bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Identification, characterization, and application of a highly sensitive lactam biosensor from *Pseudomonas putida*

Mitchell G. Thompson[1,2,3], Allison N. Pearson[1,2], Jesus F. Barajas[2,4], Pablo Cruz-Morales[1,2,5], Nima Sedaghatian[1,2], Zak Costello[1,2,6], Megan E. Garber[1,2,7], Matthew R. Incha[1,2,3], Luis E. Valencia[1,2,8], Edward E. K. Baidoo[1,2], Hector Garcia Martin[1,2,6,9], Aindrila Mukhopadhyay[1,2,7], Jay D. Keasling[1,2,8,10,11,12]

[1]Joint BioEnergy Institute, 5885 Hollis Street, Emeryville, CA 94608, USA.

[2]Biological Systems & Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

[3]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA

[4]Department of Energy Agile BioFoundry, Emeryville, California, USA

[5]Centro de Biotecnologia FEMSA, Instituto Tecnologico y de Estudios superiores de Monterrey, Mexico

[6]DOE Agile BioFoundry, Emeryville, CA, USA

[7]Comparative Biochemistry Graduate Group, University of California, Berkeley, Berkeley, California, USA

[8]Joint Program in Bioengineering, University of California, Berkeley/San Francisco, CA 94720, USA

[9]BCAM, Basque Center for Applied Mathematics, Bilbao, Spain

[10]Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[11]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Denmark

[12]Center for Synthetic Biochemistry, Institute for Synthetic Biology, Shenzhen Institutes for Advanced Technologies, Shenzhen, China

ABSTRACT

Caprolactam is an important polymer precursor to nylon traditionally derived from petroleum and produced on a scale of 5 million tons per year. Current biological pathways for the production of caprolactam are inefficient with titers not exceeding 2 mg/L, necessitating novel pathways for its production. As development of novel metabolic routes often requires thousands of designs and results in low product titers, a highly sensitive biosensor for the final product has the potential to reduce development time. Here we report a highly sensitive biosensor for valerolactam and caprolactam from *Pseudomonas putida* KT2440 which is >1000x more sensitive to exogenous ligand than previously reported sensors. Manipulating the expression of the sensor *oplR* (PP_3516) substantially altered the sensing parameters, with various vectors showing K_d values ranging from 700 nM to 1.2 mM. Our most sensitive construct was able to detect *in vivo* production of caprolactam above background at 11 μg/L. The high sensitivity and range of OplR is a powerful tool towards the development of novel routes to the biological synthesis of caprolactam.

INTRODUCTION

Caprolactam is an important chemical precursor to the polymer nylon 6, with a global demand approximately 5 million tons per year [1]. Currently the majority of caprolactam is

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

synthesized from cyclohexanone, which is derived from petroleum [2]. In addition to being inherently unsustainable, the chemical process to synthesize caprolactam requires toxic reagents and produces unwanted byproducts such as ammonium sulfate [2]. Multiple attempts have been made to produce caprolactam biologically, however the highest titers achieved to date are no greater than 1-2 mg/L [1,3]. All current published strategies to produce caprolactam rely on the cyclization of 6-aminocaproic acid (6ACA) via a promiscuous acyl-coA ligase [1,3]. While similar strategies to make C4 butyrolactam and C5 valerolactam produce gram per liter titers, it is thought that both the entropy and enthalpy properties of 7-membered ring formation of caprolactam present an inherent barrier to the cyclization of 6ACA [4,5]. Clearly, novel routes to renewable biological production of caprolactam are needed.

Genetically encoded biosensors can accelerate metabolic engineering efforts in many ways, the foremost of which is the ability to rapidly screen for desirable phenotypes beyond the throughput of analytical chemistry [6]. Multiple papers have reported transcription factors or riboswitches that respond to lactams with varying degrees of sensitivity and specificity [4,7,8]. One feature that unifies currently available lactam biosensors is that lactams are not the native ligand for any of the corresponding biosensor systems. Zhang et al. used ChnR from *Acinetobacter* sp. Strain NCIMB 9871 to sense multiple lactams with all ligands having a K_d of > 30 mM. However, ChnR natively regulates cyclohexanol catabolism and is activated by its natural ligand, cyclohexanone, at sub-millimolar concentrations [9,10]. Yeom et al. selected mutants of the NitR biosensor to sense caprolactam at concentrations as low as 100 μM when added exogenously, and leveraged this sensor to identify novel cyclases to convert 6ACA to caprolactam [8]. Natively, NitR regulates nitrile catabolism in *Rhodococcus rhodochrous* J1, and is responsive to micromolar concentrations of isovalernitrile [11,12]. Therefore, it is reasonable to assume that if

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

natural lactam catabolic pathways are identified, highly sensitive biosensors could also be found.

Recently, two groups have identified pathways of lactam degradation in both *P. putida* and *Pseudomonas jessenii* [13,14]. Work in *P. putida* demonstrated that the enzyme OplBA, putatively responsible for the hydrolysis of valerolactam, is upregulated by the lactam but not its cognate ω-amino acid [14]. These findings suggest there may be a lactam-sensitive transcription factor controlling the expression of the hydrolytic enzyme that can be used as a biosensor. In this work we demonstrate that the AraC-type regulator directly downstream of *oplBA* is indeed a lactam biosensor with unprecedented sensitivity towards both valerolactam and caprolactam. Through rational engineering we developed a suite of lactam sensing plasmids with dissociation constants ranging from 700 nM to 1.2 mM, allowing for a dramatic dynamic range of sensing. To demonstrate the utility of these sensors, we show that they are able to detect low titers of caprolactam produced biologically in an *Escherichia coli* system.

RESULTS

Identification and development of *oplR* as a lactam biosensor

In *P. putida*, the *oplBA* locus is flanked by the LysR-family regulator PP_3513 upstream, and the AraC-family regulator PP_3516 downstream. To infer if either of these transcription factors regulates *oplBA*, we used publicly available fitness data to assess if either regulator is cofit with *oplBA* (http://fit.genomics.lbl.gov) [15]. While no cofitness was observed between PP_3513 and either *oplA* or *oplB*, PP_3516 was highly cofit with both genes (0.91:PP_3514, 0.80:PP_3515). To examine the hypothesis that PP_3516 is the regulator of *oplBA*, we examined the genomic contexts of the oxoprolinase loci across multiple bacteria (Figure 1). While both regulators were found in closely related species, in more distantly related *Pseudomonads*, such as

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Pseudomonas aeruginosa, only the AraC-family regulator is conserved (Figure 1). Using Multiple EM for Motif Elicitation (MEME) [16] we attempted to identify conserved putative binding sites upstream of *oplBA* as well as PP_3516. In closely related *Pseudomonads*, including *P. aeruginosa*, a conserved motif was identified upstream of both *oplBA* as well PP_3516 (Figure 1). Attempts to confirm this as the binding site of PP_3516 were hampered by the insolubility of PP_3516 when expressed heterologously (Figure S1), a common issue with AraC-family proteins [17,18].

