A synthetic microbial biosensor for high-throughput screening of lactam biocatalysts

Soo-Jin Yeom¹, Moonjeong Kim¹,², Kil Koang Kwon¹, Yaoyao Fu¹, Eugene Rha¹, Sung-Hyun Park¹,³, Hyewon Lee¹, Haseong Kim¹,³, Dae-Hee Lee⁰¹,³, Dong-Myung Kim² & Seung-Goo Lee¹,³

Biocatalytic cyclization is highly desirable for efficient synthesis of biologically derived chemical substances, such as the commodity chemicals ε-caprolactam and δ-valerolactam. To identify biocatalysts in lactam biosynthesis, we develop a caprolactam-detecting genetic enzyme screening system (CL-GESS). The *Alcaligenes faecalis* regulatory protein NitR is adopted for the highly specific detection of lactam compounds against lactam biosynthetic intermediates. We further systematically optimize the genetic components of the CL-GESS to enhance sensitivity, achieving 10-fold improvement. Using this highly sensitive GESS, we screen marine metagenomes and find an enzyme that cyclizes ω-amino fatty acids to lactam. Moreover, we determine the X-ray crystal structure and catalytic residues based on mutational analysis of the cyclase. The cyclase is also used as a helper enzyme to sense intracellular ω-amino fatty acids. We expect this simple and accurate biosensor to have wide-ranging applications in rapid screening of new lactam-synthesizing enzymes and metabolic engineering for lactam bio-production.
identifying biocatalysts, in particular, enzymes or pathways that play key roles in biosynthesis of non-native molecules of interest, is important to the industrial synthesis of chemical products. The discovery and engineering of enzymes or pathways involved in the synthesis of a desired product are often limited by a lack of sufficiently sensitive and rapid screening tools for the identification of candidate genes from large natural or synthetic gene libraries. Genetically encoded biosensors have untapped potential as tools for screening enzymes and pathways; thus, extensive efforts have been made to develop high-throughput screening (HTS) biosensors equipped with fluorescence-based genetic circuit devices. A key component in such devices is the ligand-inducible transcription factor (TF). In nature, a wide variety of TFs can specifically recognize small molecules and alter gene transcription at their targeted promoters. Numerous TF-based biosensors with ligand specificity and dynamic detection ranges are already available for sensing various small molecules. Such biosensors are straightforward and powerful tools for detecting target molecules or their intermediates in investigations of enzyme and biosynthetic pathways. Recently, many efforts using TF-based sensors in HTS have focused on altering TF specificities or sensitivities. For instance, AraC of *Escherichia coli*, which originally senses arabinose, was engineered to detect mevalonate—a key intermediate of isoprenoid pathways—which enabled the screening of cells with increased mevalonate synthesis.

Lactams are industrially important chemicals that are used in polyamide production and ε-caprolactam biosynthesis. δ-Valerolactam and ε-caprolactam (CL) are converted from ω-amino fatty acids (5-aminovaleric acid (5-AVA) or 6-aminoacapric acid (6-ACA)) and are used as precursors for the production of nylon-5, nylon-6, and nylon-6,5, which are used as precursors for the production of nylon-5, nylon-6, and nylon-6,5, which are used to manufacture tire cords, carpeting, plastics, and food-packaging materials. ε-Caprolactam is most widely used to produce nylon-6 and is mainly produced through Beckmann rearrangement of the cyclohexanone oxime in the presence of fuming sulfuric acid, at 90–120°C. Biorenewable routes towards ε-caprolactam from fermentation-derived lysine, muconic acid, adipic acid, and 6-ACA have been discussed. Additionally, the production of 6-ACA by direct fermentation of glucose has also been reported.

Although several ω-amino fatty acids are biologically produced through biosynthetic pathways, complete biosynthetic pathways capable of synthesizing lactams are mostly unknown owing to the lack of enzymes able to catalyze the last ring-cyclization step. To date, only two enzymes that can be used for lactam biosynthesis have been reported to carry out this step: *Candida antarctica* lipase B (CALB) and *Streptomyces aizunensis* acyl-CoA ligase (ACL). However, CALB requires an anhydrous condition, high temperature, and long reaction time; thus, it is not suitable for lactam biosynthesis. ACL exhibits a broad substrate spectrum and has been used for cyclizing 4-aminobutyric acid, 5-AVA, and 6-ACA into γ-butyrolactam, δ-valerolactam, and ε-caprolactam, respectively. Both 6-ACA and its cyclized form, ε-caprolactam, are non-natural compounds; a previous report designed and proposed two biosynthetic pathways for fermentative production of 6-ACA. However, there is a lack of efficient and specific enzymes capable of producing ε-caprolactam from 6-ACA through enzymatic conversion and microbial fermentation.

In this study, we aim to investigate the lactam biosynthesis pathway in greater detail through the identification of certain enzymes involved, using HTS biosensors. To this end, we firstly designed and engineered a lactam-detecting biosensor, termed caprolactam-detectable genetic enzyme-screening system (CL-GESS), and then carried out HTS of ε-caprolactam-converting cyclases from diverse metagenomes. To improve the signal-to-noise ratio and sensing sensitivity, a transcriptional regulator, NitR, is engineered and used in the CL-GESS to identify a cyclase for ε-caprolactam or valerolactam biosynthesis from 6-ACA or 5-AVA, respectively. Finally, we determined the X-ray crystal structure of the newly identified cyclase to provide insight into its cyclization activities. Using the cyclase, we developed a genetically encoded biosensor to sense ω-amino fatty acids (5-ACA or 6-AVA). We presented a cyclase that converts 6-ACA to ε-caprolactam, which will open opportunities for the development of a bioprocess to produce lactams and nylons.

