Title
Microbial Lactam Biosynthesis and Biosensing

Author
Zhang, Jingwei

Publication Date
2015

Peer reviewed|Thesis/dissertation

Microbial Lactam Biosynthesis and Biosensing

by

Jingwei Zhang

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Joint Doctor of Philosophy

with University of California, San Francisco

in

Bioengineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jay Keasling, Chair

Professor John Dueber

Professor Shawn Douglas

Professor Wenjun Zhang

Fall 2015

Microbial Lactam Biosynthesis and Biosensing

by Jingwei Zhang

Abstract

Microbial Lactam Biosynthesis and Biosensing
by
Jingwei Zhang
Joint Doctor of Philosophy
with University of California, San Francisco
in Bioengineering
University of California, Berkeley
Professor Jay D. Keasling, Chair

Lactam is an important class of commodity chemicals used in the manufacture of nylons, with millions of tons of production every year.

2-Pyrrolidone is a valuable bulk chemical with myriad applications as a solvent, polymer precursor and active pharmaceutical intermediate. A novel 2-pyrrolidone synthase, ORF27, from *Streptomyces aizunensis* was identified to catalyze the ring closing dehydration of γ-aminobutyrate. ORF27's tendency to aggregate was resolved by expression at low temperature and fusion to the maltose binding protein (MBP). Recombinant *Escherichia coli* was metabolically engineered for the production of 2-pyrrolidone from glutamate by expressing both the genes encoding GadB, a glutamate decarboxylase, and ORF27. Incorporation of a GadB mutant lacking H465 and T466, GadB_ΔHT, improved the efficiency of one-pot 2-pyrrolidone biosynthesis in vivo. When the recombinant *E. coli* strain expressing the *E. coli* GadB_ΔHT mutant and the ORF27-MBP fusion was cultured in ZYM-5052 medium containing 9 g/L of L-glutamate, 7.7 g/L of L-glutamate was converted to 1.1 g/L of 2-pyrrolidone within 31 h, achieving 25% molar yield from the consumed substrate.

ε-caprolactam and δ-valerolactam are important commodity chemicals used in the manufacture of nylons, with millions of tons produced every year. Biological production of these high valued chemicals has not been possible due to a lack of enzymes that will cyclize the ω-amino fatty acid precursors to the corresponding lactams under ambient conditions. In this study, we demonstrated proof of these bioconversions by in vitro enzyme assays. We found that ORF27, an enzyme involved in the biosynthesis of ECO-02301 in *Streptomyces aizunensis*, has a broad substrate spectrum and can not only cyclize γ-aminobutyric acid into butyrolactam, but also 5-aminovaleric acid (5-AVA) into δ-valerolactam and 6-aminohexanoic acid (6-AHA) into ε-caprolactam. The ORF27 lactam formation reaction was characterized by product analysis, and ORF27's activity on the three ω-amino fatty acids were compared. Recombinant *E. coli* expressing ORF27 produced valerolactam and caprolactam when 5-AVA and 5-AHA, respectively, were added to the culture medium. Upon co-expressing ORF27 with a metabolic pathway that produced 5-aminovaleric acid from lysine, we were able to demonstrate production of δ-valerolactam from lysine or directly from glucose.

Biological production of ε-caprolactam, δ-valerolactam and butyrolactam were enabled by the recent discovery of lactam synthases that cyclize their ω-amino fatty acid precursors to the corresponding lactams. To facilitate strain optimization with regard to product yields, productivities, and titers, it is desirable to develop a high throughput screening system for

specific small-molecule detection and response. We took the chemical informatics concepts used in small molecule drug discovery, and adapted them into a metabolic engineering strategy for targeted scouting of protein sensor candidates, named "Analog Generation towards Catabolizable Chemicals". We discovered a lactam biosensor based on the ChnR/Pb transcription factor-promoter pair. The microbial biosensor was engineered to be a single plasmid system, demonstrating dose dependent response for ε-caprolactam, δ-valerolactam and butyrolactam with great dynamic range (1.8-3.5) and wide linear range of 1-2 orders of magnitude. The biosensor also showed specificity against intermediates of lactam biosynthetic pathways and therefore could potentially be applied for high throughput metabolic engineering for industrially important lactam biosynthesis.

This dissertation is dedicated to my father, who for the first 18 years of my life, devoted the majority of his efforts to cultivate me and teach me to be a good person. I know he would have been very proud to see this work come to completion.

Acknowledgement

I have had a great experience in graduate school because of the people I encountered, to whom I am truly grateful for their help, advice, support and friendship.

None of this would have been possible without the support of my advisor and mentor, Jay Keasling. I have been very lucky to work in a lab where the advisor is supportive of wherever my imagination and passion took me. In retrospect, you have always been encouraging me to come up with my own ideas and take risks, while silently guarded me away from the futile paths. I deeply appreciate and cherish your kind protection of my passion for creativity, especially during the early days of my graduate school, when my scientific training had not allowed me to differentiate between a wild yet feasible idea vs. an impossible idea.

I can never imagine a mentor as open-minded and tolerant as you are: In additional to my research, you allowed me to mentor undergraduate students, allowed me to gain entrepreneurship experience via a part time internship at the Lawrence Berkeley National Lab tech transfer office. You also supported me to conduct some business development for my thesis project, which landed in a joint collaboration between our lab and a biotechnology company.

I joined the Keasling Lab with minimal molecular or microbial biology experience. My first work lead, Clem Fortman, who both taught me all the techniques to get me started. Clem, you are influential to me not only for the scientific thinking, but also entrepreneurial spirit and caring for family. My passion for startup was greatly shaped by your perspectives. I wish to thank other Keasling lab rooted members who went onto industry/startup endeavors, your spirit also gave me a lot of power when I look forward. Eric Steen, Mario Oullette, Eric Luning, Nathan Hillson, Jeffery Dietrich, Matt Mattozzi, Howard Chou and Leonard Katz.

During my qualify exam, Satoshi Yuzawa, Josh Gilmore, Nathan Hillson, Greg Bokinsky, Jim Kirby, Eric Luning, Tristan de Rond, Andrew Hagen, Naoki Chiba, Sean Poust, Edward Baidoo, Chris Petzold, William Holtz, Fuzhong Zhang and Ryan Phelan offered significant advice to my thesis proposal and shaped my scientific thinking.