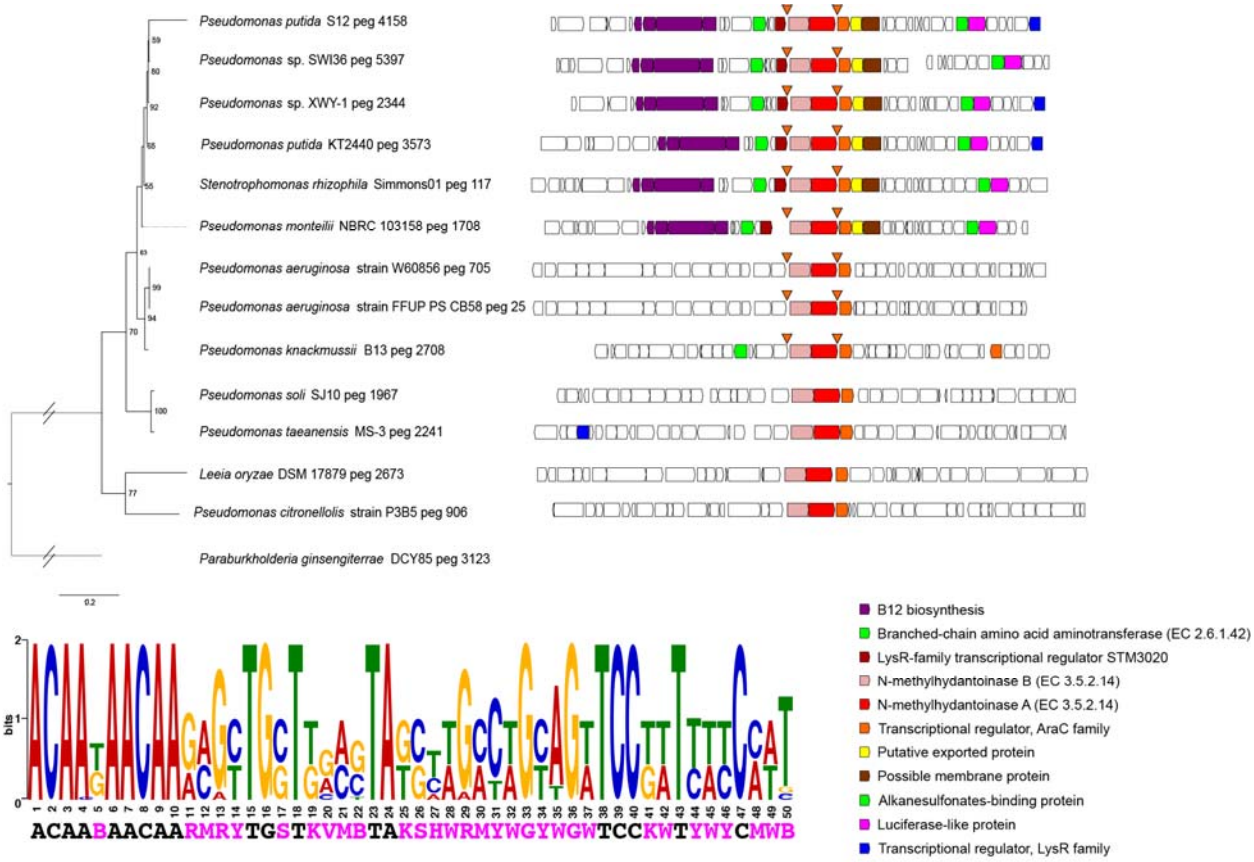


Figure 1 - Synteny analysis of *oplBA* homologs across genomes of related species. Triangles show the location of conserved putative binding sites of OplR. A consensus putative OplR binding site is shown below.

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

In order to screen the ability of PP_3516 to sense lactams, we employed a two-plasmid test system wherein PP_3516 was cloned into an arabinose inducible medium-copy p15a plasmid and the 200-bp upstream of *oplB* (Figure 2A) was cloned upstream of RFP on a compatible medium-copy pBBR plasmid (Figure 2B). The relationship between fluorescence output, *oplR* expression, and ligand induction was tested via a checkerboard assay where the levels of arabinose and valerolactam were varied independently of one another in cultures of *E. coli* that harbored both the "reporter" and "regulator" plasmids. A dose-dependent expression of RFP was observed; both arabinose and valerolactam were required for high-level expression of RFP (Figure 2C). Our initial screen also showed that the fluorescence was far above background RFP expression even at the lowest concentration of valerolactam tested (10 μM).

To quantify the sensing properties of PP_3516 the regulator was induced with a fixed concentration of arabinose at 0.0125% w/v and concentrations of either valerolactam or caprolactam were varied from 1 mM to 12 nM (Figure 2D). PP_3516 proved to be extremely sensitive to both caprolactam and valerolactam, with both ligands having a K_d ~5 uM and limits of detection ≤ 12 nM (Table 1). Based on these findings we propose PP_3516 be named *oplR* for oxoprolinase regulator, which encodes a biosensor ~5000x more sensitive towards caprolactam than the next most sensitive published biosensor [8].

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

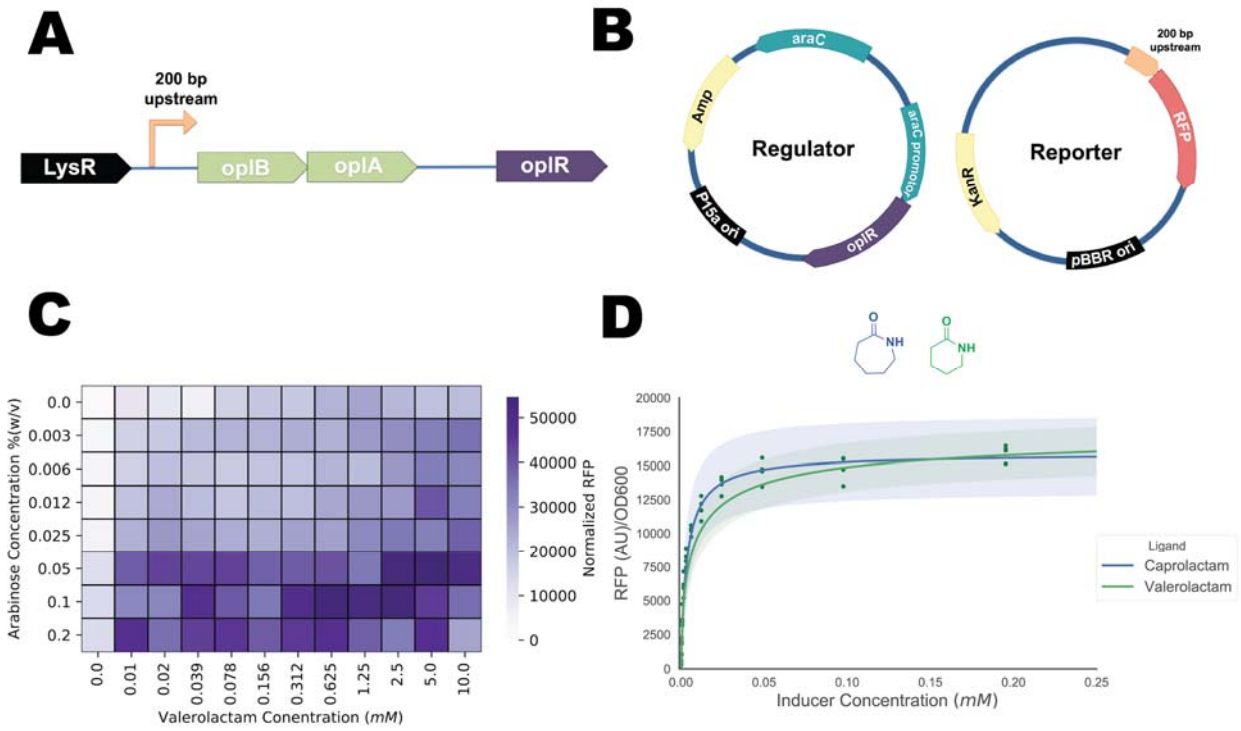


Figure 2 - Development of an *oplR*-based lactam biosensor. A) Operonic structure of *oplBA* relative to *oplR* and putative promoter region used to construct the reporter. B) Diagram of the two-plasmid system used to test OplR lactam sensing C) Checkerboard screen of OplR biosensor two-plasmid system. Y-axis shows the concentration of arabinose (%w/v), X-axis shows the concentration of valerolactam (mM). Colorbar to right shows fluorescent intensity normalized to OD_{600}. D) Fluorescence data fit to the Hill equation to derive biosensor performance characteristics for valerolactam and caprolactam. Points represent individual measurements. Shaded area represents (+/-) one standard deviation, n=4.

Table 1 - Two-plasmid biosensor parameters with caprolactam and valerolactam as ligands. Max: Predicted maximal RFP, Hill Coef: Predicted Hill coefficient, K_d: Predicted K_d in mM. LoD: Limit of detection determined experimentally. Standard deviation estimates are in parentheses.