**Results**

**Design and construction of a lactam biosensor.** The experimental strategy used to develop the CL-GESS is illustrated in Fig. 1. Our strategy included (a) engineering of a transcriptional regulator by fluorescence-activated cell sorting (FACS), (b) combinatorial analysis and optimization of a promoter or ribosomal-binding site (RBS) for controlling a regulator, (c) searching for a binding site on the regulator, and (d) optimization of a reporter gene. To develop a biosensor that responds to ε-caprolactam, we selected lactam-responsive TFs in nature and integrated them into a genetic circuit. In the actinomycete *Rhodococcus rhodochrous*, nitR is used in the industrial production of acrlyamide and nicotinamide, nitrase is strongly induced in the presence of ε-caprolactam. In *R. rhodochrous*, nitrase is encoded by the nitA gene in the nitrile degradation operon, which is positively regulated by NitR, which is in turn activated by isovaleronitrile or ε-caprolactam. As *Alcaligenes faecalis* JM3 also contains a nitrile degradation operon, we examined the applicability of the *A. faecalis* nitR regulatory subunit (nitrile degradation operon) in the development of the CL-GESS by assessing the induction of nitA gene expression in the presence of various effector molecules. NitR responded solely to ε-caprolactam, which activated nitA gene expression (Supplementary Fig. 1). Based on this result, we transformed the nitR (E. coli codon-optimized) regulatory system of *A. faecalis* in the nitrite degradation operon and the nitA promoter (P

**Optimization and characterization of the CL-GESS.** We next optimized the CL-GESS to enhance its sensitivity. We constructed five different CL-GESSs in this study (Supplementary Fig. 2). The first, CL-GESSv1, contained the E. coli codon-optimized nitR gene under the control of a constitutive J23100 promoter in the direction opposite to that of transcription of the putative P

---

**NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-07488-0**

**ARTICLE NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-07488-0**
The putative NitR-binding site was palindromically located in the –35 and untranslated regions of \( P_{nitA} \) (Supplementary Fig. 3b, c). It is important to note that sequences at positions –35 (TTTCATC) and –10 (TACACT) upstream of the transcription start site were similar to the actinomycete (mainly streptomycete) consensus promoter sequence \( \text{TTGAC(A/G)-17 bp- TAg(A/G)} \). In addition, a single transcription initiation site was identified 74-bp upstream of the first ATG codon of GFP (Supplementary Fig. 3d). Deletion of NitR from the CL-GEES plasmid resulted in complete loss of \( P_{nitA} \)-induced green fluorescence (Supplementary Fig. 4), indicating that NitR activates the transcription of the target gene at its targeted promoter \( P_{nitA} \).

To further optimize NitR expression in CL-GEESv3, which contained sfGFP and truncated \( P_{nitA} \), we substituted the promoter and RBS with various synthetic promoters and RBSs of different strengths (rank order of promoter strength: J23100>J23106>J23114; rank order of RBS strength: B0030>B0034, and T7RBS) (http://parts.igem.org/Promoters/Catalog/Anderson, Supplementary Table 1). The plasmids CL-GEES J23114-B0034 (CL-GEESv4) showed the highest fold change in fluorescence in the presence of ε-caprolactam (Fig. 2b).

To quantitatively assess the response of CL-GEESv4 to ε-caprolactam, we measured fluorescence at the single-cell level in \( E. \) coli DH5α cells harboring CL-GEESv4 in Luria-Bertani (LB) medium, at various ε-caprolactam concentrations (up to 30 mM). Fluorescence was observed only in the presence of ε-caprolactam (Fig. 2c). The minimal concentration of ε-caprolactam required to activate NitR was 500 μM; moreover, a tight correlation was observed between fluorescence intensity and ε-caprolactam concentration. To examine the ligand specificity of CL-GEESv4, we applied it to various substrate and precursor molecules involved in lactam biosynthesis. CL-GEESv4 sensed ε-caprolactam, ε-caprolactone, cyclohexanone, N-acetylcaprolactam, δ-valerolactam, δ-valerolactone, benzonitrile, and isovaleronitrile, whereas detected none of the intermediates in the proposed ε-caprolactam biosynthesis pathway (Fig. 2d, Supplementary Fig. 7a). Even at high concentrations (30 mM), intermediates, such as l-lysine, 5-AVA, and 6-ACA did not induce sfGFP expression above the background level. To investigate the relationship between the intracellular ε-caprolactam level and sensor output, we measured intracellular ε-caprolactam. Intracellular ε-caprolactam was 0.035–0.27 μmole/mg wet cells when we added 1–30 mM caprolactam extracellularly (Supplementary Table 2). As the concentration of external ε-caprolactam increased, the intracellular ε-caprolactam concentration also increased (Supplementary Table 2). When we washed the cells twice or three times with 50 ml of saline, intracellular ε-caprolactam was not detected. These observations supported that the uptake and release of ε-caprolactam appear to be equilibrated by passive intracellular transport.

In addition, we evaluated the applicability of CL-GEESv4 to various Gram-negative bacteria, including \( Pseudomonas putida \) strains KT2440 and S12, and \( Ralstonia eutropha \). The backbone
vector of CL-GESSv4 was replaced with pBBRBB or pSEVA (http://seva.cnb.csic.es) containing RK2 or pBBR1 ori, respectively. *P. putida* S12 containing pBBRBB-based CL-GESSv4 showed a three-fold increase in signal in the presence of ε-caprolactam, whereas S12 transformed with pSEVA-CL-GESSv4 exhibited a 1.5-fold increase (Fig. 2e). *P. putida* KT2440 and *R. eutropha* harboring pBBRBB-based or pSEVA-based CL-GESS showed no fluorescence response to ε-caprolactam (Fig. 2e). These results show the possibility that CL-GESS system can be applied in *P. putida* with various advantages for natural product biosynthesis, such as a versatile intrinsic metabolism with diverse enzymatic capacities, and outstanding tolerance to xenobiotics.

For these reasons, the genetic circuit-based sensor in *P. putida* should be further studied.

**Engineer the NitR TF.** To improve the sensitivity of CL-GESSv4 further, we engineered the NitR transcriptional regulator. To this end, we designed a single cell-based screening strategy for a library of NitR with random mutations. We selected hits with a high fluorescence signal in the presence of ε-caprolactam and removed false positives that exhibited fluorescence in the absence of ε-caprolactam. After three rounds of screening, the highest 0.5% of phenotypes were selected as putative hits and their fluorescence signals were verified against ε-caprolactam in liquid LB medium (Supplementary Fig. 5a). Among them, a NitR double mutant, CL-GESSNitR-L117F/P133S, showed increased fluorescence signal when exposed to ε-caprolactam at a low concentration of 50 μM as minimum while wild-type CL-GESS showed increased fluorescence signal more than 500 μM (Fig. 3a).