My research was greatly helped by members of the polyketide synthase minigroup in the Keasling lab. I owe a lot to Satoshi Yuzawa, Olga Sekurova, Sergey Zotchev, Woncheol Kim, Andrew Hagen, Leonard Katz, Naoki Chiba, Sean Poust, Bob Haushalter, Anna lechner, Peter Kelly, Clara Eng and Ryan Phelan. You have spent a lot of time discussing every aspect of my project with me, from the goal, to experimental design, to data and to publication. Some people I have worked side by side with, Veronica Fok, Edward Baidoo, George Wang, Chijioke Joshua, Ouwei Wang, Suzan Yilmaz, Freeman Lan, Tuan Tran, Harry Beller, Taek Soon Lee and Ryan McAndrew. I would like to thank Dr. Bo Chen, from China National Cereals, Oils and Foodstuffs Corporation (COFCO). Thank you for putting faith in my project and established the collaborative project. I gained a great deal of insight into how biotechnology companies view R&D.

I was lucky to have worked with a group of talented and enthusiastic undergraduate students: Sandra Diaz, Anirudh Sapru, Matt Chen, Emily Kao, Bryton Dias, Kenny Hsu, Mehmet Burdu.

To Bill Shelander, your perspective on technology commercialization and entrepreneurship will inspire and influence me for many years to come.

I love my families. To my wife, Yue Liu, having your accompany during graduate school was the best thing that happened to me. I am grateful for all your sacrifices; for your bringing two lovely children to our family; for spending your precious five years staying with me, far away from your home, your family and friends. To my parents, and the family on my wife's side, thank you for your support and guidance for my pursuing my PhD, alleviating a lot of the life's burdens so that I could concentrate on my study. To my grandma on my father's side, I am about to complete an important milestone you assigned to me during my childhood. Thank you for teaching me all the important lessons in life, they have been giving me courage whenever I encounter difficulties. I wish you enjoy a healthy and happy life. To my grandparents on my mum's side, I wish you peaceful up there in heaven. Now I am about to obtain my PhD in bioengineering. I hope I make you proud.

Thank you for reading and commenting on my dissertation draft.

Jingwei Zhang Curriculum Vitae

Email: harryharry424@gmail.com

Education:
Univ. of California, Berkeley 2010-present
Ph.D. candidate in Joint Graduate Group in Bioengineering
University of California, Berkeley - Walter A. Haas School of Business
Certificate, Management of Technology
Univ. of California, Berkeley, 2008-2010, B.S., Chemical Biology
Univ. of California, Los Angeles, 2007-2008, Chemistry and Biochemistry
The University of Hong Kong, 2006-2007, Biochemistry
Tsinghua University, 2005-2006, Chemical Engineering

Current Position:
Graduate Researcher, 2010-present
Keasling Lab, Department of Bioengineering, UC Berkeley
Combine Metabolic Engineering Technologies to Develop Biorenewable Chemicals for a Sustainable Future.

Honors/Awards:
Berkeley Fellowship, 2010-2012 (UCB)
George C. Pimentel Award, 2010 (UCB)
HKU Exchange Scholarships, 2008 (HKU)
Fan Charitable Foundation Scholarship, 2007 (UCLA)
Patrick Chow Lum Wong Memorial Prize in Biochemistry, 2007 (HKU)
Award of Scholarships Scheme For Mainland Students, 2006-2008 (HKU)

Publications and Inventions
1. Metabolic Engineering of *Escherichia coli* for the Biosynthesis of 2-Pyrrolidone
(Accepted, Metabolic Engineering Communications 3 (2016) 1–7)
Jingwei Zhang, Emily Kao, George Wang, Edward E.K. Baidoob, Matthew Chen, Jay. D. Keasling*

2. Application of an Acyl-CoA Synthetase from *Streptomyces aizunensis* for Lactam Biosynthesis (in preparation)
Jingwei Zhang, Mehmet Burdu, George Wang, Edward E.K. Baidoob, Jay. D. Keasling*

3. Development of a Transcription Factor Based Lactam Biosensor (submitted)
Jingwei Zhang, Mehmet Burdu, Bryton Dias, Jay. D. Keasling*

Hybrid Polyketide Synthases PCT/US2011/058660 filed on 31-Oct-2011
A Host Cell Modified to Produce 2-Pyrrolidone US Patent Application No. 62/147,461 filed on 14-Apr-2015
Invention Disclosure: Microbial Production of Valerolactam and Caprolactam
Invention Disclosure: Transcription Factor Based Biosensor

Technology Licensing
The Hybrid Polyketide Synthases technology was licensed to Lygos Inc.

Table of Contents

Chapter 1 Metabolic Engineering of *Escherichia coli* for the Biosynthesis of 2-Pyrrolidone *

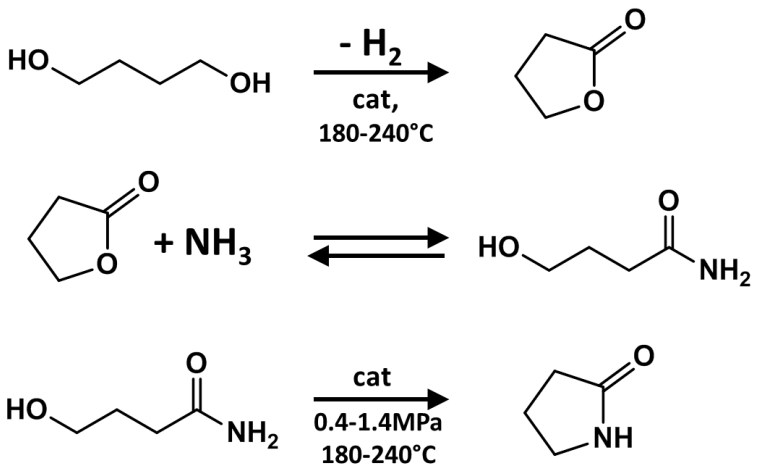
1.1. Background

2-Pyrrolidone was identified by the US Department of Energy as an important C4 "Top Value-Added Chemical from Biomass" that can potentially be derived from glutamate (1). 2-Pyrrolidone is currently used as precursor for the production of N-vinylpyrrolidone, a solvent for animal injection, a building block for active pharmaceutical ingredients, optical co-solvent for water-based ink formulation, process solvent for membrane filters and a copolymer for floor polish (2). 2-Pyrrolidone can also be used in ring-opening polymerization to produce nylon-4, a fiber material with better thermal stability and the highest hydrophilicity in the nylon family of materials (3). With a variety of applications, 2-pyrrolidone continues to be a product of huge commercial interest.