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Ligand	**Max**	**Hill Coef**	**K_d (mM)**	**LoD (mM)**
Valerolactam	17817 (414)	0.62 (0.05)	0.007 (0.001)	≤ 0.000012
Caprolactam	15943 (483)	0.94 (0.16)	0.004 (0.001)	≤ 0.000012

Development of One Plasmid Systems

As a two-plasmid system is not convenient for engineering biological systems, we then sought to consolidate both the reporter and regulator into a single vector. Initial screening of OplR in the checkerboard assay suggested that varying the level of expression of OplR could dramatically influence the resulting sensing properties of the system (Figure 2C). We therefore constructed a family of plasmids, pLACSENS, where *oplR* was constitutively expressed from five promoters of increasing strength divergent from the RFP reporter (Figure 3A).

Biosensor performance of pLACSENS vectors was then assessed using valerolactam as a ligand across concentrations from 12.5 mM to 12.5 nM Figure (Figure 3B). As seen in the two-plasmid system, varying the strength of *oplR* expression dramatically changed the characteristics of the biosensor (Table 2). The most sensitive vector, pLACSENS3, had an experimentally determined limit of detection (LoD) of ≤12 nM, and a K_d of 700 nM. The least sensitive vector, pLACSENS5, had a limit of detection of 1.5 μM, and a K_d of 1.5 mM, but drives the expression of *oplR* with the strongest predicted promoter. The maximal RFP expression also varied greatly with *oplR* expression, with the highest and lowest RFP expression observed in pLACSENS3 and pLACSENS5, respectively (26300 vs. 793 RFP (AU)/OD_{600}). Induction over background expression was also highly variable; pLACSENS1 and pLACSENS4 were both maximally induced at ~250x over background, while pLACSENS5 was only induced ~25x over background.

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Given the wide range of biosensing parameters within the pLACSENS vectors, we sought to characterize which ligand concentration ranges each vector is most suited to detect. To do this, we utilized a recently developed model to probabilistically relate inducer concentration and fluorescence data via Markov Chain Monte Carlo (MCMC) sampling [19]. A resolution window is defined as the concentrations of inducer that are statistically compatible with the fluorescence data fit to the Hill function at a 95% confidence interval (cI). Resolution windows for each pLACSENS plasmid were graphed from their experimentally determined LoD to ligand concentrations compatible with 75% maximal fluorescence (Figure 3C). Overall, the family of vectors showed resolution from 5 nM to 1 mM valerolactam, with pLACSENS3 having the highest resolution at the lowest concentrations and pLACSENS5 having the best resolution at high concentrations (Figure 3C).

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

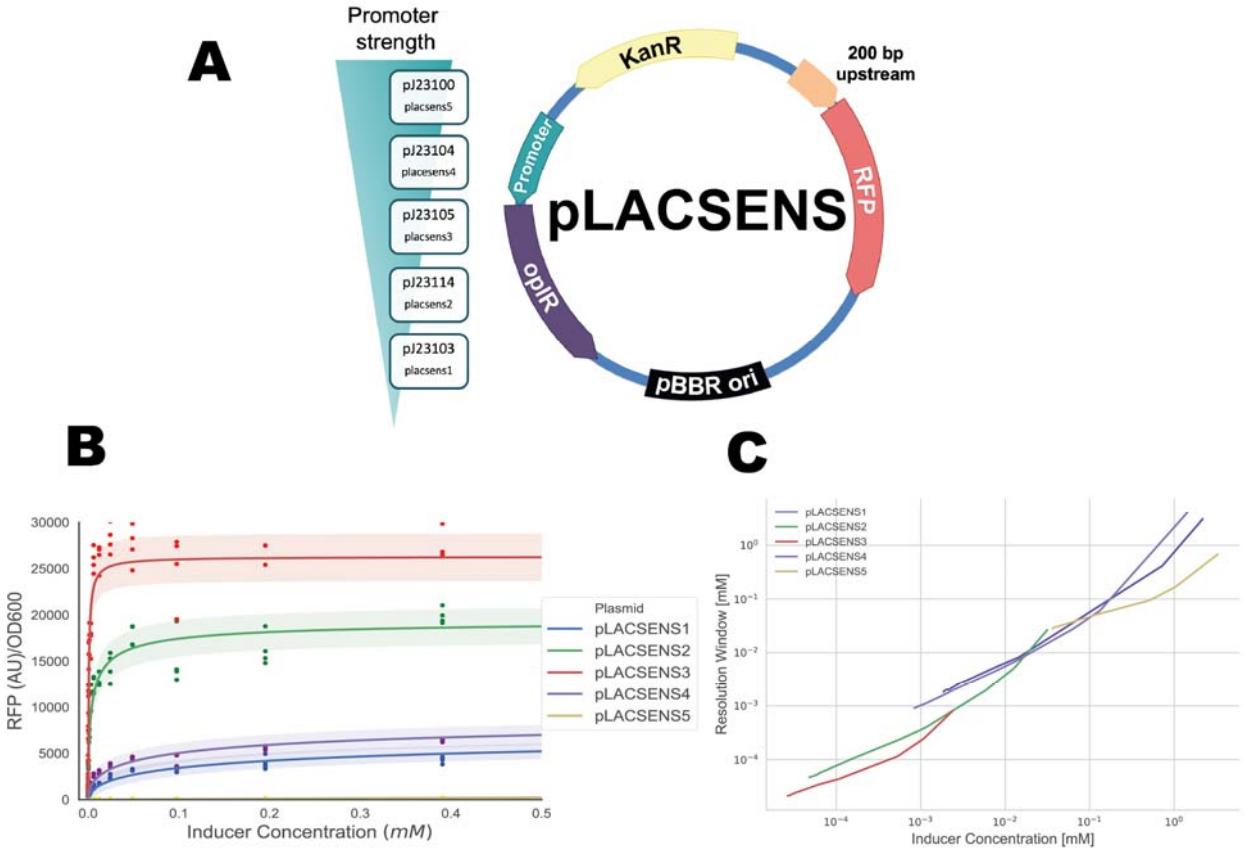


Figure 3 - Development of pLACSENS lactam biosensors A) Diagram of the pLACSENS vector design, with the relative predicted strength of the promoter driving *oplR* on the left B) Fluorescence data fit to the Hill equation to derive biosensor performance characteristics for valerolactam against all 5 pLACSENS vectors. Points represent individual measurements. Shaded area represents (+/-) one standard deviation, n=4. C) Resolution window of each pLACSENS vector over a range of valerolactam concentrations.

Table 2 - One-plasmid biosensor parameters: Max: Predicted maximal RFP, Hill Coef: Predicted Hill coefficient, K_d: Predicted K_d in mM, Exp. LoD: Limit of detection determined experimentally. Induction: maximal induction over background based on experimental data.

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Plasmid	**Max**	**Hill Coef**	**K_d (mM)**	**LoD (mM)**	**Induction**
pLACSENS1	8284 (358)	0.54 (0.05)	0.19 (0.05)	0.000024	253x
pLACSENS2	19614 (419)	0.63 (0.05)	0.004 (0.0006)	≤ 0.000012	93x
pLACSENS3	26300 (390)	0.90 (0.07)	0.0007 (0.0001)	≤ 0.000012	38x
pLACSENS4	9877 (421)	0.52 (0.05)	0.09 (0.02)	0.000024	256x
pLACSENS5	793 (25)	0.96 (0.10)	1.23 (0.15)	0.0015	25x

<u>OplR is selective for valerolactam and caprolactam</u>

To assess lactam specificity of OplR, we measured fluorescence induction of pLACSENS3 in the presence of lactams (laurolactam, caprolactam, valerolactam, butyrolactam, 5-oxoproline), ω-amino acids (4-aminobutyrate, 5AVA, 6ACA), the lactone valerolactone, and piperidine (Figure 4). Robust fluorescence induction was observed with caprolactam, valerolactam, and 5AVA (Table 3), but no other tested chemicals were capable of induction at the concentrations tested in the work (data not shown).

Given the dissimilarity in chemical structure of 5AVA and lactams, the ability of pLACSENS3 to detect 5AVA was surprising. Previously it has been shown in *E. coli* that 5AVA can spontaneously be converted into valerolactam by the activity of native acyl-coA ligases [1]. This led us to believe that rather than detecting 5AVA, pLACSENS3 was detecting valerolactam derived from the added 5AVA. LC-TOF analysis revealed that when 1 mM of 5AVA was fed to

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

E. coli cultures harboring pLACSENS3 ~600 nM of valerolactam was produced, while no valerolactam could be detected in *E. coli* not supplemented with 5AVA (Figure S2). Valerolactam concentration at this level would explain the high levels of fluorescence observed when 5AVA is added to cultures.