We further engineered the double mutant (CL-GESSNitR-L117F/P133S) to maximize the sensor sensitivity. To find a mutation to generate NitR mutants more sensitive to ε-caprolactam, we firstly tested the P133 location. However, when we changed the P133 to serine, there was no change in sensitivity of CL-GESS to ε-caprolactam, but the background fluorescence was increased. The single-point mutant CL-GESSNitR-L117F/P133S showed increased fluorescence signal when exposed to ε-caprolactam at a low concentration of 50 μM as minimum while wild-type CL-GESS showed increased fluorescence signal more than 500 μM (Fig. 3a).

We further engineered the double mutant (CL-GESSNitR-L117F/P133S) to maximize the sensor sensitivity. To find a mutation to generate NitR mutants more sensitive to ε-caprolactam, we firstly tested the P133 location. However, when we changed the P133 to serine, there was no change in sensitivity of CL-GESS to ε-caprolactam, but the background fluorescence was increased. The single-point mutant CL-GESSNitR-L117F/P133S showed increased fluorescence signal when exposed to ε-caprolactam at a low concentration of 50 μM as minimum while wild-type CL-GESS showed increased fluorescence signal more than 500 μM (Fig. 3a).

We further engineered the double mutant (CL-GESSNitR-L117F/P133S) to maximize the sensor sensitivity. To find a mutation to generate NitR mutants more sensitive to ε-caprolactam, we firstly tested the P133 location. However, when we changed the P133 to serine, there was no change in sensitivity of CL-GESS to ε-caprolactam, but the background fluorescence was increased. The single-point mutant CL-GESSNitR-L117F/P133S showed increased fluorescence signal when exposed to ε-caprolactam at a low concentration of 50 μM as minimum while wild-type CL-GESS showed increased fluorescence signal more than 500 μM (Fig. 3a).

**Engineering the NitR TF.** To improve the sensitivity of CL-GESSv4 further, we engineered the NitR transcriptional regulator. To this end, we designed a single cell-based screening strategy for a library of NitR with random mutations. We selected hits with a high fluorescence signal in the presence of ε-caprolactam and removed false positives that exhibited fluorescence in the absence of ε-caprolactam. After three rounds of screening, the highest 0.5% of phenotypes were selected as putative hits and their fluorescence signals were verified against ε-caprolactam in liquid LB medium (Supplementary Fig. 5a). Among them, a NitR double mutant, CL-GESSNitR-L117F/P133S, showed increased fluorescence signal when exposed to ε-caprolactam at a low concentration of 50 μM as minimum while wild-type CL-GESS showed increased fluorescence signal more than 500 μM (Fig. 3a).

We further engineered the double mutant (CL-GESSNitR-L117F/P133S) to maximize the sensor sensitivity. To find a mutation to generate NitR mutants more sensitive to ε-caprolactam, we firstly tested the P133 location. However, when we changed the P133 to serine, there was no change in sensitivity of CL-GESS to ε-caprolactam, but the background fluorescence was increased. The single-point mutant CL-GESSNitR-L117F/P133S showed increased fluorescence signal when exposed to ε-caprolactam at a low concentration of 50 μM as minimum while wild-type CL-GESS showed increased fluorescence signal more than 500 μM (Fig. 3a).
ε-caprolactam (Fig. 3b). Interestingly, the substitution of L117 with aromatic residues, as in CL-GESS\textsubscript{NitR-L117F}, CL-GESS\textsubscript{NitR-L117Y}, and CL-GESS\textsubscript{NitR-L117W}, induced a marked increase in the fluorescence signal of 11.4-fold, 5.6-fold, and 4.9-fold, respectively, in the presence of ε-caprolactam as compared to the 3-fold increase observed for wild-type CL-GESS (Fig. 3b). CL-GESS\textsubscript{NitR-L117F} also showed enhanced fluorescence for cyclohexanone, δ-valerolactone, ε-caprolactone, ε-caprolactam, and N-acetylcaprolactam (Fig. 3c). CL-GESS\textsubscript{NitR-L117F} but not wild-type CL-GESS, was able to detect δ-valerolactone, δ-valerolactam, and N-acetylcaprolactam (Fig. 3c). CL-GESS\textsubscript{NitR-L117} variants (L117F, L117Y, and L117W) also yielded strong signals in the presence of various lactones or lactams rather than nitrile compounds (Supplementary Fig. 6). These results suggest that bulky aromatic residues in position 117 may interact with lactam or lactone substrates. To determine the amino acid residues of NitR that are important for the interaction with ε-caprolactam, a homology model of \textit{A. faecalis} NitR was constructed based on the crystal structure of Vibrio cholerae O395 ToxT (Protein Data Bank (PDB) entry 3GBG)\textsuperscript{22} as the closest sequence among the known structures of AraC-type regulators. Although the sequence identity between NitR and ToxT was relatively low (18.1% identity, 38.2% similarity), the sequence identity (23% identity, 43% similarity) of the N-terminal domain (substrate-binding site) between two regulators can result in useful homology model. Molecular docking was attempted to estimate the substrate interactions in the \textit{A. faecalis} NitR active site (Fig. 3d). Amino acid L117 of NitR was estimated to be located on one side of the active pocket (Fig. 3d) and the mutations to aromatic amino acids was estimated to result in a shorter distance between a docked ε-caprolactam and the aromatic side chains in mutant NitR (Fig. 3d). These results suggest that an aromatic amino acid at position 117 closely interacts with lactam or lactone substrates in the active pocket. Although the L117W mutant with larger side chains appeared more sensitive to ε-caprolactam (Supplementary Fig. 6), it also showed a higher background signal. Therefore, we used L117F mutant for the screening of ε-caprolactam producing enzymes because high background of L117W mutant could be a problem during the screening of large library.

