatory node with a different architecture often results in devices that do worse than the naturally-occurring setup (38, 39). But the question still remained of whether we could erase altogether the effect of *crc* on the performance of the AlkS/*P_{alkB}* module, not by playing with the genetic background but by rewriting the DNA sequence of the regulator. To address this, an additional *alkS* derivative was synthesized in which the Hfq/Crc binding site of the corresponding mRNA 5'-AATAAAAATAATAAA-3' that overlaps amino acid positions 4-9 of the protein was edited to 5'-GATCAAGATCATCAA-3'. As shown in Supplementary Fig. S1, these changes are predicted to abolish altogether the target site for Hfq/Crc, but they still keep the wild-type primary amino acid sequence of the encoded AlkS. As before, the resulting DNA (that bears the variant hereafter called *alkS^{CR}* for Crc-free) was formatted as a NcoI-PacI fragment and coupled to the spacer shown in Fig. 1B and the *P_{alkB}* promoter. The resulting cargo was subsequently placed in the frame of plasmid pSEVA421 as PacI-AvrII insert, resulting in expression vector pSEVA429 (Supplementary Fig. S2). For the sake of comparing its performance, a promoterless GFP gene identical to that of pARalkS and pSEVA429 *crc*⁺ → GFP, was added to pSEVA429, thereby

generating pSEVA429 → GFP. This *alkS^{CR}*-containing plasmid was then placed in *crc⁺* and *crc⁻* strains of *P. putida* and the readout of the fluorescent reporters followed in LB medium with or without DCPK induction as before. As shown in Fig. 5, removal of the *crc* site from the *alkS* sequence had only a moderate effect on the performance of the expression device, as the induction patterns were quite similar when pSEVA429 → GFP was placed in isogenic *P. putida* strains with or without the factor.

To gain some insight into this apparently minor influence of *crc* we constructed additional derivatives of pSEVA429 *crc⁺* and pSEVA429 with a promoterless *luxCDABE* operon instead of GFP. This luminiscent reporter is way more sensitive than GFP and therefore a better proxy of transcriptional output from *P_{alkB}* at a population level. The new construct was placed as before in equivalent *crc⁺* and *crc⁻* strains and the cognate transformants grown in LB with different inducer concentrations. In this case (Fig. 6) both the overall weight of Crc on the system and the effect of removing its target site in *alkS* mRNA became clear. Working in an entirely Crc-free strain exposed to 0.05% DCPK equalized *P_{alkB}* activity at high levels regardless of whether the *alkS* variant of the expression vector had the *crc* site or not (Fig. 6C). In contrast, when placed in a wild-type strain, the construct with *alkS^{CR}* produced luminescence levels approximately twice as high those as the same with *alkS^{ED}*. While this accredited an inhibitory role of Crc on AlkS, removal of the related sequence failed to make the AlkS/*P_{alkB}* altogether blind to catabolite repression—surely because this physiological control influences additional components of the system other than the transcriptional regulator

CONCLUSIONS

In this work we have used the inducer dependent and AlkS-mediated activation of the *P_{alkB}* promoter of the pOCT plasmid as an example of the roadmap that should be followed for reshaping a naturally-occurring regulatory node into an standardized device for engineering heterologous expression (40). As with any standard, there is a somewhat arbitrary but still reasonable and scientifically justifiable choice of a given composition rule (41). The one we propose in this work is summarized in Fig. 1 and explained in detail above. As is also the case of other standards, this particular choice will certainly limit flexibility but will foster interoperability, parameterization and comparative metrology (42, 43). The work above exemplifies how the same

device, still after formatting, may go through successive, improved versions of the same functional DNA segments even if a prefixed arrangement is kept constant. In the cases examined above we can consider plasmids pARalkS and pSEVA429 *crc*⁺ as beta versions of what we propose to be the *bona fide* standardized AlkS/*P*_{alkB} expression device apt for inclusion as a cargo in the SEVA collection: pSEVA429 (Supplementary Fig. S2). Note however that—as shown above—there is still room for improvement and it is likely that other versions will follow, an issue that is contemplated in the updated nomenclature of the SEVA collection (6). For instance, the system could be refactored to make it more digital (e.g. suppressing its considerable basal level; 27) and making it still more independent of catabolite repression. But, in the meantime, pSEVA429 is an altogether standardized off-the-shelf expression vector with a large number of benefits, including the possibility of comparing faithfully its performance with other expression modules that follow the same arrangement. We ultimately expect such standardization to ease engineering of complex systems and encourage other genetic tool developers to follow suit.