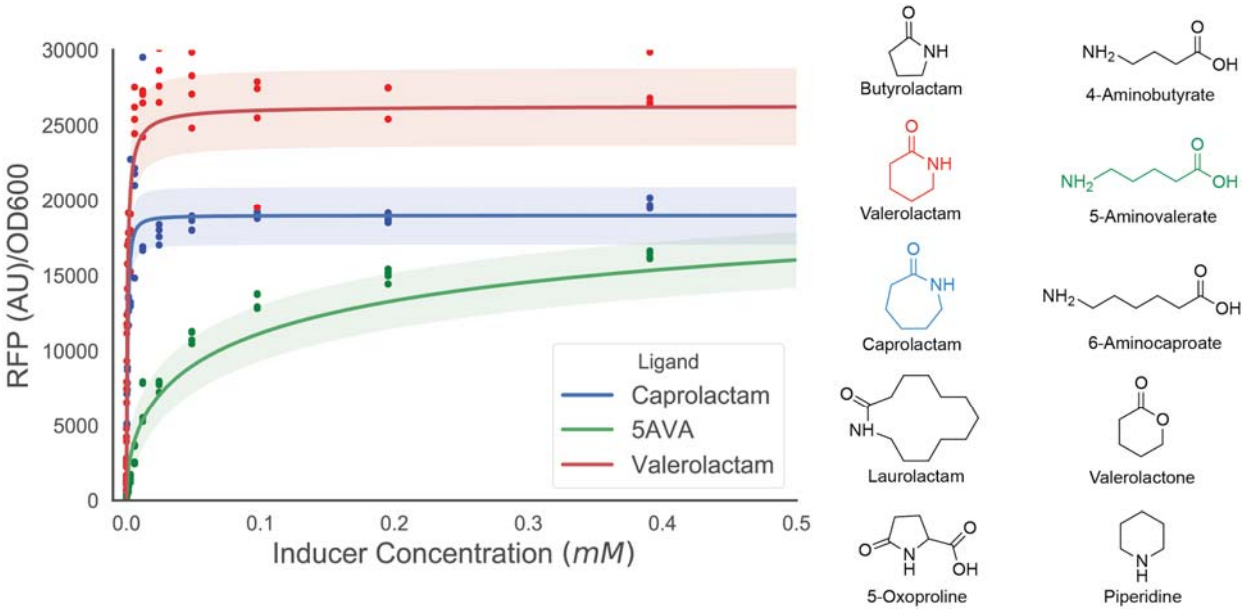


Figure 4 - Ligand range of pLACSENS3. Fluorescence data fit to the Hill equation to derive biosensor performance characteristics for ligands that activated pLACSENS3. Points represent individual measurements. Shaded area represents (+/-) one standard deviation, n=4. To the right, chemical structures of ligands that were tested.

Ligand	**Max**	**Hill Coef**	**K_d (mM)**
Valerolactam	26300 (390)	0.90 (0.07)	0.0007 (0.0001)
Caprolactam	18974 (272)	1.36 (0.14)	0.0009 (0.0001)

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

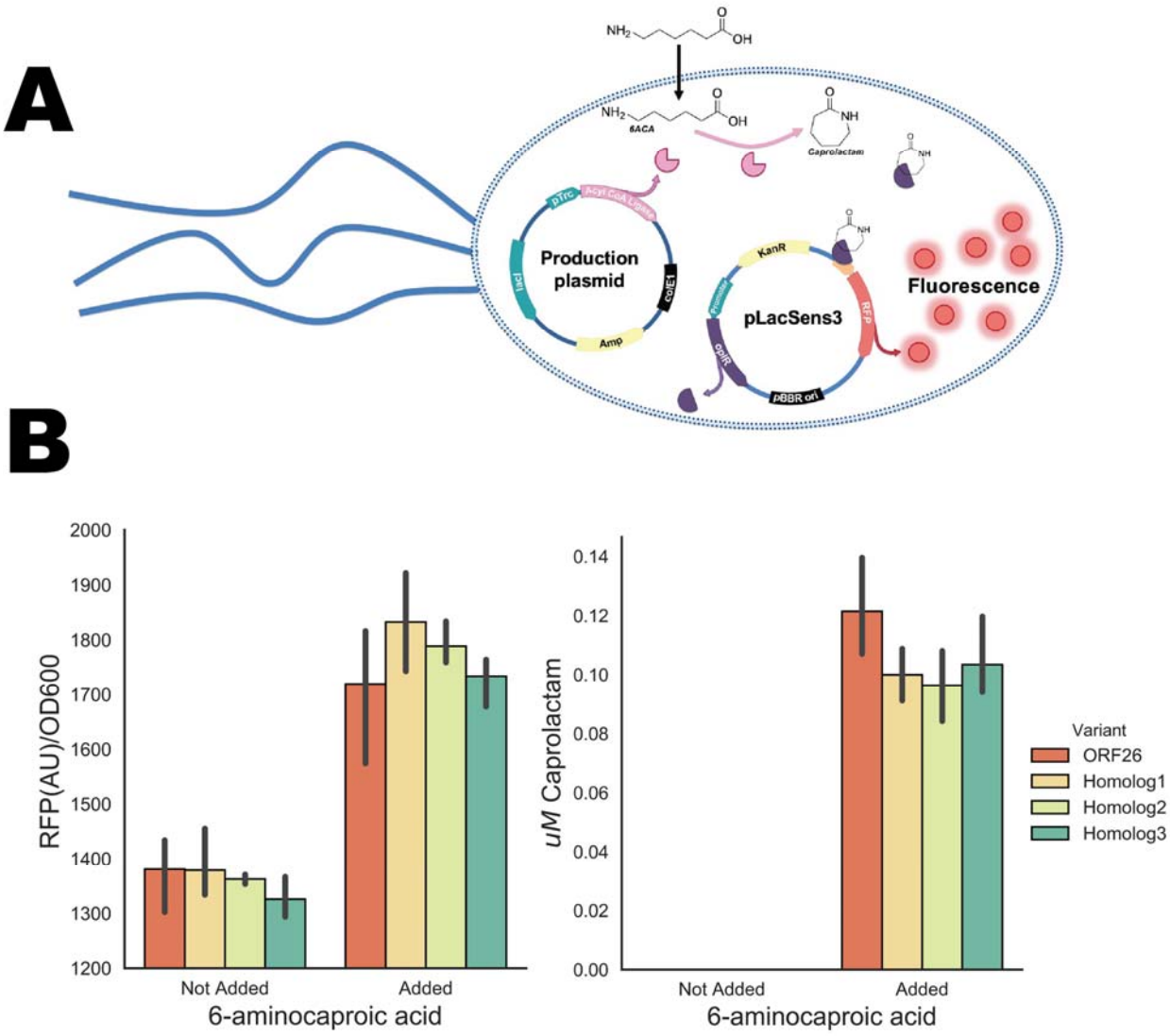
5AVA	23766 (1093)	0.52 (0.05)	0.12 (0.04)

Table 3 - pLACSENS3 biosensor parameters against different ligands: Max: Predicted maximal RFP, Hill Coef: Predicted Hill coefficient, K_d: Predicted K_d in mM.

Detection of caprolactam production *in vivo*

To demonstrate the utility of *oplR* based systems for metabolic engineering applications we introduced our most sensitive pLACSENS plasmid (pLACSENS3) into *E. coli* harboring various acyl-coA ligases on an IPTG-inducible orthogonal plasmid (Figure 5A). Multiple reports have utilized acyl-coA ligases to cyclize exogenously added 6ACA to produce low titers of caprolactam, with production ranging from 0.8 - 2 mg/L [1,3]. Strains harboring both plasmids were grown in LB medium with or without 10 mM 6ACA added for 24 hours. Cells grown in the presence of 6ACA demonstrated fluorescence greater than cells grown without 6ACA (Figure 5B). LC-MS analysis confirmed that no caprolactam was produced in cells grown without 6ACA, while cells grown with 6ACA had produced 11 ug/L (Figure 6B). The ability of pLACSENS3 to detect such minute production validates its utility as a means to rapidly and accurately screen novel pathways for caprolactam production.