**HTS of a metagenome library using CL-GESS\textsubscript{NitR-L117F}**. A metagenome is a potentially substantial reservoir of valuable enzymes or biocatalysts that could be targeted by CL-GESS\textsubscript{NitR-L117F}. Various intermediates in the ε-caprolactam synthesis pathway, including l-lysine, 5-AVA, and ACA did not induce sfGFP expression of CL-GESS\textsubscript{NitR-L117F} in solid or liquid LB medium, suggesting that the highly sensitive CL-GESS\textsubscript{NitR-L117F} can be used for HTS of new enzymes in the ε-caprolactam biosynthesis pathway (Supplementary Fig. 7). We carried out CL-GESS-based HTS of a metagenomic fosmid library constructed using tidal flat sediments. To this end, we firstly transformed a plasmid encoding CL-GESS\textsubscript{NitR-L117F} into \textit{E. coli} cells to screen a cyclase that could convert 6-ACA into ε-caprolactam. Cells grown in LB medium containing 10 mM 6-ACA were sorted by flow cytometry. After removing false positives by culturing recovered cells in fresh LB medium without 6-ACA, a total of 10\textsuperscript{6} cells were
obtained (Supplementary Fig. 5). These cells were cultured in LB medium containing 10 mM 6-ACA and sorted again based on fluorescence intensity, yielding 27 hits showing stronger fluorescence on LB solid plate than the control. The 27 hits were grown on solid LB medium containing 6-ACA, and one colony (named fosmid #3), with the strongest fluorescence, was selected by microscopy (Fig. 4a). Flow-cytometric analysis revealed that fosmid #3 with CL-GESE<sub>NifR</sub>-L<sub>117F</sub> exhibited stronger fluorescence in the presence of 6-ACA than did cells harboring the empty fosmid vector (Fig. 4a, Supplementary Fig. 8). This suggested that fosmid #3 contained at least one open-reading frame (ORF) encoding an enzyme that catalyzed the biosynthesis of ε-caprolactam from 6-ACA. Therefore, we isolated DNA from fosmid #3 cells, and then completely sequenced it using the shotgun method.

The DNA from fosmid #3 contained a 31-kb DNA fragment that showed up to 88% identity to <i>Citrobacter</i> genome sequences (Supplementary Table 3). At least one ORF encoding a putative cyclase was identified, supporting the results of the functional screen. Of the 25 ORFs in clone #3, we selected 11 ORFs that were up to 500 bp (Supplementary Table 4) and subcloned them into the pET28a<sup>+</sup> plasmid. <i>E. coli</i> cells containing CL-GESE<sub>NifR</sub>-L<sub>117F</sub> were transformed with the pET28a<sup>+</sup> plasmid. E. coli cells containing CL-GESE<sub>NifR-L<sub>117F</sub></sub> were transformed with the pET28a<sup>+</sup> plasmid encoding an individual ORF to detect the fluorescence signal corresponding to the conversion of 6-ACA to ε-caprolactam. We found that an ORF encoding 3-hydroxybutyrate dehydrogenase from <i>Citrobacter freundii</i> (CF3HBD) was positive for this activity (number 2 in Fig. 4b).

Characterization of CF3HBD. To corroborate its cyclase activity, CF3HBD was expressed and purified as a 27-kDa protein (Supplementary Fig. 10a), and its cyclization activity was calibrated by detection of ε-caprolactam by liquid chromatography–mass spectrometry (LC–MS; Fig. 5a and Supplementary Fig. 9b) and nuclear magnetic resonance (NMR; Fig. 5b). Importantly, this reaction was performed using purified enzyme and in the absence of any cofactors, such as ATP and NADH. Furthermore, the reaction mixture was extracted with ethyl acetate to purify the reaction product from the other materials. LC–MS analysis revealed that the purity of the obtained caprolactam was >99% (Supplementary Fig. 9a). Thus, based on the product confirmed by NMR, LC–MS, and IR (infrared spectroscopy) analyses, we concluded that the promiscuous enzyme catalyzed the cyclization of 6-ACA in the absence of any cofactor (Supplementary Fig. 9a, c, d). CF3HBD showed 81%, 82%, 81%, 56%, and 28% sequence identity with 3HBD from <i>Serratia marcescens</i>, <i>Klebsiella pneumoniae</i>, <i>Enterobacter</i> sp., <i>Rhodobacter sphaeroides</i>, and <i>Pseudomonas lemoignei</i>, respectively (Supplementary Fig. 10b). We cloned two additional 3HBDs from <i>S. marcescens</i> (SM3HBD) and <i>Enterobacter</i> sp. (ES3HBD) and determined their cyclase activity after purification (Supplementary Fig. 10a) using 6-ACA as a substrate; both enzymes showed cyclization activity towards 6-ACA to produce ε-caprolactam as well (Supplementary Fig. 10c). However, the 3HBDs from <i>R. sphaeroides</i> and <i>P. lemoignei</i>, with low sequence identity, showed no cyclization activities.

Basically, the screened CF3HBDs belong to the NAD(P)H-dependent short-chain dehydrogenase family, and they have cyclization activity as unexpected promiscuous activity. 3-Hydroxybutyrate dehydrogenase catalyzed the reversible and stereospecific oxidation of 3-hydroxybutyric acid to acetoacetate using NAD<sup>+</sup> as a coenzyme<sup>23</sup>. Thus, we further investigated simple enzyme characteristics in terms of the effects of pH and temperature, metals, and coenzymes on enzyme activities. The maximum dehydrogenation activity of CF3HBD towards 3-hydroxybutyrate was observed at pH 8.0 and at 40 °C (Supplementary Fig. 11a, b). The maximum cyclization activity of CF3HBD towards 3-hydroxybutyric acid was observed at pH 7.5 and 40 °C (Supplementary Fig. 11a, b). The cyclization activity of CF3HBD did not require metals or cofactors in vitro (Supplementary Fig. 12a, b). The time course analyses for the production of ε-caprolactam were conducted using 1 mM 6-ACA and CF3HBD. The enzyme concentrations used were 0.1, 0.25, and 0.5 mg/ml and yielded molar conversions of 17.9%, 34.6%, and 47.3%, respectively (Supplementary Fig. 13). No detectable intermediate was obtained during this reaction. In addition, we found that CF3HBD could catalyze the degradation of the cyclic amide group in ε-caprolactam, generating 6-ACA as a reaction product (Supplementary Fig. 14).