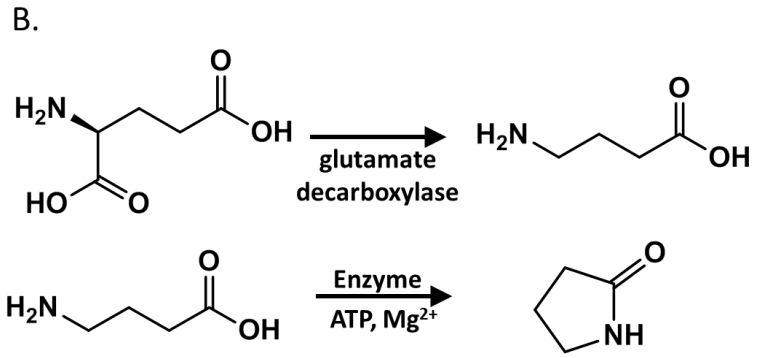
Current industrial production of 2-pyrrolidone involves the dehydrogenation of 1,4-butanediol to form γ–butyrolactone, followed by reacting aqueous γ-butyrolactone with ammonia (Fig. 1.1A) (4). By using low cost glutamate as starting material, as well as avoiding harsh reaction conditions, biological production of 2-pyrrolidone offers the potential for a cheaper and more environmentally friendly synthesis route. Therefore, we propose a two-step enzymatic process for 2-pyrrolidone biosynthesis from glutamate: 1) decarboxylation of glutamate to form γ-aminobutyrate (GABA), 2) enzymatic ring closing of GABA into 2-pyrrolidone (Fig. 1.1B).

Figure 1.1. Routes for production of 2-pyrrolidone. (A) BASF petrochemical route for 2-pyrrolidone production. (B) Microbial 2-pyrrolidone biosynthetic route.

A.



* Reproduced with permission from Jingwei Zhang, Emily Kao, George Wang, Edward E.K. Baidoob, Matthew Chen, Jay. D. Keasling. "Metabolic Engineering of *Escherichia coli* for the Biosynthesis of 2-Pyrrolidone" (Metabolic Engineering Communications 3 (2016) 1–7, Accepted)



While the first enzymatic step is known (3, 5-8), the second step has not been demonstrated under any fermentation conditions (9). We employed a targeted strategy to identify appropriate enzyme candidates for the GABA activation step by conducting retro-biosynthetic analysis of polyketides. Here we report the discovery in *Streptomyces aizunensis* of ORF27, an auxiliary enzyme in the linearmycin A biosynthetic cluster that performs the GABA activation step to form 2-pyrrolidone under mild fermentation conditions. Coupling this 2-pyrrolidone synthase with glutamate decarboxylase, which forms GABA from glutamate, we achieved the first demonstration of the full 2-pyrrolidone biosynthetic pathway in *E. coli*.

1.2. Materials and Methods

1.2.1. Enzyme scouting by retro-biosynthetic analysis of PKS natural products

It was hypothesized that an enzyme employing a GABA activation mechanism on GABA's acid group followed by spontaneous irreversible intramolecular cyclization could be utilized as a 2-pyrrolidone synthase. However, there has been no documentation of AMP-activating enzymes using GABA as a substrate. Type I polyketide synthases (PKSs) are mega-synthases whose enzymatic domain organization predictably correlate with their natural product's chemical structure – usually referred to as the colinearity rule (10-19). PKSs are responsible for biosynthesis of the characteristic polyketide aglycone backbone, and auxiliary enzymes are involved in starter unit activation and/or post aglycone modification (20-26). The close positioning of PKS genes with auxiliary enzymes in the same gene cluster makes it relatively easy to identify the substrates upon which these auxiliary enzymes act. PKS databases were explored to identify polyketides with positively charged, amine-containing starter units and auxiliary enzymes clustered near the PKS (see Table S1.1 for lists of PKS databases). Desertomycin A/B, Eco-0501, linearmycin A were selected for further retrobiosynthetic analysis (Fig. S1.1A). From the arrangement of the polyketide synthase genes, 4-aminobutyrate-like starter unit to be loaded onto the ACP0 domain (Fig. S1.1B). Since polyketide starter unit loading requires the substrate to be activated by an ATP dependent mechanism, the linearmycin A biosynthetic gene cluster was further analyzed for auxiliary enzymes with this function. Two ORFs in the gene cluster, ORF27 and ORF36, were predicted to be AMP-dependent synthetases (Fig. S1.1B). The hypothesized candidates were reverse transcribed according to the optimized *E. coli* codon usage using DNA 2.0's algorithm and synthesized.

1.2.2. Bacterial strains and chemicals

E. coli strain DH10B [F− mcrA Δ(mrr-hsdRMS-mcrBC) φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ− rpsL (Strr) nupG] was used for all molecular biology manipulations. DH10B or BL21 Star (DE3) [F− ompT hsdSB (rB− mB−) gal dcm rne131 (DE3)] were used as hosts for production of 2-pyrrolidone. All the strains and plasmids utilized in this study are listed in Table 1.1. For high-density shake flask cultures, Studier's autoinduction ZYM-5052 medium was prepared according to the published protocol (27). Chloramphenicol (25 μg/mL), kanamycin (20 μg/ml) and ampicillin (100 μg/ml) were added where desired to provide selective pressure for plasmid maintenance. During 2-pyrrolidone production, the following antibiotic concentration was used: chloramphenicol (6.25 μg/mL), kanamycin (5 μg/ml) and ampicillin (25 μg/ml). 2-Pyrrolidone, glutamic acid, GABA, and ATP were purchased from Sigma-Aldrich (St. Louis, Missouri).

Table 1.1. *E. coli* strains, plasmids and oligonucleotides used.