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



222

223 **Figure 5 - Detection of caprolactam production in *E. coli* via pLACSENS3. A) Diagram of**

224 **the pLACSENS vector design, with the relative predicted strength of the promoter driving**

225 ***oplR* on the left B) Fluorescence data fit to the Hill equation to derive biosensor**

226 **performance characteristics for valerolactam against all five pLACSENS vectors. Shaded**

227 **area represents (+/-) one standard deviation, n=4. Different acyl-CoA ligases were tested**

228 **for the *in vivo* cyclization of 6ACA to caprolactam. These included the previously studied**

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

ORF26 CoA-ligase (red) [1] **and three additional homologs. Homolog 1 (yellow), Homolog 2 (light green) and homolog 3 (dark green).**

DISCUSSION

Previously, we have leveraged shotgun proteomics to infer local regulation within the lysine catabolism of *P. putida* [14,19–21]. From these data, multiple glutarate biosensors were engineered and used to measure relative metabolite amounts in the native host [19]. Here we again leveraged previously published proteomics data that showed OplBA to be specifically upregulated in the presence of valerolactam to develop a biosensor [14]. The AraC-family regulator tentatively named OplR was shown to have limits of detection for exogenously added valerolactam or caprolactam ≤ 12 nM when expressed at particular levels. This is remarkably more sensitive than the previously published valerolactam and caprolactam sensors, which showed limits of detection between 10-100 μM [8]. We attribute this high degree of sensitivity to the fact that, to our knowledge, it is the first sensor to control the catabolism of lactams specifically.

The range of detection of either caprolactam or valerolactam was highly dependent on the expression of *oplR*, with the K_d varying from 700 nM to 1.23 mM depending on the constitutive promoter used to drive expression. Previous work has also demonstrated that the sensing parameters of transcription factors can be readily modulated by changing the strength of transcription factor expression [22]. At both the highest and lowest levels of predicted expression OplR was less sensitive to lactams than when *oplR* was expressed more moderately. This may be explained by the fact that some AraC-family regulators are known to act as both positive and negative regulators, thus overexpression of OplR could result in hyper-repression [18].

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Unfortunately, insoluble expression of OplR prevented further examination of the biochemical means of transcriptional control.

OplR was shown to be highly specific for valerolactam and caprolactam as ligands, but not butyrolactam or laurolactam. These findings are consistent with previous observations that showed *oplBA* mutants were not defective in their ability to hydrolyze butyrolactam. The inability of *oplR* to sense the annotated substrate of OplBA, 5-oxoproline, suggests that the natural function of the amidohydrolase is not that of a 5-oxoprolinase. While the 5-membered lactam rings tested here were not able to induce OplR-mediated expression, further work should be conducted to test derivatives of valerolactam and caprolactam. Additional functional groups added to these lactams could be used to produce both pharmaceutical precursors [23] and polymers with novel nylon properties. Identifying biological routes to their synthesis is highly attractive. Furthermore, recent work has shown that directed evolution may also be applied to broaden the ligand range [24], which could allow OplR to accommodate different lactam ligands.

Current published routes to caprolactam biosynthesis rely on the cyclization of 6ACA via the activity of promiscuous CoA-ligase activity [1,3]. While *oplR*-based biosensors may be able to aid in the selection of mutant acyl-coA ligases with enhanced activity, there remains a sizeable thermodynamic barrier for the cyclization of the 7-membered ring [1]. Novel routes to caprolactam or naturally occurring caprolactam-containing natural products that can mitigate this barrier would be ideal for high-level production. For example, a better understanding of pestalactam A-C biosynthesis may provide new and more efficient chemoenzymatic routes to 7-ring cyclization [25].

In addition to the value of OplR as a biosensor for metabolic engineering purposes, it may also be useful as an inducible system. Valerolactam is inexpensive (~$2/gram), and highly

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

water soluble (291 mg/mL). The vector pLACSENS2 demonstrated ~100x induction over background, with a K_d toward valerolactam of 4 uM, and the second highest maximal expression of any single vector tested. Additional engineering of the system could improve upon these qualities. Furthermore, as OplR works well in *E. coli* and is derived from the distantly related *P. putida*, it may work well in other bacterial systems. Future work could evaluate which hosts are suitable for this inducible system.

Since naturally-occuring genetically encoded biosensors for chemicals of interest have the potential to be much more sensitive than those repurposed or evolved in the laboratory, it is critical to pursue rapid and efficient means of identifying them. The recent development of other high-throughput methods to associate genotypes with phenotypes, such as RB-TnSeq and CRISPRi, has created a large reservoir of data that can be easily mined for transcription factors useful in synthetic biology [20,26–28]. Bacteria often locally regulate catabolism, thus allowing inference of genetic control by adjacent transcription factors once a catabolic pathway has been discovered. Empirical evidence of catabolism is critical for assigning transcription factor function as orthologous transcription is often utilized differently by different species [29]. Future work to generalize approaches to develop useful synthetic biology tools from genome-wide fitness data has the potential to dramatically increase the genetically encoded chemical sensor space.

METHODS

Media, chemicals, and culture conditions

General *E. coli* cultures were grown in Lysogeny Broth (LB) Miller medium (BD Biosciences, USA) at 37 °C. When indicated, *E. coli* was also grown on EZ-RICH medium

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(Teknova, Hollister, CA) supplemented with 1% glucose. Cultures were supplemented with kanamycin (50 mg/L, Sigma Aldrich, USA), or carbenicillin (100mg/L, Sigma Aldrich, USA), when indicated. All other compounds were purchased through Sigma Aldrich (Sigma Aldrich, USA) .

Strains and plasmids

All bacterial strains and plasmids used in this work are listed in Table 4. All strains and plasmids created in this work are available through the public instance of the JBEI registry. (https://public-registry.jbei.org/folders/XXX). All plasmids were designed using Device Editor and Vector Editor software, while all primers used for the construction of plasmids were designed using j5 software [30–32]. Plasmids were assembled via Gibson Assembly using standard protocols [33], or Golden Gate Assembly using standard protocols [34]. Plasmids were routinely isolated using the Qiaprep Spin Miniprep kit (Qiagen, USA), and all primers were purchased from Integrated DNA Technologies (IDT, Coralville, IA).

Expression and purification of proteins

Proteins were purified as described previously [35]. The cultures were grown at 37 °C until the OD_{600} nm reached 0.8 and cooled on ice for 20 min. 1 mM IPTG was added to induce overexpression for 16 h at 18 °C. The cells were harvested by centrifugation (8000g, 10 min, 4 °C), resuspended in 40 mL of lysis buffer (50 mM HEPES, pH 8.0, 0.3 M NaCl, 10% glycerol (v/v) and 10 mM imidazole), and lysed by sonication on ice. Cellular debris was removed by centrifugation (20000g, 60 min, 4 °C).

The supernatant was applied to a fritted column containing Ni-NTA resin (Qiagen, USA) and the proteins were purified using the manufacturer's instructions. Fractions were collected and analyzed via SDS-PAGE.

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Fluorescence biosensor assays

All assays were conducted in 96-deep well plates (Corning Costar, 3960), with each well containing 500 μL of medium with appropriate ligands, antibiotics, and/or inducers inoculated at 1% v/v from overnight cultures. Plates were sealed with AeraSeal film (Excel Scientific, AC1201-02 and incubated at 37 C in a 250 rpm shaker rack. After 24 hours, 100 μL from each well was aliquoted into a black, clear-bottom 96-well plate for measurements of optical density and fluorescence using an Infinite F200 (Tecan Life Sciences, San Jose, CA) plate reader. Optical density was measured at 600 nm (OD_{600}), while fluorescence was measured using an excitation wavelength of 535 nm, an emission wavelength of 620 nm, and a manually set gain of 60.

For the checkerboard assay of the two-plasmid system, LB medium supplemented with both kanamycin and carbenicillin was inoculated with E. coli containing the regulator and reporter plasmids grown overnight in the same medium. Arabinose concentration was decreased from 0.2 to 0% w/v along the y-axis, while valerolactam concentration was increased from 0-10 mM along the x-axis.