Crystal structure of the CF3HBD and homology modeling. To further characterize the cyclization activity, we determined the crystal structure of CF3HBD in complex with NAD<sup>+</sup> at a
Determination of the active site residues by mutagenesis. To identify the catalytic residues involved in the dual activities of CF3HBD, we conducted a ligand-docking study of the linear form of 6-ACA with the crystal structure. The docking results suggested that 10 residues (Q91, S139, V140, H141, K149, Y152, Q193, W184, V190, and Q193) located within 4.0 Å of the center of the docked substrate are active-site residues. These residues were selected as candidate determinant residues for enzyme activity. The selected 10 residues were separately replaced with alanine or glutamate, and the wild-type and all mutant 3HBD were expressed and purified by His tag affinity chromatography as a single band with a molecular mass of ~27 kDa in sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (Supplementary Fig. 16). Alanine substitution of the catalytic residues, such as Q91, S139, and H141 completely abolished catalytic activity towards both 3-hydroxybutyric acid and 6-ACA, which suggested that dehydrogenation activity and cyclization activity share catalytic site (Table 2). Alanine substitution at V140, Y152, and W184 resulted in lack of activity towards 3-

**CL-GESS as a sensor for ω-amino fatty acid cyclization.** We investigated whether CL-GESS<sub>L117F</sub>can be used as a biosensor that reflects the amount of ε-caprolactam formed from 6-ACA as a fluorescence signal detectable by flow cytometry (Fig. 5c). We co-transformed *E. coli* EPI(DE3) cells with the CL-GESS<sub>L117F</sub> sensor and the other plasmid expressing CF3HBD under the gratitude.
control of the T7 promoter and lacI regulation. In cells harboring CL-GESS_{L117F} and CF3HBD which were termed ACA-GESS, strong fluorescence was observed in the presence of 5-AVA and 6-ACA (Supplementary Fig. 18). The CF3HBD enzyme can serve as a helper enzyme for substrate detection in the CL-GESS system.

Discussion

The identification of new enzymes or evolution of enzymes for desired activities is a crucial task for industrial biotechnology. TF-based genetic screening methods have received much attention as enzyme selection systems because they enable the rapid screening of libraries comprising innumerable genetic variants. We previously identified a phenolic compound with GESS, a reporter system that can be used to screen metagenomic or mutation libraries for enzymes. Additionally, HTS for the generation of new activity or directed evolution by TF-based biosensor has been reported. The TF-based genetic screening method could be useful to generate new activity that would alleviate the current burden on HTS. In this report, we described the design and validation of CL-GESS, a genetically encoded biosensor for the screening and engineering of lactam synthesis enzymes. The engineering of NitR as a TF for higher sensitivity for screening was successful and helpful to screen for a cyclase that has promiscuous activity. We believe that this tool will enable the scientific community to study lactam biosynthesis and further develop an economical bioprocess for lactam production.

NitR is related to the bacterial XylS/AraC family of TFs that are involved in metabolism of carbon sources and have been linked to pathogenesis, with sequence similarity at the carboxyl terminus containing a helix-turn-helix DNA-binding motif. Deletion of NitR resulted in the complete loss of P_{nitA}-induced green fluorescence (Supplementary Fig. 4), indicating that NitR activates the transcription of the target gene in the operon. Using engineered NitR in the CL-GESS allowed us to detect target molecules with low background signal levels, which is a desirable feature for HTS systems. Single-cell technologies based on high signal-to-noise ratios have been used to explore cellular mechanisms. We applied this strategy to the GESS, although there was a low background signal that was likely due to leaky reporter expression or some other factor. A recent study described the prediction of transcriptional noise based on intracellular physical distance between a regulator (XylS) and the target promoter (Pm). By random mutagenesis and computational modeling, we identified position 117 in the active pocket of NitR as a critical site for sensitive ligand detection in GESS (Fig. 3d). Wild-type NitR showed a slightly longer distance between a regulator (XylS) and the target promoter (Pm) than did the L117F mutant, suggesting that an aromatic amino acid with a bulky functional group at position 117 favors ε-caprolactam detection.

In this study, we applied CL-GESS to a metagenomic library derived from sea tidal flat sediments to identify enzymes that produce lactam from ω-amino fatty acids. The resulting isolate, CF3HBD, is a unique biocatalyst with lactam-synthesizing activity as promiscuous activity that does not require a cofactor or energy source in an in vitro system (Fig. 5a). Along with the CF3HBD, other 3HBD enzymes with high sequence similarity also exhibited the cyclization activity towards 6-ACA (Supplementary Fig. 10). Here, we report an enzyme that can produce ε-caprolactam from 6-ACA identified using metagenomic library screening rather than did the L117F mutant, suggesting that an aromatic amino acid with a bulky functional group at position 117 favors ε-caprolactam detection.
Table 1 Data collection and refinement statistics for CF3HBD–NAD\(^+\) complex

| CF3HBD–NAD\(^+\) | Data collection | | | | | Rwork = 19.6(26.9) | Rfree = 23.5(31.2) |
|---|---|---|---|---|---|---|---|
| Space group | P2\(_1\) | | | | | | |
| Cell dimensions (Å) | 62.5, 148.3, 62.4 | 90, 101.5, 90 | | | | | |
| Resolution (Å) | 50.2(2.29) | | | | | | |
| Rsym (%) | 11.8(89.1) | | | | | | |
| I/\(\alpha\) | 28.9(2.4) | | | | | | |
| Rmerge (%) | 12.8(96.8) | | | | | | |
| Complement (%) | 91.8(82.3) | | | | | | |
| Redundancy | 6.8(6.0) | | | | | | |