Name	Relevant genotype	Reference
Strains		
DH10B	F⁻ mcrA crmrr-hsdRMS-mcrBC) r-hsdRMS-mcrBC) and oligonucleotide139 Δ(ara, leu)7697 galU galK alrpsL nupG	Life Technologies (Carlsbad, CA)
BL21 Star (DE3)	F− ompT hsdSB (rB− mB−) gal dcm rne131 (DE3)	Life Technologies (Carlsbad, CA)
JW2637-4	F-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, ΔgabT743::kan, rph-1, Δ(rhaD-rhaB)568, hsdR514	CGSC #11775
JW0036-1	F-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, ΔCaiC750::kan, rph-1, Δ(rhaD-rhaB)568, hsdR514	CGSC #8352
Plasmids		
pDNA2.0-ORF27	pUC *ori* , Kan^R, ORF27 with Biobrick overhang	DNA 2.0 (Menlo Park, CA)
pBbE2C- RFP	ColE *ori*, Cm^R, *tetR*, RFP	TS. Lee (2011)(28)
pBbE2C-ORF27	ColE *ori*, Cm^R, *tetR*, ORF27	This study
pBbE2C-CaiC	ColE *ori*, Cm^R, *tetR*, CaiC	This study
pBbS2C-ORF27	SC101 *ori*, Cm^R, *tetR*, ORF27	This study
pET28b	pBR322 *ori*, Kan^R , lacI, MCS after P_{T7}	EMD Millipore (Billerica, MA)
pET28a-MBP	pBR322 *ori*, Kan^R , lacI, MCS after MBP fusion driven by P_{T7}	EMD Millipore (Billerica, MA)
pET28a-MBP-ORF27	pET28a-MBP with ORF27 inserted into NdeI and XhoI	This study
pBbE2C-MBP-ORF27	ColE *ori*, Cm^R, *tetR*, MBP-ORF27	This study
pBbA5a-GadB	p15A ori, AmpR ,Placuv5, GadB	This study
pBbA5a-GadB_ΔHT	p15A ori, AmpR , Placuv5, GadB_ΔHT	This study
pBbA5a-GadB_Δ1-14	p15A ori, AmpR , Placuv5, GadB_ Δ1-14	This study
pBbA5a-GadB_Δ1-14, ΔHT	p15A ori, AmpR , Placuv5, GadB_ Δ1-14, ΔHT	This study
pBbA7a-RFP	p15A ori, AmpR ,PT7, RFP	TS. Lee (2011)
pBbA7a-GadB	p15A *ori*, Amp^R ,P_{T7}, GadB	This study
pBbA7a-GadB_ΔHT	p15A *ori*, Amp^R ,P_{T7}, GadB_ΔHT	This study
pBbA7a-GadB_Δ1-14	p15A *ori*, Amp^R ,P_{T7}, GadB_ Δ1-14	This study
pBbA7a-GadB_Δ1-14_ΔHT	p15A *ori*, Amp^R ,P_{T7}, GadB_ Δ1-14_ΔHT	This study
Oligonucleotides	5'→ 3' Sequence, restriction site underlied, synthesized by Integrated DNA Technologies, Inc (Coralville, IA)	Target Gene

JZ_MBP-ORF27_f	GCGCGgaattcaaaagatcttttaagaaggagatatacatatgggcagcagccatcatca	MBP-ORF27
JZ_MBP-ORF27_r	GCGCGctcgagtttggatccTCATTCTGCCGCCATACGGG	MBP-ORF27
JZ_GadB_f	gcgcgCATATGgataagaagcaagtaacg	GadB, GadB_ΔHT,
JZ_GadB_r	gcgcgGGATCCTTAtcaggtatgtttaaagctgtt	GadB, GadB_Δ1-14
JZ_GadBΔ1-14_f	gcgcgCATATGGGTTCACGTTTTGGTGCGA	GadB_Δ1-14, GadB_Δ1-14_ΔHT
JZ_GadBΔHT_r	gcgcgGGATCCTTAtcatttaaagctgttctgttggg	GadB_Δ1-14_ΔHT GadB_ΔHT
JZ_CaiC_f	GCGCgaattcaaaagatcttttaagaaggagatatacatATGGATAGAGGTGCAATGGAT	CaiC
JZ_CaiC_r	GCGCG ctcgagtttggatccTTATTTCAGATTCTTTCTAATTATTTTCCCC	CaiC

Strains	plasmids	Host
JZ-171	pBbE2C-ORF27	JW2637-4
JZ-172	pBbE2C-RFP	JW2637-4
JZ-298	pBbS2C-ORF27	JW2637-4
JZ-299	pBbE2C-MBP-ORF27	JW2637-4
JZ-336	pBbE2C-MBP-ORF27+pBbA5a-GadB WT	BL21 Star (DE3)
JZ-338	pBbE2C-MBP-ORF27 + pBbA5a-GadB_ΔHT	BL21 Star (DE3)
JZ-339	pBbE2C-MBP-ORF27 + pBbA5a-GadB_Δ1-14	BL21 Star (DE3)
JZ-340	pBbE2C-MBP-ORF27 + pBbA5a-GadB_Δ1-14, ΔHT	BL21 Star (DE3)
JZ-342	pET28a-MBP-ORF27 + pBbA7a-GadB WT	BL21 Star (DE3)
JZ-344	pET28a-MBP-ORF27 + pBbA7a-GadB_ΔHT	BL21 Star (DE3)
JZ-345	pET28a-MBP-ORF27 + pBbA7a-GadB_Δ1-14	BL21 Star (DE3)
JZ-346	pET28a-MBP-ORF27 + pBbA7a-GadB_Δ1-14, ΔHT	BL21 Star (DE3)
JZ-348	pBbE2C-MBP-ORF27+pBbA7a-GadB WT	BL21 Star (DE3)
JZ-350	pBbE2C-MBP-ORF27 + pBbA7a-GadB_ΔHT	BL21 Star (DE3)
JZ-351	pBbE2C-MBP-ORF27 + pBbA7a-GadB_Δ1-14	BL21 Star (DE3)
JZ-352	pBbE2C-MBP-ORF27 + pBbA7a-GadB_Δ1-14, ΔHT	BL21 Star (DE3)
JZ-370	pBbE2C-CaiC	JW2637-4
JZ-386	pET28b-N-CaiC + pBbA7a-GadB_ΔHT	BL21 Star (DE3)
JZ-393	pBbE2C-RFP	JW0036-1

1.2.3. Plasmid construction

Genes encoding ORF27 (GenBank: AAX98201.1) and ORF36 (GenBank: AAX98210.1) were recoded using *E. coli* codon usage with biobrick overhangs 5'-gaattcaaaAGATCTAGGAGGCAT-3' on the 5' end and 5'-TAAGGATCCAAACTCGAG-3' on the 3' end. DNA 2.0 (Menlo Park, CA) cloned the genes into plasmid vectors creating pDNA2.0-ORF27 and pDNA2.0-ORF36, respectively. The genes encoding wild-type GadB and the variant GadB_ΔHT, were amplified from *E. coli* MG1655 genomic DNA using the primers specified. GadB_ΔHT lacks histidine 465 and threonine 466 of *E. coli* GadB. The construction of each plasmid is described in Table 1.2 and Table S1.2.

Table 1.2. Plasmid construction.