To find the Hill fit to the two-plasmid system, EZ-RICH medium containing kanamycin, carbenicillin, and 0.0125 w/v% arabinose was inoculated with an overnight culture of the two-plasmid system in LB medium supplemented with both kanamycin and carbenicillin. Both valerolactam and caprolactam were tested at concentrations ranging from 0 to 50 mM.

Characterization of the 5 variations of the one-plasmid pLACSENS plasmids was conducted using EZ-rich medium containing kanamycin and inoculated with overnight cultures of the appropriate E. coli strain. Valerolactam concentrations was varied from 0 to 50 mM. This same assay was repeated on pLACSENS3 with various ligands—using concentrations between 0

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

to 50 mM of caprolactam and 5AVA, and 0 to 10 mM of butyrolactam, 5-oxoproline, gamma-aminobutyric acid, 6ACA, valerolactone, and piperidine. The maximum concentration of laurolactam tested was 0.1 mM due to its poor solubility.

Production assays and analytical methods

Caprolactam production assays were carried out in 10 mL of LB medium supplemented with 10 mM 6ACA, as well as kanamycin, carbenicillin, and 1 mM IPTG. Cultures were inoculated 1:100 with overnight culture harboring both pLACSENS3 and expression vectors for acyl-coA ligases and grown at 30 °C shaking at 200 rpm for 24 hours. After 24 hours optical density was measured at 600 nm (OD_{600}), while fluorescence was measured using an excitation wavelength of 535 nm, an emission wavelength of 620 nm, and a manually set gain of 60. To sample for caprolactam production 200 µL of culture was quenched with an equal volume of ice-cold methanol and then stored at -20 °C until analysis. The different acyl-CoA ligases tested for the *in vivo* cyclization of 6ACA to caprolactam included the previously studied ORF26 CoA-ligase (red) [1] and three additional homologs. Homolog 1 refers to A0A0D4DX08_9ACTN, homolog 2 refers to G2NX2_STRVO and homolog 3 refers to A0A0C1VDH3_9ACTN.

Valerolactam and caprolactam were measured via LC-QTOF-MS as described previously [14]. Liquid chromatographic separation was conducted at 20°C with a Kinetex HILIC column (50-mm length, 4.6-mm internal diameter, 2.6-µm particle size; Phenomenex, Torrance, CA) using a 1260 Series HPLC system (Agilent Technologies, Santa Clara, CA, USA). The injection volume for each measurement was 5 µL. The mobile phase was composed of 10 mM ammonium formate and 0.07% formic acid in water (solvent A) and 10 mM ammonium formate and 0.07% formic acid in 90% acetonitrile and 10% water (solvent B) (HPLC grade, Honeywell

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Burdick & Jackson, CA, USA). High purity ammonium formate and formic acid (98-100% chemical purity) were purchased from Sigma-Aldrich, St. Louis, MO, USA. Lactams were separated with the following gradient: decreased from 90%B to 70%B in 2 min, held at 70%B for 0.75 min, decreased from 70%B to 40%B in 0.25 min, held at 40%B for 1.25 min, increased from 40%B to 90%B for 0.25 min, held at 90%B for 1 min. The flow rate was varied as follows: 0.6 mL/min for 3.25 min, increased from 0.6 mL/min to 1 mL/min in 0.25 min, and held at 1 mL/min for 2 min. The total run time was 5.5 min.

The HPLC system was coupled to an Agilent Technologies 6520 quadrupole time-of-flight mass spectrometer (QTOF MS) with a 1:6 post-column split. Nitrogen gas was used as both the nebulizing and drying gas to facilitate the production of gas-phase ions. The drying and nebulizing gases were set to 12 L/min and 30 lb/in^2, respectively, and a drying gas temperature of 350°C was used throughout. Fragmentor, skimmer and OCT 1 RF voltages were set to 100 V, 50 V and 300 V, respectively. Electrospray ionization (ESI) was conducted in the positive-ion mode for the detection of $[M + H]^+$ ions with a capillary voltage of 4000 V. The collision energy voltage was set to

0 V. MS experiments were carried out in the full-scan mode (75–1100 *m/z*) at 0.86 spectra/s. The QTOF-MS system was tuned with the Agilent ESI-L Low concentration tuning mix in the range of 50-1700 *m/z*. Lactams were quantified by comparison with 8-point calibration curves of authentic chemical standards from 0.78125 µM to 100 µM. R^2 coefficients of ≥0.99[EB1] were achieved for the calibration curves. Data acquisition was performed by Agilent MassHunter Workstation (version B.05.00), qualitative assessment by Agilent MassHunter Qualitative Analysis (version B.05.00 or B.06.00), and data curation by Agilent Profinder (version B.08.00)

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Bioinformatic Analysis

For the phylogenetic reconstructions, the best amino acid substitution model was selected using ModelFinder as implemented on IQ-tree [36] phylogenetic trees were constructed using IQ-tree, nodes were supported with 10,000 bootstrap replicates. The final tree figures were edited using FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). Orthologous syntenic regions of OplBA were identified with CORASON-BGC [37] and manually colored and annotated. DNA-binding sites were predicted with MEME [16].

Analysis of Biosensor Performance

Fluorescent measurements were fit to the Hill equation and biosensor parameters we estimated as described previously [19]. Briefly, fluorescent values were OD_{600} normalized, and had their background fluorescent subtracted. A probabilistic model relating inducer concentrations (C) and fluorescence measurements (F) to characterize the performance of a biosensor was used where F :

$$P(F|C,\theta,\alpha) = \mathcal{N}(h_\theta(C),\sigma)$$

where is the Hill function, are the parameters of the hill function, is the estimated standard deviation, and N represents a gaussian (normal) distribution. Using the probabilistic model which captures our constraints on the problem the log likelihood function is expressed as:

$$\ell(\theta,\sigma|\mathcal{D}) = \sum_{i=1}^{N} \log P(F = f_i|C = c_i,\sigma,\theta).$$

The log likelihood is used to express the maximum likelihood estimation (MLE) problem:

$$\hat{\theta},\hat{\sigma} = \arg\max_{\theta,\sigma} \ell(\theta,\sigma|\mathcal{D})$$

which when solved results in the optimal parameters of the model given the characterization data. In order to estimate the distribution of ligand concentrations that are compatible with

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

experimental fluorescence data, MCMC sampling was used to solve the following MLE problem:

$$\hat{c} = \arg\max_{c} \sum_{i=1}^{N} \log P(F = f_i | C = c, \hat{\theta}, \hat{\sigma})$$

We determined biosensor resolution by solving the above maximum likelihood estimation problem iteratively over the range of observed fluorescences during the biosensor characterization process. This can determine the relationship between an inducer concentration estimate and the estimated standard deviation. The standard deviation of the estimate of inducer concentration can be interpreted as the resolution window. Here, two standard deviations is considered the resolution window of the sensor, as 95% of the compatible inducer concentration estimates fall within the interval .

Induction above background was calculated by dividing the maximal experimental normalized RFP expression by background fluorescence in uninduced cultures. The experimental limit of detection was defined as the minimal concentration of inducer that produced normalized fluorescence that was statistically above uninduced cultures harboring the same plasmid via Student's t-test ($p<0.05$) .

Acknowledgements

We would like to thank the Koret Research Scholars Program for providing funding to ANP to conduct summer research. This work was part of the DOE Joint BioEnergy Institute (https://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, supported by the U.S. Department of Energy, Energy Efficiency and Renewable Energy, Bioenergy Technologies Office, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Energy. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan). HGM was also supported by the Basque Government through the BERC 2018-2021 program and by Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2017-0718.

Contributions

Conceptualization, M.G.T.; Methodology, M.G.T.,A.N.P.,J.F.B, P.C.M.,E.E.K.B.,N.S,Z.C.; Investigation, M.G.T., L.E.V., J.F.B, P.C.M, M.R.I.,M.E.G.,A.N.P., ; Writing – Original Draft, M.G.T.; Writing – Review and Editing, All authors.; Resources and supervision, H.G.M,A.M.,J.D.K.