Table 2 Specific activities of wild-type CF3HBD and mutant enzymes of the active site residues

| 3-Hydroxybutyric acid | 6-Aminocaproic acid |
|---|---|
| Specific activity (U/mg) | Relative activity (%) | Specific activity (U/mg) | Relative activity (%) |
| WT | 79.2 ± 2.5 | 100.0 ± 3.2 | 2.1 ± 0.0 | 100.0 ± 0.3 |
| Q91A | - | - | - | - |
| Q91E | 90.9 ± 2.3 | 114.9 ± 2.9 | 2.9 ± 0.1 | 139.5 ± 6.4 |
| S139A | - | - | - | - |
| V140A | - | - | 1.7 ± 0.1 | 83.2 ± 4.3 |
| H141A | - | - | - | - |
| H141E | 0.9 ± 0.03 | 11.0 ± 0.04 | 1.2 ± 0.0 | 56.0 ± 0.7 |
| K149A | 7.1 ± 0.4 | 8.9 ± 0.5 | 1.8 ± 0.1 | 88.0 ± 4.8 |
| K149E | 23.5 ± 0.3 | 29.7 ± 0.4 | 1.6 ± 0.0 | 76.5 ± 5.0 |
| Y152A | - | - | 6.5 ± 0.5 | 312.9 ± 23.2 |
| Y152E | 28.5 ± 1.3 | 36.0 ± 1.6 | 2.1 ± 0.7 | 101.5 ± 3.5 |
| W184E | - | - | 1.7 ± 0.2 | 84.2 ± 7.5 |
| W184A | - | - | 1.8 ± 0.1 | 88.0 ± 4.0 |
| W189A | 16.4 ± 0.5 | 20.8 ± 0.7 | 2.2 ± 0.0 | 106.6 ± 1.3 |
| V190A | 77.7 ± 3.6 | 98.2 ± 4.6 | 1.8 ± 0.1 | 86.0 ± 4.7 |
| Q193A | 14.6 ± 0.2 | 18.4 ± 0.3 | 1.0 ± 0.0 | 50.0 ± 1.8 |
| Q193E | 168.1 ± 5.8 | 212.4 ± 7.4 | 2.6 ± 0.2 | 127.5 ± 7.2 |

*No activity at the specified assay conditions*
Table 3 Kinetic parameters of the wild-type and Y152A mutant enzymes toward 6-aminocaproic acid

|        | Wild type | Y152A |
|--------|-----------|-------|
| \(K_m\) (mM) | 3.3 ± 0.3 | 2.9 ± 0.1 |
| \(k_{cat}\) (min\(^{-1}\) × 10\(^{-3}\)) | 72.2 ± 2.0 | 208.5 ± 1.6 |
| \(k_{cat}/K_m\) (min\(^{-1}\) M\(^{-1}\)) | 21.8 ± 1.8 | 73.1 ± 2.1 |

AVAs). Consequently, our CL-GESS could prove a powerful tool for the development of a bioprocess to produce lactams and nylons.

Methods

Materials. All chemical reagents used in this study were purchased from Sigma-Aldrich (St. Louis, MO, USA). Restriction endonucleases, DNA cloning kits, pUC19 plasmid, and E. coli DH5α cells were purchased from New England Biolabs (Ipswich, MA, USA). The Diversify PCR Random Mutagenesis kit and Quik-Change II XL Site-Directed Mutagenesis kit were purchased from Stratagene (La Jolla, CA, USA) and Agilent Technologies (San Diego, CA, USA), respectively. Plasmid DNA isolation and DNA extraction were conducted using plasmid preparation kits (Promega, Madison, WI, USA). Oligonucleotides (Supplementary Table 5) were synthesized and sequenced by Macrogen (Daejeon, Korea).

Construction of the CL-GESS plasmid. We designed an e-caprolactam detection system consisting of the TR NitR, a NiIA promoter, and GFP as a reporter protein. nitR and the nitA promoter fragments were amplified by PCR from the genomic DNA of A. faecalis JM3 KCTC2867 (ATCC8750) using primer pairs 1/2 and 3/4, respectively (Supplementary Table 5). nitR was subcloned downstream of the constitutively active promoter J23016 (http://parts.igem.org), and nitA was inserted on the opposite side of nitR to avoid transcriptional noise. The egfp gene from the plasmid pMGP63 was cloned downstream of the nitR promoter. The transcription terminator rrrT1T2 from plasmid pFCEIIb and l3 from plasmid pKd46 were inserted into the C-termini of transcription terminator rrnBT1T2 from plasmid pHCEIIB and tL3 from plasmid pKD46, respectively. Plasmid DNA isolation and DNA extraction were conducted using plasmid preparation kits (Promega, Madison, WI, USA). Oligonucleotides (Supplementary Table 5) were synthesized and sequenced by Macrogen (Daejeon, Korea).

Identification of the transcription start site in Pcaprolactam. Total RNA was extracted from CL-GESS cells cultured with 10 mM caprolactam using the ecaprolactam-expressing E. coli (54°C, 1 h). cDNA was synthesized using the SMARTer RACE cDNA Amplification kit (Clontech). The transcription start site was identified by using the Switching Mechanism at the 5′ end of RNA transcript (SMART) RACE cDNA Amplification kit (Clontech) following the manufacturer’s protocol. Two reverse gene-specific primers (GSPs) for ampicillin and caprolactam were designed, including a primer for amplification of the target sequence with its template using the ModeLLer software program by applying the default model-building routine model with fast refinement. The advantage of this procedure is that it allows selection of the best model from among several candidates and that variability among models can be used to evaluate model reliability. Energy minimization was applied using a consistent valence force field and DS CHARMM with the steepest descent and conjugated gradient algorithms. The quality of the models was analyzed by the PROCHECK program. The CL-GESS was docked as the ligand in the NitR model using AutoDock Vina. Docking pocket residues were searched using the Pockit pocket detection program (http://schwarzbenjamin.free.fr/Work/Pock/home.html). The lowest energy conformation was selected for further analyses.