Constructed Plasmid	**Backbone Source (restriction site)**	**Gene Source (direct digestion or PCR)**	**PCR primers**
pBbE2C-ORF27	pBbE2C-RFP (BglII, XhoI)	pDNA2.0-ORF27 (BglII, XhoI)	N/A

pBbS2C-ORF27	pBbS2C-RFP (BglII, XhoI)	pDNA2.0-ORF27 (BglII, XhoI)	N/A
pET28b-N-CaiC	pET28b (NdeI, XhoI)	pBbE2C-CaiC (NdeI, XhoI)	N/A
pBbE2C-ORF27	pBbE2C-RFP (BglII, XhoI)	pDNA2.0-ORF27 (BglII, XhoI)	N/A
pBbE2C-CaiC	pBbE2C-RFP (EcoRI, XhoI)	*E. coli* MG1655 gDNA PCR	JZ_CaiC_f, JZ_CaiC_r
pET28a-MBP-ORF27	pET28a-MBP (NdeI, XhoI)	pDNA2.0-ORF27 (NdeI, XhoI)	N/A
pBbE2C-MBP-ORF27	pBbE2C-RFP (BglII, XhoI)	pET28a-MBP-ORF27 PCR	JZ_MBP-ORF27_f, JZ_MBP-ORF27_r
pBbA5a-GadB	pBbA5a-RFP (NdeI, BamHI)	*E. coli* MG1655 gDNA PCR	JZ_GadB_f, JZ_GadB_r
pBbA5a-GadB_ΔHT	pBbA5a-RFP (NdeI, BamHI)	*E. coli* MG1655 gDNA PCR	JZ_GadB_f, JZ_GadB_ΔHT _r
pBbA5a-GadB_Δ1-14	pBbA5a-RFP (NdeI, BamHI)	*E. coli* MG1655 gDNA PCR	JZ_GadBΔ1-14_f, JZ_GadB_r
pBbA5a-GadB_Δ1-14, ΔHT	pBbA5a-RFP (NdeI, BamHI)	*E. coli* MG1655 gDNA PCR	JZ_GadBΔ1-14_f, JZ_GadB_ΔHT _r
pBbA7a-GadB	pBbA7a-RFP (NdeI, BamHI)	pBbA5a-GadB (NdeI, BamHI)	N/A
pBbA7a-GadB_ΔHT	pBbA7a-RFP (NdeI, BamHI)	pBbA5a-GadB_ΔHT(NdeI, BamHI)	N/A
pBbA7a-GadB_Δ1-14	pBbA7a-RFP (NdeI, BamHI)	pBbA5a-GadB_Δ1-14(NdeI, BamHI)	N/A
pBbA7a-GadB_Δ1-14_ΔHT	pBbA7a-RFP (NdeI, BamHI)	pBbA5a-GadB_Δ1-14, ΔHT(NdeI, BamHI)	N/A

1.2.4. 2-Pyrrolidone, GABA and Glutamate Product Assays

1. 2.4.1 Liquid chromatography method for 2-pyrrolidone, GABA and glutamate separation

Liquid chromatography (LC) separation of 2-pyrrolidone was conducted at 55°C with an Inertsil ODS-3 reverse-phase C18 column (250 mm length, 2.1 mm internal diameter, 3 μM particle size; GL Sciences) using a 1100 series high-performance LC system (Agilent Technologies). The mobile phase was composed of 0.1% formic acid in H_2O (solvent A) and 0.1% formic acid in MeOH (solvent B). Butyrolactam was separated with the following gradient: 40% to 60% B for 4.5 min, 60% to 100% B for 0.5 min, 100% to 40% B for 0.5 min, held at 10% B for 8.5 min. A flow rate of 0.18 mL/min was used throughout.

1.2.4.2 Mass Spectrometry Analysis of 2-pyrrolidone, GABA and Glutamate

The LC system was coupled to an Agilent Technologies LC-MSD SL electrospray ionization mass spectrometer (ESI MS). 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 10 L/min and 20 psig, respectively, and a drying gas temperature of 300°C was used throughout. ESI was conducted in the positive-ion mode with a capillary voltage of 4 kV. Mass measurements were carried out in the selected ion monitoring (SIM) mode (2-pyrrolidone, m/z 86; GABA, m/z 104; glutamate, m/z 148) for the detection of [M + H]+ ions. Data acquisition and processing were performed using ChemStation (Agilent technologies).

1.2.5. 2-Pyrrolidone production titer determination

600 μL of culture was cooled on ice and centrifuged at 18,000×g for 5 min at 4°C. 250 μL of the supernatant was mixed with 250 μL methanol to a final concentration of 50% (v/v), and the mixed solution filtered through 10 K Amicon Ultra-0.5 mL Centrifugal Filters (Millipore) by centrifuging at 20,000×g for 15 min. The filtered solution was diluted into the respective linear range of detection for 2-pyrrolidone and analyzed by LC-MS in the SIM mode described above.

1.2.6. Culture conditions

1.2.6.1 Inducible 2-pyrrolidone production from GABA

E. coli strains (JZ-298, JZ-171, JZ-299, JZ-370, JZ-393) harboring plasmids containing genes encoding the proposed GABA activating enzyme (ORF27, MBP-ORF27 or CaiC) were inoculated into 25 mL LB medium with chloramphenicol (25 μg/mL) and grown at 37°C. *E. coli* expressing RFP was utilized as a negative control (JZ-172). When the OD_{600} reached around 0.5, the culture was cooled to various temperatures (18°C - 37°C). 50 ng/mL of anhydrotetracycline (aTc) was added for protein production and GABA was supplied to a final concentration of 0-10 mM. The 2-pyrrolidone titer was analyzed 24 h after induction.

1.2.6.2 Inducible 2-pyrrolidone production from glutamate

E. coli strains (JZ-336, JZ-338~JZ-340, JZ-342, JZ-344~JZ-346, JZ-348, JZ-350~ JZ-352, JZ-386) harboring plasmids containing genes encoding both glutamate decarboxylase and a GABA-activating enzyme were inoculated into 25 mL LB medium containing various concentrations of glutamate (0 to 9 g/L) with appropriate antibiotics and grown at 37°C. When the OD_{600} reached around 0.6, the culture was cooled to 25°C. IPTG and aTet were added to a final concentration of 500 μM and 50 ng/mL, respectively to induce protein expression. The pH was titrated by adding 0.6 N HCl solution, and the culture was placed in a 25°C incubator. 2-Pyrrolidone titer was analyzed at 24 h after induction.

1.2.6.3 Autoinducible 2-pyrrolidone production from glutamate

E. coli strains (JZ-344, JZ-386) harboring plasmids containing genes encoding both glutamate decarboxylase and a GABA-activating enzyme were inoculated into 10 mL of LB or LB plus 5 g/L glutamate overnight. On day 2, the overnight culture was inoculated 1:100 (v/v) into 25 mL Studier's autoinduction ZYM-5052 medium with various concentrations of glutamate (0 g/L to 9 g/L) and appropriate antibiotics (27). The culture was incubated at 37°C. When the OD_{600} reached around 0.6, the culture was cooled to 25°C. The pH was titrated to 5.25 by adding 0.6N HCl solution. The culture was then placed at 25°C, and 2-pyrrolidone titer was analyzed at 24h later.