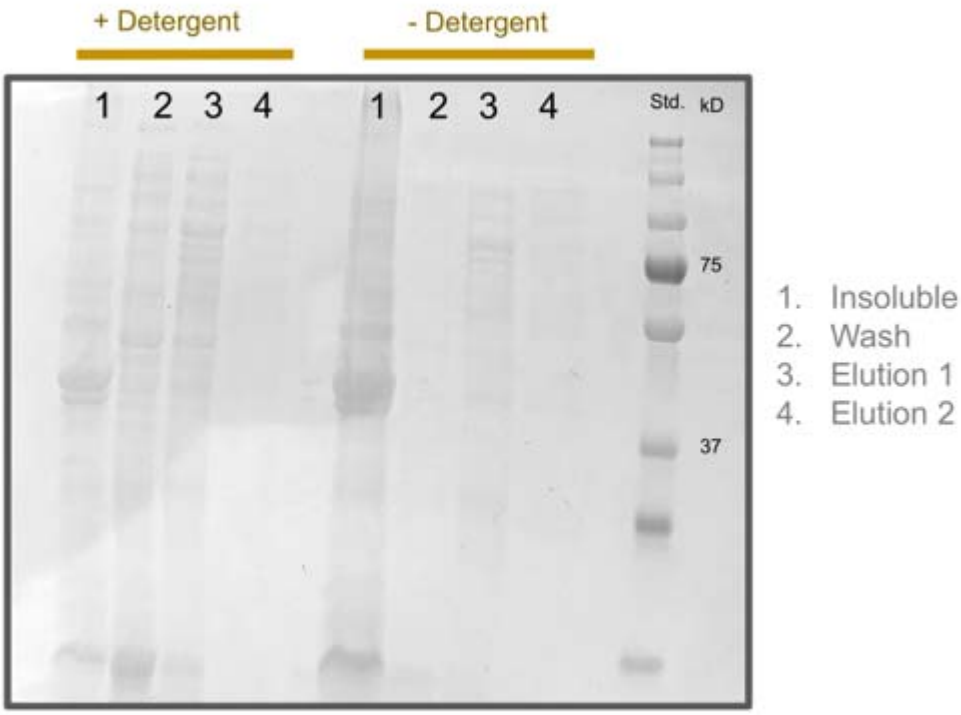
bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Competing Interests

J.D.K. has financial interests in Amyris, Lygos, Constructive Biology, Demetrix, Napigen and Maple Bio.

Supplementary Figures

Figure S1: Insolubility of OplR expressed heterologously in *E. coli*: OplR-6xHis expressed in the insoluble fraction (~40 kD) of *E. coli* in the presence and absence of the detergent Tween 20 (0.5% w/v).



bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S2: Valerolactam produced from 5AVA in *E. coli* culture. Bars show the amount of valerolactam produced in *E. coli* harboring pLACSENS3 when fed either 1 mM or 0 mM 5AVA. Error bars represent 95% cI.

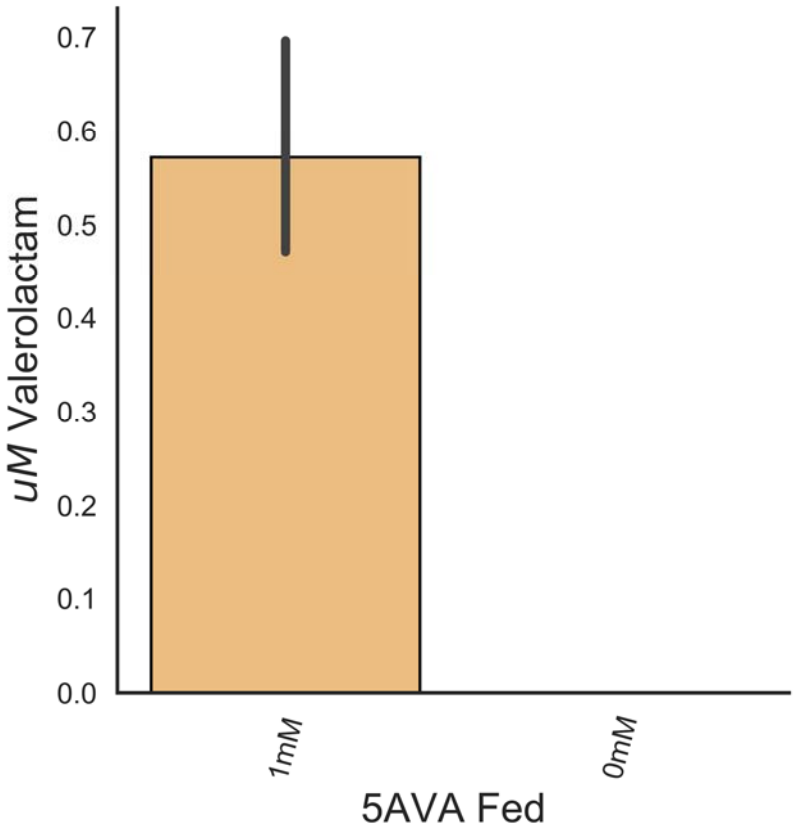


Table 4: Strains and plasmids used in this study

Strain	JBEI Part ID	Reference
E. coli DH10B		38

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

E. coli BL21(DE3)		Novagen
P. putida KT2440		ATCC 47054
Plasmids		
pBADT		[39]
pBADT-PP_3516p-RFP	JBEI-104519	This work
pBbA8a-PP_3516	JBEI-104517	This work
pLacSens1	JBEI-104506	This work
pLacSens2	JBEI-104505	This work
pLacSens3	JBEI-104504	This work
pLacSens4	JBEI-104503	This work
pLacSens5	JBEI-104502	This work
pBbE7a-ORF26	JBEI-104322	This work
pBbE7a-G2NX2_STRVO	JBEI-104323	This work
pBbE7a-A0A0D4DX08_9ACTN	JBEI-104324	This work
pBbE7a-A0A0C1VDH3_9ACTN	JBEI-104325	This work

Bibliography

(1) Zhang, J., Barajas, J. F., Burdu, M., Wang, G., Baidoo, E. E., and Keasling, J. D. (2017) Application of an Acyl-CoA Ligase from Streptomyces aizunensis for Lactam Biosynthesis. *ACS Synth. Biol. 6*, 884–890.

(2) Thomas, J. M., and Raja, R. (2005) Design of a “green” one-step catalytic production of epsilon-caprolactam (precursor of nylon-6). *Proc Natl Acad Sci USA 102*, 13732–13736.

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(3) Chae, T. U., Ko, Y.-S., Hwang, K.-S., and Lee, S. Y. (2017) Metabolic engineering of Escherichia coli for the production of four-, five- and six-carbon lactams. *Metab. Eng. 41*, 82–91.

(4) Zhang, J., Barajas, J. F., Burdu, M., Ruegg, T. L., Dias, B., and Keasling, J. D. (2017) Development of a Transcription Factor-Based Lactam Biosensor. *ACS Synth. Biol. 6*, 439–445.

(5) Casadei, M. A., Galli, C., and Mandolini, L. (1984) Ring-closure reactions. 22. Kinetics of cyclization of diethyl (.omega.-bromoalkyl)malonates in the range of 4- to 21-membered rings. Role of ring strain. *J. Am. Chem. Soc. 106*, 1051–1056.

(6) Dietrich, J. A., McKee, A. E., and Keasling, J. D. (2010) High-throughput metabolic engineering: advances in small-molecule screening and selection. *Annu. Rev. Biochem. 79*, 563–590.

(7) Jang, S., Jang, S., Im, D.-K., Kang, T. J., Oh, M.-K., and Jung, G. Y. (2019) Artificial caprolactam-specific riboswitch as an intracellular metabolite sensor. *ACS Synth. Biol.*

(8) Yeom, S.-J., Kim, M., Kwon, K. K., Fu, Y., Rha, E., Park, S.-H., Lee, H., Kim, H., Lee, D.-H., Kim, D.-M., and Lee, S.-G. (2018) A synthetic microbial biosensor for high-throughput screening of lactam biocatalysts. *Nat. Commun. 9*, 5053.