Mutagenesis of nitR and 3HBD. A NitR mutant library was constructed by error-prone PCR using a PCR mutagenesis kit with a mutation rate of 2–4 mutations per 1000 base pairs and using CL-GESS as a template. Fragments used to construct the mutant library were prepared by standard PCR and were ligated into the genetic circuit using the Gibson Assembly Master Mix (New England Biolabs). Mutant libraries were generated using the QuikChange II XL Site-Directed Mutagenesis kit at 37°C for 16 h. In the second round of screening, false-positive cells showing high fluorescence intensity in the absence of e-caprolactam were removed by sorting non-fluorescent cells. Approximately 1.5 × 10⁶ (bottom 3%) non-fluorescent cells in 3 × 10⁶ negatively sorted cells were collected and were grown in LB medium. In the third round, 1000 fluorescent cells with the highest fluorescence in 2 × 10⁶ (top 0.1%) in the presence of 100 μM e-caprolactam were sorted and were grown on LB agar, and a single clone exhibiting higher fluorescence intensity than cells expressing wild-type NitR was selected for further examination.

NiIR structure modeling and docking simulations. Homology modeling of A. faecalis NitR was carried out using Discovery Studio 3.1 (Accelrys, San Diego, CA, USA) based on the X-ray structure of ToC. Homology searches and sequence alignment were conducted using sequence analysis and multiple sequence alignment modules, respectively. Five models were generated based on the alignment of the target sequence with its template using the MODELLER software program by applying the default model-building routine model with fast refinement. The advantage of this procedure is that it allows selection of the best model from among several candidates and that variability among models can be used to evaluate model reliability. Energy minimization was applied using a consistent valence force field and DS CHARMM with the steepest descent and conjugated gradient algorithms. The quality of the models was analyzed by the PROCHECK program. The CL-GESS was docked as the ligand in the NitR model using AutoDock Vina. Docking pocket residues were searched using the Pockit pocket detection program (http://schwarzbenjamin.free.fr/Work/Pock/home.html). The lowest energy conformation was selected for further analyses.

Fluorescence analysis. E. coli DH5α cells harboring CL-GESS and CL-GESS mutants were grown on LB agar containing 100 μg/ml ampicillin and various substrates including e-caprolactam, l-lysine, and 6-ACA at 37°C for 16 h. Fluorescent colonies were visualized on a 2% agarose gel and were purified by PCR using CL-GESS as a template. Fragments used to construct the mutant library were amplified by PCR using primers 1/2 and 3/4, respectively (Supplementary Table 5). To investigate the NiIR-binding site and mechanisms of transcription, truncated promoter fragments were amplified using overlap PCR. The CL-GESS plasmid was constructed by ligating the DNA fragment into plasmid pUC19.

Mutant library construction and NiIR screening. E. coli DH5α cells were transformed by electroporation with ligation mixtures containing mutant genes. Transformants were spread on LB agar plates containing 100 μg/ml ampicillin and were incubated at 37°C for 16 h. The library (size: 4 × 10⁶) was stored at −70°C in storage buffer (1× TY consisting of 8 g tryptone, 5 g yeast extract, and 2.5 g NaCl in 11 of distilled water) containing 15% (v/v) glycerol until screening by flow cytometry with a FACSAria III instrument (BD Biosciences, Franklin Lakes, NJ, USA). The library (1 × 10⁶ cells) was inoculated in 2 ml of LB medium containing 100 μM e-caprolactam and 100 μM ampicillin and incubated at 37°C for 4 h prior to flow cytometry. A blue laser (488 nm) and bandpass filter (530/30 nm) were used to analyze fluorescence intensity of the mutant library. Approximately 20,000 cells with the highest fluorescence intensity were collected and were recovered in LB medium at 37°C for 16 h. The second round of screening, false-positive cells showing high fluorescence intensity in the absence of e-caprolactam were removed by sorting non-fluorescent cells. Approximately 1.5 × 10⁶ (bottom 3%) non-fluorescent cells in 3 × 10⁶ negatively sorted cells were collected and were grown in LB medium. In the third round, 1000 fluorescent cells with the highest fluorescence in 2 × 10⁶ (top 0.1%) in the presence of 100 μM e-caprolactam were sorted and were grown on LB agar, and a single clone exhibiting higher fluorescence intensity than cells expressing wild-type NitR was selected for further examination.

Cyclase purification and LC-MS analysis. After screening, the putative cyclase gene (CF3HBD) was cloned from the metagenic hit and cloned into the
Measurement of activity and kinetic parameters. The dehydrogenation activity of CF3HBD was determined based on NADH formation from NAD, by measuring the rate of CF3HBD enzyme after treatment with 10 mM EDTA, followed by dialysis against 50 mM HEPES (pH 7.5) at 4°C for 16 h. The enzyme activity was assayed with 1 mM 6-ACA in 50 mM HEPES (pH 7.5) containing 1 mM metal ion at 35°C for 10 min. In the kinetic study, various amounts of 6-ACA (0.1–20 mM) were incubated in 50 mM HEPES (pH 7.5) containing purified 3HBD enzyme at 35°C for 10 min. The reaction was stopped by the addition of HCl at a final concentration of 200 mM. The rates of the enzyme kinetic parameters kcat and km for substrates were determined by fitting the data to the Michaelis-Menten equation.

Analytical methods. 6-Caprolactam and 6-ACA contents in the reaction mixture were determined by LC-MS. To compare extraction efficiencies, we used the PhaserMR program in the CCP4 software suite40. The crystal structure was determined by molecular replacement using the PhaserMR program in the CCP4 software suite40. The P. putida 3-hydroxybutyrate dehydrogenase (PDB ID 2Q2Q) sequence was processed using CHAINSAW in the CCP4 suite according to the corresponding sequence of the CF3HBD, and then was employed as the search model for CF3HBD41. The sequence identity of CF3HBD and P. putida 3-hydroxybutyrate dehydrogenase is 70.31%. Crystallographic refinement was carried out with PHENIX41, and model building was performed using COOT42.

Data availability

The refined models of CF3HBD have been deposited in the Protein Data Bank (https://www.rcsb.org/structure/SYSS) with PDB code SYSS. All data that support the findings of this study are included in this article and in Supplementary Information. They are available from the corresponding author upon reasonable request.