SUPPLEMENTARY DATA

Supplementary Figure S1. Alignment of the wild-type *alkS* sequence with its edited variants *alkS*^{ED} and *alkS*^{CR}.

Supplementary Figure S2. Organization of AlkS-based expression vectors

Supplementary Figure S3. Functional insert of plasmids pJAMA30 and pARalkS

Acknowledgments

Authors are indebted to Fernando Rojo and Renata Moreno for strains and valuable materials.

Funding

This work was funded by the SETH (RTI2018-095584-B-C42) (MINECO/FEDER) and SyCoLiM (ERA-COBIOTECH 2018 - PCI2019-111859-2) Projects of the Spanish Ministry of Science and Innovation, the MADONNA (H2020-FET-OPEN-RIA-2017-1-766975), BioRoboost (H2020-NMBP-BIO-CSA-2018-820699), SynBio4Flav (H2020-NMBP-TR-IND/H2020-NMBP-BIO-2018-814650)

and MIX-UP (MIX-UP H2020-BIO-CN-2019-870294) Contracts of the European Union and the InGEMICS-CM (S2017/BMD-3691) Project of the Comunidad de Madrid - European Structural and Investment Funds - (FSE, FECER).

REFERENCES

1. Terpe K (2006) Overview of bacterial expression systems for heterologous protein production: From molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotech* 72, 211-22.
2. Browning DF, Godfrey RE, Richards KL, Robinson C, Busby SJW (2019) Exploitation of the *Escherichia coli* lac operon promoter for controlled recombinant protein production. *Biochem Soc Trans* 47, 755-763.
3. Marschall L, Sagmeister P, Herwig C (2017) Tunable recombinant protein expression in *E. coli*: promoter systems and genetic constraints. *Appl Microbiol Biotech* 101, 501-512.
4. Martínez-García E, de Lorenzo V (2017) Molecular tools and emerging strategies for deep genetic/genomic refactoring of *Pseudomonas*. *Curr Opin Biotechnol* 47, 120-132.
5. Silva-Rocha R, Martinez-Garcia E, Calles B, Chavarria M, Arce-Rodriguez A, de Las Heras A, Paez-Espino AD, Durante-Rodriguez G, Kim J, Nikel PI, Platero R, de Lorenzo V (2013) The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucl Acids Res* 41, D666-D675.
6. Martínez-García E, Goñi-Moreno A, Bartley B, McLaughlin J, Sánchez-Sampedro L, Del Pozo HP, Hernández CP, Marletta AS, De Lucrezia D, Sánchez-Fernández G, Fraile S, de Lorenzo V (2020) SEVA 3.0: an update of the Standard European Vector Architecture for enabling portability of genetic constructs among diverse bacterial hosts. *Nucleic Acids Res* 48, 3395.
7. Rojo F (2009) Degradation of alkanes by bacteria. *Environ Microbiol* 11, 2477-2490.
8. van Beilen JB, Panke S, Lucchini S, Franchini AG, Röthlisberger M, Witholt B (2001) Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the *alk* genes. *Microbiology* 147, 1621-1630.

9. Canosa I, Sánchez-Romero JM, Yuste L, Rojo F (2000) A positive feedback mechanism controls expression of AlkS, the transcriptional regulator of the *Pseudomonas oleovorans* alkane degradation pathway. *Mol Microbiol* 35, 791-799.
10. Sticher P, Jaspers MC, Stemmler K, Harms H, Zehnder AJ, van der Meer JR (1997) Development and characterization of a whole-cell bioluminescent sensor for bioavailable middle-chain alkanes in contaminated groundwater samples. *Appl Environ Microbiol* 63, 4053–4060.
11. Jaspers MC, Meier C, Zehnder AJ, Harms H, van der Meer JR (2001) Measuring mass transfer processes of octane with the help of an *alkS alkB::gfp*-tagged *Escherichia coli*. *Environ Microbiol* 3, 512-524.
12. Makart S, Heinemann M, Panke S (2007) Characterization of the AlkS/PalkB-expression system as an efficient tool for the production of recombinant proteins in *Escherichia coli* fed-batch fermentations. *Biotechnol Bioeng* 96, 326-336.
13. Reed B, Blazeck J, Alper H (2012) Evolution of an alkane-inducible biosensor for increased responsiveness to short-chain alkanes. *J Biotechnol* 158, 75-79.
14. Hernández-Arranz S, Sánchez-Hevia D, Rojo F, Moreno R (2016) Effect of Crc and Hfq proteins on the transcription, processing, and stability of the *Pseudomonas putida* CrcZ sRNA. *RNA* 22, 1902-1917.
15. Dinamarca MA, Ruiz-Manzano A, Rojo F (2002) Inactivation of cytochrome o ubiquinol oxidase relieves catabolic repression of the *Pseudomonas putida* GPo1 alkane degradation pathway. *J Bacteriol* 184, 3785-3793.
16. Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor, New York.
17. de Lorenzo V, Timmis KN (1994) Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Meth Enzymol* 235, 386-405.
18. Choi KH, Kumar A, Schweizer HP (2006) A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J Microbiol Meth* 64, 391-397.