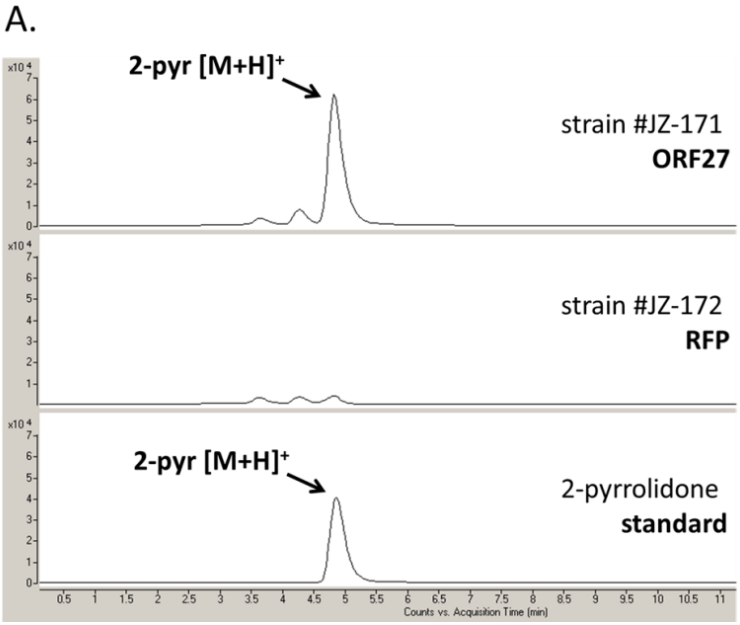
1.3. Results

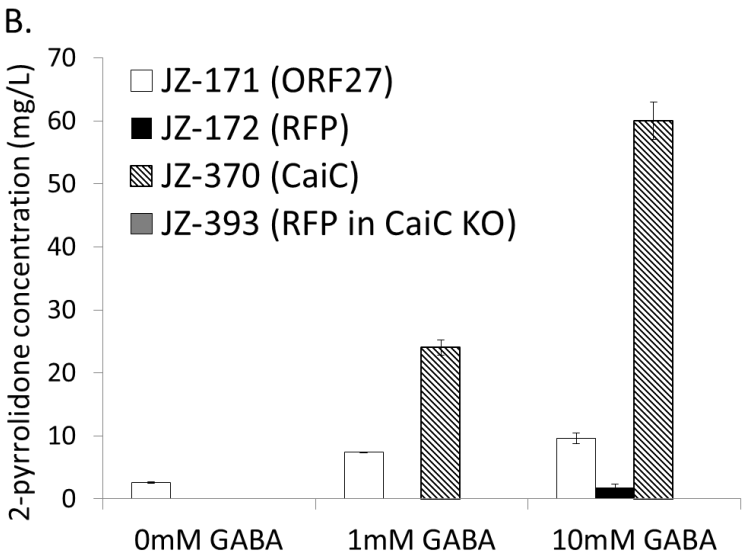
1.3.1. ORF27 and *E. coli*'s native CaiC catalyze 2-pyrrolidone formation in vivo

E. coli JW2637-4 from the KEIO knockout library contains a knockout of *gabT*, which encodes a GABA transaminase involved in GABA catabolism (29). This host was initially used to confirm

production of 2-pyrrolidone in vivo. ORF36 overexpression did not result in 2-pyrrolidone production (data not shown). *E. coli* JZ-171 (ORF27) and JZ-172 (RFP negative control) were grown in LB medium containing 0, 1, and 10 mM GABA. At 1 mM GABA, 2-pyrrolidone was observed only in the presence of ORF27 (Fig. 1.2A). Surprisingly, at 10 mM GABA, slight 2-pyrrolidone production was observed in the RFP control strain, indicating that nonspecific activation of GABA by a native *E. coli* enzyme also contributes to 2-pyrrolidone formation. To determine which native *E. coli* enzymes were responsible for the activity, genes encoding various AMP activating enzymes, such as *acs*, *prpE* and *caiC*, were cloned and overexpressed. Overexpression of one of these enzymes, CaiC (strain JZ-370), a previously described betaine-CoA ligase (AMP activating enzyme family), led to 2-pyrrolidone formation (30). *E. coli* JW0036-1, the KEIO collection *caiC* deletion mutant, was transformed with pBbE2C-RFP, creating strain *E. coli* JZ-393, which overexpresses RFP; this strain was no longer able to produce 2-pyrrolidone when fed 10 mM GABA (Fig. 1.2B).

Figure 1.2. 2-Pyrrolidone production in vivo. (A) At 1mM GABA, 2-pyrrolidone was observed in the strain where the ORF27 gene (JZ-171) is expressed. RFP expression strain (JZ-172) serves as negative control. (B) At 10mM GABA concentration, 2-pyrrolidone production is still observed in JZ-172 due to background CaiC expression, when *caiC* was knocked out (JZ-393), 2-pyrrolidone is no long produced.





1.3.2. Increased ORF27 functional expression improves GABA to 2-pyrrolidone conversion in vivo

ORF27 had the tendency to aggregate and become insoluble one day after induction depending on the temperature (Fig. S1.2). By balancing reaction rate and enzyme functional expression, the optimal temperature was determined to be 25°C for ORF27 to maximize 2-pyrrolidone titer from GABA. An MBP-ORF27 fusion increased solubility of ORF27 and resulted in a 1.5X -2X improvement in overall titer. ORF27 was expressed from a low copy number plasmid harboring the SC101 origin of replication (pBbS2C-ORF27). However, use of this plasmid decreased the titer, presumably due to lower expression of the enzyme (Fig. S1.3). Chaperone coexpression was tested as another strategy to improve solubility (related strains listed in Table S1.2). This strategy, however, also resulted in decreased titer (Fig. S1.4).