Received: 25 October 2017 Accepted: 1 November 2018
Published online: 29 November 2018

References

1. Xiong, D. et al. Improving key enzyme activity in phenylpropanoid pathway with a designed biosensor. Metab. Eng. 40, 115–123 (2017).
2. Mahr, R. & Frunzke, J. Transcription factor-based biosensors in biotechnology: current state and future prospects. Appl. Microbiol. Biotechnol. 100, 79–90 (2016).
3. Kim, H. et al. A cell-cell communication-based screening system for novel microbes with target enzyme activities. ACS Synth. Biol. 5, 1231–1238 (2016).
4. Eggeling, L., Bott, M. & Marienhagen, J. Novel screening methods—biosensors. Carr. Opin. Biotechnol. 35, 30–36 (2015).
5. Schallmein, M., Frunzke, J., Eggeling, L. & Marienhagen, J. Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. Carr. Opin. Biotechnol. 26, 148–154 (2014).
6. Raman, S., Rogers, J. K., Taylor, N. D. & Church, G. M. Evolution-guided optimization of biosynthetic pathways. Proc. Natl Acad. Sci. USA 111, 17803–17808 (2014).
7. Choi, S. L. et al. Toward a generalized and high-throughput enzyme screening system based on artificial genetic circuits. ACS Synth. Biol. 3, 163–171 (2014).
8. Van Rossum, T., Kengen, S. W. & Van der Oost, J. Reporter-based screening and selection of enzymes. FEBS J. 290, 2979–2996 (2013).
9. Helm, Evd., Genee, H. J. & Sonner, M. O. A. The evolving interface between synthetic biology and functional metagenomics. Nat. Chem. 14, 752–759 (2018).
10. Tang, S. Y., Fazelinia, H. & Cinco, P. C. AraC regulatory protein mutants with altered effector specificity. J. Am. Chem. Soc. 130, 5267–5271 (2008).
11. Zhang, J. et al. Application of an Acyl-CoA ligase from Streptomyces to the production of medium and long-chain emulsifiers. ACS Synth. Biol. 5, 884–891 (2017).
12. Chae, T. U., Ko, Y. S., Hwang, K. S. & Lee, S. Y. Metabolic engineering of Escherichia coli for the production of four-, five- and six-carbon lactams. Metab. Eng. 41, 82–91 (2017).
13. Thomas, J. M. & Raja, R. Design of a “green” one-step catalytic production of epoxid-caprolactam (precursor of nylon-6). Proc. Natl Acad. Sci. USA 102, 13732–13736 (2005).
14. Beethrus, R., Rothenberg, G. & Shiju, N. R. Catalytic routes towards acrylic acid, adipic acid and caprolactam starting from biorenewables. Green Chem. 17, 1341–1361 (2015).
15. Turk, S. C. et al. Metabolic Engineering toward sustainable production of nylon-6. ACS Synth. Biol. 5, 65–73 (2016).
16. Stavila, E. & Loos, K. Synthesis of lactams using enzyme-catalyzed aminolysis. Tetrahedron Lett. 54, 370–372 (2013).
17. Komeda, H., Hori, Y., Kobyashi, M. & Shimizu, S. Transcriptional regulation of the Rhodococcus rhodochrous J1 nitA gene encoding a nitrilase. Proc. Natl Acad. Sci. USA 93, 10572–10577 (1996).
18. Stohr, W. R. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. Nucleic Acids Res. 20, 961–974 (1992).
19. Nagasawa, T., Nakamura, T. & Yamada, H. E-caprolactam, a new powerful inducer for the formation of Rhodococcus rhodochrous J1 nitrilease. Arch. Microbiol. 155, 13–17 (1990).
20. Vick, J. E. et al. Optimized compatible set of BioBrick™ vectors for metabolic pathway engineering. Appl. Microbiol. Biotechnol. 92, 1275–1286 (2011).
21. Loeschcke, A. & Thies, S. Pseudomonas putida-a versatile host for the production of natural products. Appl. Microbiol. Biotechnol. 99, 6197–6214 (2015).

22. Lowden, M. J. et al. Structure of Vibrio cholerae ToxT reveals a mechanism for fatty acid regulation of virulence genes. Proc. Natl Acad. Sci. USA 107, 2860–2865 (2010).

23. Paithankar, K. S. et al. Cosubstrate-induced dynamics of D-3-hydroxybutyrate dehydrogenase from Pseudomonas putida. FEBS J. 274, 5767–5779 (2007).

24. Kim, Y. I. et al. Improved metagenome screening efficiency by random insertion of T7 promoters. J. Biotechnol. 230, 47–53 (2016).

25. Choi, S.-L. et al. Toward a generalized and high-throughput enzyme screening system based on artificial genetic circuits. ACS Synth. Biol. 3, 163–171 (2013).

26. Kwon, K. K. et al. Evolution of enzymes with new specificity by high-throughput screening using DmpR-based genetic circuits and multiple flow cytometry rounds. Sci. Rep. 8, 2659 (2018).

27. Goñi-Moreno, Á., Benedetti, I., Kim, J. & Lorenzo, V. D. Deconvolution of gene expression noise into spatial dynamics of transcription factor-promoter interplay. ACS Synth. Biol. 6, 1359–1369 (2017).

28. Goswami, A. & Van Lanen, S. G. Enzymatic strategies and biocatalysts for aryl bond formation: tricks of the trade outside of the ribosome. Mol. Biostyst. 11, 338–353 (2015).

29. Rauwerdink, A. & Kazlauskas, R. J. How the same core catalytic machinery catalyzes 17 different reactions: the serine-histidine-aspartate catalytic triad of a/b-hydrolase fold enzymes. ACS Catal. 5, 6153–6176 (2015).

30. Hu, J. S. et al. Design and application of highly responsive fluorescence resonance energy transfer biosensors for detection of sugar in living Saccharomyces cerevisiae cells. Appl. Environ. Microbiol. 73, 7408–7014 (2007).

31. Silva-Rocha, R. et al. The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. Nucleic Acids Res. 41(Database issue), D666–D675 (2013).

32. Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. 24, 79–88 (2006).

33. Marti-Renom, M. A. et al. Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. 29, 291–325 (2000).

34. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291 (1993).

35. Trotter, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 453–461 (2010).

36. Winn, M. D. et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D 67, 235–242 (2011).