(9) Iwaki, H., Hasegawa, Y., Teraoka, M., Tokuyama, T., Bergeron, H., and Lau, P. C. (1999) Identification of a transcriptional activator (ChnR) and a 6-oxohexanoate dehydrogenase (ChnE) in the cyclohexanol catabolic pathway in Acinetobacter sp. Strain NCIMB 9871 and localization of the genes that encode them. *Appl. Environ. Microbiol. 65*, 5158–5162.

(10) Steigedal, M., and Valla, S. (2008) The Acinetobacter sp. chnB promoter together with its cognate positive regulator ChnR is an attractive new candidate for metabolic engineering applications in bacteria. *Metab. Eng. 10*, 121–129.

(11) Komeda, H., Hori, Y., Kobayashi, M., and Shimizu, S. (1996) Transcriptional regulation of the Rhodococcus rhodochrous J1 nitA gene encoding a nitrilase. *Proc Natl Acad Sci USA 93*, 10572–10577.

(12) Pandey, A. K., Raman, S., Proff, R., Joshi, S., Kang, C.-M., Rubin, E. J., Husson, R. N., and Sassetti, C. M. (2009) Nitrile-inducible gene expression in mycobacteria. *Tuberculosis (Edinb) 89*, 12–16.

(13) Otzen, M., Palacio, C., and Janssen, D. B. (2018) Characterization of the caprolactam degradation pathway in Pseudomonas jessenii using mass spectrometry-based proteomics. *Appl. Microbiol. Biotechnol. 102*, 6699–6711.

(14) Thompson, M., Valencia, L. E., Blake-Hedges, J., Cruz-Morales, P., Velasquez, A., Pearson, A., Sermeno, L., Sharpless, W., Benites, V., Chen, Y., Baidoo, E., Petzold, C. J., Deutschbauer, A., and Keasling, J. D. (2019) Host engineering for improved valerolactam production in Pseudomonas putida. *BioRxiv.*

(15) Wetmore, K. M., Price, M. N., Waters, R. J., Lamson, J. S., He, J., Hoover, C. A., Blow, M. J., Bristow, J., Butland, G., Arkin, A. P., and Deutschbauer, A. (2015) Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. *MBio 6*, e00306-15.

(16) Bailey, T. L., Williams, N., Misleh, C., and Li, W. W. (2006) MEME: discovering and

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

analyzing DNA and protein sequence motifs. *Nucleic Acids Res. 34*, W369-73.
(17) Porter, M. E., and Dorman, C. J. (2002) In vivo DNA-binding and oligomerization properties of the Shigella flexneri AraC-like transcriptional regulator VirF as identified by random and site-specific mutagenesis. *J. Bacteriol. 184*, 531–539.
(18) Schleif, R. (2010) AraC protein, regulation of the l-arabinose operon in Escherichia coli, and the light switch mechanism of AraC action. *FEMS Microbiol. Rev. 34*, 779–796.
(19) Thompson, M. G., Cruz-Morales, P., Krishna, R. N., Blake-Hedges, J. M., Incha, M. R., and Keasling, J. D. (2019) Glutarate metabolism in Pseudomonas putida is regulated by two distinct glutarate sensing transcription factors. *BioRxiv*.
(20) Thompson, M. G., Blake-Hedges, J. M., Cruz-Morales, P., Barajas, J. F., Curran, S. C., Eiben, C. B., Harris, N. C., Benites, V. T., Gin, J. W., Sharpless, W. A., Twigg, F. F., Skyrud, W., Krishna, R. N., Henrique Pereira, J., Baidoo, E. E. K., Petzold, C. J., Adams, P. D., Arkin, A. P., Deutschbauer, A. M., and Keasling, J. D. (2019) Massively Parallel Fitness Profiling Reveals Multiple Novel Enzymes in Pseudomonas putida Lysine Metabolism. *MBio 10*.
(21) Chen, Y., Vu, J., Thompson, M. G., Sharpless, W. A., Chan, L. J. G., Gin, J. W., Keasling, J. D., Adams, P. D., and Petzold, C. J. (2019) A rapid methods development workflow for high-throughput quantitative proteomic applications. *PLoS ONE 14*, e0211582.
(22) Rogers, J. K., Guzman, C. D., Taylor, N. D., Raman, S., Anderson, K., and Church, G. M. (2015) Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Res. 43*, 7648–7660.
(23) Thai, K., Wang, L., Dudding, T., Bilodeau, F., and Gravel, M. (2010) NHC-catalyzed intramolecular redox amidation for the synthesis of functionalized lactams. *Org. Lett. 12*, 5708–5711.
(24) Kasey, C. M., Zerrad, M., Li, Y., Cropp, T. A., and Williams, G. J. (2018) Development of Transcription Factor-Based Designer Macrolide Biosensors for Metabolic Engineering and Synthetic Biology. *ACS Synth. Biol. 7*, 227–239.
(25) Davis, R. A., Carroll, A. R., Andrews, K. T., Boyle, G. M., Tran, T. L., Healy, P. C., Kalaitzis, J. A., and Shivas, R. G. (2010) Pestalactams A-C: novel caprolactams from the endophytic fungus Pestalotiopsis sp. *Org. Biomol. Chem. 8*, 1785–1790.
(26) Wang, T., Guan, C., Guo, J., Liu, B., Wu, Y., Xie, Z., Zhang, C., and Xing, X.-H. (2018) Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nat. Commun. 9*, 2475.
(27) Price, M. N., Wetmore, K. M., Waters, R. J., Callaghan, M., Ray, J., Liu, H., Kuehl, J. V., Melnyk, R. A., Lamson, J. S., Suh, Y., Carlson, H. K., Esquivel, Z., Sadeeshkumar, H., Chakraborty, R., Zane, G. M., Rubin, B. E., Wall, J. D., Visel, A., Bristow, J., Blow, M. J., and Deutschbauer, A. M. (2018) Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature 557*, 503–509.
(28) Sagawa, S., Price, M. N., Deutschbauer, A. M., and Arkin, A. P. (2017) Validating regulatory predictions from diverse bacteria with mutant fitness data. *PLoS ONE 12*, e0178258.
(29) Price, M. N., Dehal, P. S., and Arkin, A. P. (2007) Orthologous transcription factors in

bioRxiv preprint doi: https://doi.org/10.1101/700484; this version posted July 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

bacteria have different functions and regulate different genes. *PLoS Comput. Biol. 3*, 1739–1750.

(30) Ham, T. S., Dmytriv, Z., Plahar, H., Chen, J., Hillson, N. J., and Keasling, J. D. (2012) Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. *Nucleic Acids Res. 40*, e141.

(31) Chen, J., Densmore, D., Ham, T. S., Keasling, J. D., and Hillson, N. J. (2012) DeviceEditor visual biological CAD canvas. *J. Biol. Eng. 6*, 1.

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

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

(34) Engler, C., Kandzia, R., and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE 3*, e3647.

(35) Jackson, D. R., Tu, S. S., Nguyen, M., Barajas, J. F., Schaub, A. J., Krug, D., Pistorius, D., Luo, R., Müller, R., and Tsai, S.-C. (2016) Structural Insights into Anthranilate Priming during Type II Polyketide Biosynthesis. *ACS Chem. Biol. 11*, 95–103.

(36) Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods 14*, 587–589.

(37) Cruz-Morales, P., Ramos-Aboites, H. E., Licona-Cassani, C., Selem-Mójica, N., Mejía-Ponce, P. M., Souza-Saldívar, V., and Barona-Gómez, F. (2017) Actinobacteria phylogenomics, selective isolation from an iron oligotrophic environment and siderophore functional characterization, unveil new desferrioxamine traits. *FEMS Microbiol. Ecol. 93*.

(38) Grant, S. G., Jessee, J., Bloom, F. R., and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proc Natl Acad Sci USA 87*, 4645–4649.

(39) Bi, C., Su, P., Müller, J., Yeh, Y.-C., Chhabra, S. R., Beller, H. R., Singer, S. W., and Hillson, N. J. (2013) Development of a broad-host synthetic biology toolbox for Ralstonia eutropha and its application to engineering hydrocarbon biofuel production. *Microb. Cell Fact. 12*, 107.