19. Yuste L, Canosa I, Rojo F (1998) Carbon-source-dependent expression of the *PalkB* promoter from the *Pseudomonas oleovorans* alkane degradation pathway. *J Bacteriol* 80, 5218-5226.
20. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77, 61-68.
21. Benedetti I, de Lorenzo V, Nikel PI (2016) Genetic programming of catalytic *Pseudomonas putida* biofilms for boosting biodegradation of haloalkanes. *Metab Eng* 33, 109-118.
22. Alon U. 2019. *An introduction to systems biology: design principles of biological circuits*. CRC press.
23. Martínez-García E, Calles B, Arévalo-Rodríguez M, de Lorenzo V (2011) pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. *BMC Microbiol* 11, 38.
24. Miller WG, Lindow SE (1997) An improved GFP cloning cassette designed for prokaryotic transcriptional fusions. *Gene* 191, 149-53.
25. Amann E, Ochs B, Abel K-J (1988) Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* 69, 301-315.
26. Wu H-Y, Shyy S, Wang JC, Liu LF (1988) Transcription generates positively and negatively supercoiled domains in the template. *Cell* 53, 433-440.
27. Calles B, Goñi-Moreno Á, de Lorenzo V (2019) Digitalizing heterologous gene expression in Gram-negative bacteria with a portable ON/OFF module. *Mol Syst Biol* 15, e8777.
28. Kærn M, Elston TC, Blake WJ, Collins JJ (2005) Stochasticity in gene expression: from theories to phenotypes. *Nature Rev Genet* 6, 451-464.
29. Costello A, Badran AH (2020) Synthetic biological circuits within an orthogonal central dogma. *Trends Biotechnol* <https://doi.org/10.1016/j.tibtech.2020.05.013>
30. Bryant JA, Sellars LE, Busby SJW, Lee DJ (2014) Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucl Acids Res* 42, 11383-11392.
31. Borkowski O, Ceroni F, Stan G-B, Ellis T (2016) Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Curr Op Microbiol* 33, 123-130.

32. Cases I, De Lorenzo V (2005) Promoters in the environment: transcriptional regulation in its natural context. *Nature Revs Microbiol* 3, 105-118.
33. Shingler V (2011) Signal sensory systems that impact σ^{54} -dependent transcription. *FEMS Microbiol Revs* 35, 425-440.
34. Rojo F (2010) Carbon catabolite repression in *Pseudomonas*: optimizing metabolic versatility and interactions with the environment. *FEMS Microbiol Revs* 34, 658-684.
35. Yuste L, Rojo F (2001) Role of the *crc* gene in catabolic repression of the *Pseudomonas putida* GPo1 alkane degradation pathway. *J Bacteriol* 183, 6197-6206.
36. Molina L, La Rosa R, Nogales J, Rojo F (2019) Influence of the Crc global regulator on substrate uptake rates and the distribution of metabolic fluxes in *Pseudomonas putida* KT2440 growing in a complete medium. *Environ Microbiol* 21, 4446-4459.
37. Molina L, Rosa RL, Nogales J, Rojo F (2019) *Pseudomonas putida* KT2440 metabolism undergoes sequential modifications during exponential growth in a complete medium as compounds are gradually consumed. *Environ Microbiol* 21, 2375-2390.
38. Chan LY, Kosuri S, Endy D (2005) Refactoring bacteriophage T7. *Mol Sys Biol* 1, 2005.0018.
39. Temme K, Zhao D, Voigt CA (2012) Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc Natl Acad Sci USA* 109, 7085-7090.
40. Martínez-García E, Benedetti I, Hueso A, de Lorenzo V (2015) Mining environmental plasmids for synthetic biology parts and devices. *Microbiol Spectr* 3, Plas-0033-2014.
41. Porcar M, Danchin A, de Lorenzo V (2015) Confidence, tolerance, and allowance in biological engineering: the nuts and bolts of living things. *Bioessays* 37, 95-102.
42. de Lorenzo V, Schmidt M (2018) Biological standards for the Knowledge-Based BioEconomy: What is at stake. *New Biotechnol* 40, 170-180.
43. Beal J, Goñi-Moreno A, Myers C, Hecht A, de Vicente MDC, Parco M, Schmidt M, Timmis K, Baldwin G, Friedrichs S, Freemont P, Kiga D, Ordozgoiti E, Rennig M, Rios L, Tanner K, de Lorenzo V, Porcar M (2020) The long journey towards standards for engineering biosystems: Are the Molecular Biology and the Biotech communities ready to standardise? *EMBO Rep* 21, e50521.