1.3.3. Inducible 2-pyrrolidone production from glutamate

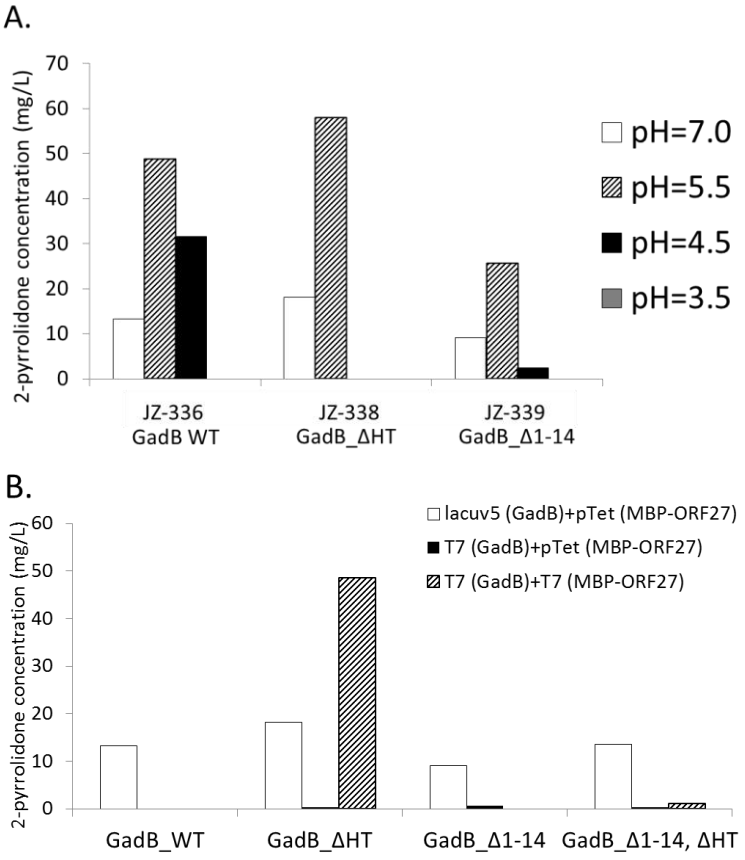
1.3.3.1. pH profile experiment

A two-step 2-pyrrolidone biosynthetic route from glutamate consists of the *E. coli* glutamate decarboxylase GadB and MBP-ORF27 as the 2-pyrrolidone synthase. *E. coli*'s GadB has a pH optimum of 3.8 and its catalytic activity significantly decreases when the pH rises above 5.0 (31). In GadB, H465 plays an important role in blocking the enzyme active site at pHs near neutral by forming aldamine between its distal nitrogen on the imidazole ring with GadB's PLP-L276 schiff base; GadB mutants, such as H465A or GadB_ΔHT (lacking H465 and T466), are able to maintain relatively high activity at more alkaline pH (31). As shown in the extracellular pH profile for 2-pyrrolidone biosynthesis strains carrying either wild type or one of two GadB mutants (JZ-336, JZ-338, JZ-339), pH ~5.5 resulted in the highest 2-pyrrolidone titer for both GadB wild type and variants (Fig. 1.3A). Sensitivity analysis showed that strains carrying GadB ΔHT robustly gave optimum yield within the pH range of 5.0 – 5.5, and declined when the exogenous pH drifted above 6.0 (Fig. S1.5).

1.3.3.2. Promoter engineering

To increase the flux through the pathway, various inducible promoters (lacUV5, Tet, T7) were tested (JZ-336, JZ-338~JZ-340, JZ-342, JZ-344~JZ-346, JZ-348, JZ-350~ JZ-352). Production of 2-pyrrolidone was highest when strong promoters were used, so the T7 promoter was chosen to drive expression of GadB ΔHT and MBP-ORF27 for optimal 2-pyrrolidone production (Fig. 1.3B).

Figure 1.3. (A) GadB mutants and the effect of extracellular pH on 2-pyrrolidone titer when feeding 10 mM glutamic acid. (B) The effect of promoter strength on 2-pyrrolidone titer when feeding 10mM glutamic acid at an extracellular pH of 7.0.

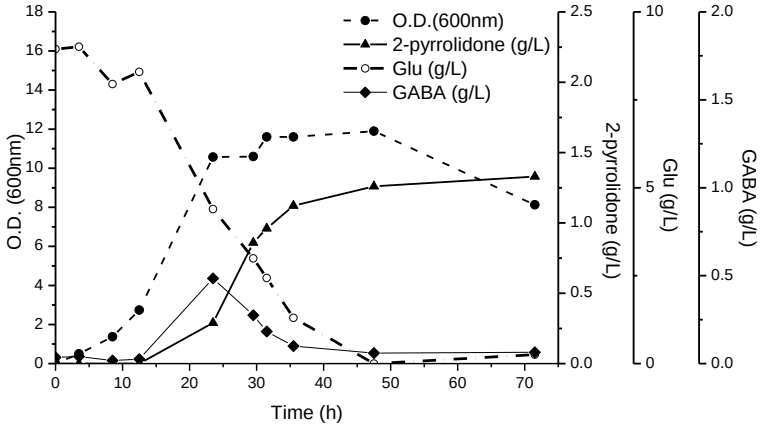


1.3.4. Autoinducible 2-pyrrolidone production from glutamate

Studier's ZYM-5052 is a high density culture medium that enables autoinduction of protein expression in *E. coli* (27). Strain JZ-344, grown overnight in LB medium, was tested through 2-stage autoinducible 2-pyrrolidone production from glutamate: stage I, cell growth to OD_{600} ~ 0.6 at 37°C; stage II, pH titration to 5.25 and incubation at 25°C for 2-pyrrolidone production. Production in medium with 9 g/L and 0 g/L of glutamate showed that over 98% of 2-pyrrolidone was converted from glutamate supplied in the medium (Fig. S1.6). Overall, 1.1 g/L of 2-pyrrolidone was produced by converting 7.7 g/L of L-glutamate within 31 h, achieving 25% molar yield from consumed substrate (Fig. 1.4). The formation of GABA from glutamate likely occurred during 12-40 h, since during this time, the pH increased from 5.36 to 8.07 (Fig. S1.7).

There was an approximate 12-hour time delay between extracellular GABA accumulation and glutamate consumption, presumably due to delayed autoinduction of enzyme expression. GABA accumulated around 12-30 h, yet was consumed quickly. This could result from transcriptional activation of GABA catabolic pathway in *E. coli*. GABA was rapidly transformed into 2-pyrrolidone when MBP-ORF27 was solubly expressed in day 1. 2-Pyrrolidone formation slowed down, presumably due to inactivation at approximately 48 h (Fig. 1.4). The OD_{600} of the culture decreased from 12 to 8, presumably due to either cell clumping or lysis.

Figure 1.4. *E. coli* strain #344 Production 2-pyrrolidone from glutamic acid (closed triangle) and OD_{600} of *E. coli* growth (closed circle). Glutamic acid feed consumption (open circle), GABA intermediate accumulation (closed diamond).



1.3.5. *E. coli* CaiC

E. coli CaiC was also tested (30). Although CaiC has better thermostability and outperformed ORF27 in 2-pyrrolidone biosynthesis during GABA feeding, expression of *E. coli* CaiC under glutamate feeding conditions resulted in only 18.7% of the 2-pyrrolidone titer of that achieved when engineered ORF27 fused with MBP was used (data not shown).