- 1 44. Manoil C, Beckwith J (1985) *TnphoA*: a transposon probe for protein export signals. *Proc*
2 *Natl Acad Sci USA* 82, 8129-8133.
- 3 45. Hanahan D, Meselson M (1983) Plasmid screening at high colony density. *Methods Enzymol*
4 100, 333-342.
- 5 46. Nelson KE, Weinl C, Paulsen IT, Dodson RJ, Hilbert H, Martins dos Santos VA, Fouts DE,
6 Gill SR, Pop M, Holmes M, Brinkac L, Beanan M, DeBoy RT, Daugherty S, Kolonay J,
7 Madupu R, Nelson W, White O, Peterson J, Khouri H, Hance I, Chris Lee P, Holtzapple E,
8 Scanlan D, Tran K, Moazzez A, Utterback T, Rizzo M, Lee K, Kosack D, Moestl D, Wedler
9 H, Lauber J, Stjepandic D, Hoheisel J, Straetz M, Heim S, Kiewitz C, Eisen JA, Timmis KN,
10 Dusterhoft A, Tummier B, Fraser CM (2002) Complete genome sequence and comparative
11 analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ Microbiol* 4,
12 799-808.
- 13 47. Moreno R, Rojo F (2008) The target for the *Pseudomonas putida* Crc global regulator in the
14 benzoate degradation pathway is the *benR* transcriptional regulator. *J Bacteriol* 190, 1539–
15 1545.
- 16 48. Kessler B, de Lorenzo V, Timmis KN (1992) A general system to integrate *lacZ* fusions into
17 the chromosomes of Gram-negative eubacteria: regulation of the *Pm* promoter of the TOL
18 plasmid studied with all controlling elements in monocopy. *Mol Gen Genet* 233, 293-301.

19

20

1 Table 1. Strains and plasmids used in this work.

2

Strain/plasmid	Description / relevant characteristics	Reference
<i>E. coli</i> strains		
CC118	F ⁻ , $\Delta(ara-leu)7697$, $araD139$, $\Delta(lac)X74$, $phoA\Delta20$, $galE$, $galK$, thi , $rpsE$, $rpoB$, $argE(Am)$, $recA1$	(44)
DH5 α	F ⁻ , $supE44$, $\Delta lacU169$, ($\phi 80 lacZDM15$), $hsdR17$, ($rk-mk+$), $recA1$, $endA1$, $thi1$, $gyrA$, $relA$	(45)
HB101	Sm ^R , $hsdR^-$ M ⁺ , pro , leu , thi , $recA$	(16)
<i>P. putida</i> strains		
KT2440	$rsdR$; prototrophic, wild-type strain derived of <i>P. putida</i> mt-2 without pWWO plasmid	(46)
KT2442	Spontaneous rifampicin resistant derivative of KT2440	(46)
<i>P. putida</i> PBS4	KT2442 inserted chromosomally with $alkS$ gene and P_{alkB} promoter	(19).
KT2442-C1	Rif ^R $crc::tet$ derivative of <i>P. putida</i> KT2442	(47)
Plasmids		
pRK600	Cm ^R ; $oriV$ ColE1, tra^+mob^+ of RK2; helper plasmid for mobilization in tripartite conjugations	(48)
pMA	Cloning vector for synthetic DNA	GeneArt ¹
pAlkS3	Ap ^R ; pMA cloning vector inserted with the optimized sequence of $alkS$ gene	This work
pBAM1	Mini-Tn5 suicide delivery vector, source of P_{neo} promoter	(23)
pSEVA429 crc^+	Sm ^R , $oriRK2$, $oriT$; pSEVA421-derivative carrying an $alkS^{ED}/P_{alkB}$ expression system. The gene $alkS$ is edited for SEVA-incompatible restriction sites, but encodes the wild type primary amino acid sequence of the AlkS protein and keeps the Crc/Hfq binding sequence in the cognate transcript.	This work
pSEVA429	Sm ^R , $oriRK2$, $oriT$; pSEVA 421-derivative carrying the $alkS^{CR}/P_{alkB}$ expression system. $alkS$ sequence is same as $alkS^{ED}$ (Supplementary Fig. S1) but 5'-end edited for	This work

¹ <https://www.thermofisher.com/content/dam/LifeTech/Documents/geneart/geneart-vector-map.pdf>

	removing the Crc/Hfq binding sequence of the corresponding mRNA.	
pSEVA421	Sm ^R , <i>oriV</i> RK2, <i>oriT</i> , standard MCS	(5)
pSEVA426	Sm ^R , <i>oriV</i> RK2, <i>oriT</i> , <i>luxCDABE</i> reporter system	(5)
pSEVA429 → <i>luxCDABE</i>	Sm ^R , <i>oriV</i> RK2, <i>oriT</i> ; pSEVA 429 cloned with the <i>luxCDABE</i> reporter system	This work
pSEVA427	Sm ^R , <i>oriV</i> RK2, <i>oriT</i> , <i>GFP</i> reporter system	(5)
pSEVA429 <i>crc</i> ⁺ → <i>GFP</i>	pSEVA429 <i>crc</i> ⁺ inserted with promoterless <i>GFP</i> gene as a transcriptional reporter	This work
pSEVA429 <i>crc</i> ⁺ → <i>luxCDABE</i>	pSEVA429 <i>crc</i> ⁺ inserted with promoterless <i>luxCDABE</i> operon as a transcriptional reporter	This work
pSEVA429 → <i>GFP</i>	pSEVA429 inserted with promoterless <i>GFP</i> gene as a transcriptional reporter	This work
pJAMA30	Ap ^R , <i>oriV</i> ColE1; native <i>P_{alkST}-alkST/P_{alkB}</i> expression system driving the transcription of <i>GFP tir</i> . This segment is flanked by NotI sites.	(11)
pARalkS	Sm ^R , <i>oriV</i> RK2, <i>oriT</i> ; pSEVA421 cloned with the NotI fragment from pJAMA30 carrying the native alkane/DCPK-responsive reporter system	This work

1
2
3
4
5

FIGURES

Figure 1. Organization of standardized inducible expression modules.

(A) Arrangement of DNA portion, boundaries and their roles as the frame for inserting genes for inducer-responsive transcriptional regulators and cognate target promoters. (B) Blowup of the standardized DNA sequence that acts and the scaffold for the rest of the functional parts. (C) Configuration of reporter plasmids (with GFP or any other gene of interest, GoI) used in this work. Note constant and variable segments.

Figure 3. Transcriptional activity of *AlkS*/ *P_{alkB}* node before and after standardization as SEVA cargo.

P. putida transformants with either pAR*alkS* (wild-type *alkS*) and pSEVA429 *crc⁺* → GFP (edited *alkS* variant ED) were grown in LB until mid-exponential phase and added with 0.05% *AlkS* effector DCPK. Fluorescent emission was then measured in a cytometer for the next 4 hours as explained in the Materials and Methods section. (A) Control, no induction. (B) Cultures supplemented with inducer.

Figure 4. Effect of Crc on performance of the standardized *AlkS*/*P_{alkB}* module.

Plasmids pAR*alkS* (*alkS^{WT}*) and pSEVA429 *crc⁺* → GFP (*alkS^{ED}*) were placed in a *crc::tet* mutant of *P. putida*, the transformants grown in LB and treated as indicated in the legend of Fig. 3. (A) Control, no induction. (B) Cultures supplemented with inducer 0.05 % DCPK.

Figure 5. Behavior of an *alkS* gene variant erased of the Crc/Hfq binding site to its mRNA.

Plasmid pSEVA429 → GFP (with the *AlkS/P_{alkB}* device bearing the regulator gene devoid of the Crc binding site in its mRNA) was transformed in *crc*⁺ and *crc*⁻ strains of *P. putida*. Transformants were grown in LB and treated as indicated in the legend of Fig. 3. (A) Control, no induction. (B) Cultures supplemented with inducer 0.05 % DCPK.

Figure 6. Influence of Crc at in *P. putida* populations bearing the AlkS/*P*_{alkB} module with the regulator gene with or without the Crc site.

Plasmids pSEVA429 *crc*⁺ → *luxCDABE* (*alkS^{ED}*, with Crc-binding site) and pSEVA429 → *luxCDABE* (*alkS^{CR}*, without Crc site) were transformed in *crc*⁺ and *crc*⁻ strains of *P. putida*, grown in LB until mid-exponential phase and added with the DCPK concentrations indicated in each case. Luminiscent emission was then recorded after 4 h as a proxy of transcriptional activity as explained in the Materials and Methods section. (A) Sketch, not to scale, of functional segments in reporter plasmids. (B) Readout of reporter constructs borne by wild-type *P. putida*. (C) Same in *crc::tet* cells. Note different scales in Y axes.