1.4. Discussion

Through years of engineering, the petrochemical industry created over 4,000 bulk chemicals (from ICIS). Currently, around 200 products are made via microbial fermentation (excluding food and beverages), including amino acids, bioactive compounds, etc. With advances in synthetic biology, it is now possible to engineer microbes to produce chemicals traditionally made via petrochemical processes at lower cost.

An important step towards building molecules of increasing complexity is to discover enzymes for substrate activation, such as C-C bond and C-N bond formation (32). Here we successfully demonstrated an efficient methodology to scout for enzymes from the secondary metabolite linearmycin A biosynthetic gene cluster that activate γ-aminobutyrate. The huge diversity of secondary metabolites and the functional groups embedded in these molecules made them a rich source of gene candidates (15, 33). Because polyketide biosynthesis follows the co-linearity rule

and is highly predictable in terms of its biosynthetic pathway, polyketide biosynthetic clusters are convenient sources of substrate activating enzymes.

Although the dehydration of γ-aminobutyrate to form 2-pyrrolidone is thermodynamically favorable, due to its high activation barrier, a significant temperature (>200°C) is required for the reaction to proceed even in the presence of Al_2O_3 catalyst. This reaction requires several days to complete, and the harsh conditions also lead to off pathway reactions, producing oligomers or cyclic GABA dimer or trimer (9). The discovery of *S. aizunensis* ORF27 enables the first demonstration of 2-pyrrolidone production from γ-aminobutyrate at mild fermentation conditions.

S. aizunensis ORF27 was prone to aggregation when heterologously overexpressed in *E. coli*. Utilizing the MBP fusion alleviated ORF27 aggregation and maintained enzymatic activity for prolonged 2-pyrrolidone production, thus leading to an increased final titer. However ORF27 still suffers from a tendency to aggregate. For prolonged 2-pyrrolidone biosynthesis to be sustained during a production process, it would be desirable to evolve ORF27 to be more soluble and stable. In addition, use of a glutamate or GABA overproduction host, such as *Corynebacterium glutamicum*, would be desirable for production of 2-pyrrolidone from glucose or another carbon source (34, 35).

ORF27 and CaiC exhibited different comparative 2-pyrrolidone formation behavior. During 10mM GABA feeding, CaiC had better performance, while during direct production from glutamate, ORF27 performed better. ORF27 might have a low K_m for GABA, therefore high intracellular GABA production during glutamate feeding experiment result in much higher 2-pyrrolidone formation for ORF27, although more enzymology work needs to be conducted.

1.5. Conclusion

In this study, we utilized retro-biosynthetic analysis of polyketide natural products as a targeted method to prospect for novel γ-aminobutyrate activating enzymes for performing unprecedented reactions such as GABA conversion to 2-pyrrolidone. *E. coli*'s native CaiC, a betaine-CoA ligase, was also discovered to be able to catalyze 2-pyrrolidone formation. Protein modification, such as MBP fusion, increased the activity of expressed ORF27. Metabolic engineering and process optimization collectively improved 2-pyrrolidone titer from glutamate. 1.1 g/L of 2-pyrrolidone was produced from 7.7 g/L of glutamate, representing a 25% yield from consumed substrate.

Supplemental Material

Supplementary Figures

Figure S1.1. Linearmycin A structure and genomic arrangement coding for linearmycin A biosynthesis. A) The structure of natural products that contain 4-aminobutyrate like starter unit, Desertomycin A/B, Eco-0501, linearmycin A. The red box refers to starter units loaded onto polyketide synthases. B) genomic arrangement of the locus coding for linearmycin A PKS aglycone biosynthesis. Open reading frame (ORFs) 10 through 18 encode the type I polyketide synthase, the predicted domains of these ORFs are shown. Based on the structure of Linearmycin A, 4-aminobutyryl-CoA was predicted to be loaded as a starter unit. (Adapated from McAlpine, 2005)

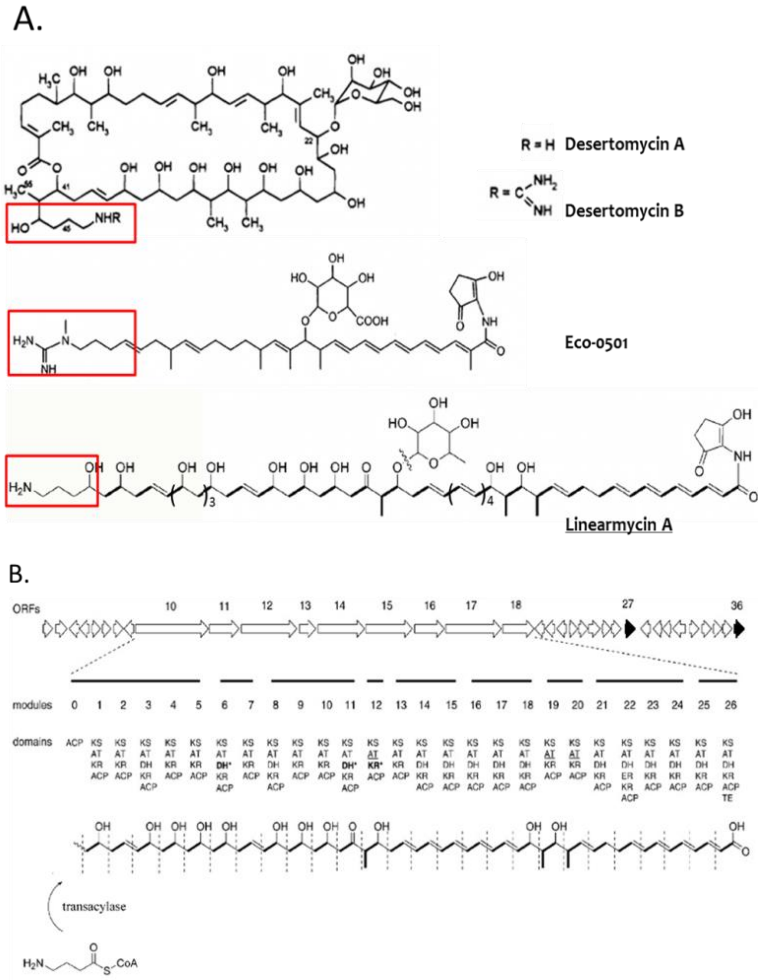


Figure S1.2. Time dependent expression solubility of C-6xHis-ORF27. A) Day 1, solubility of C-6xHis-ORF27 expression at 18°C, 22°C, 25°C and 30°C. B) Day 2 - 3, aggregation of C-6xHis-ORF27 in 18°C and °C. "s" stands for soluble fraction, "in" stands for insoluble fraction. C) Day 2 - 3, aggregation of C-6xHis-ORF27 in 25°C and 30°C.

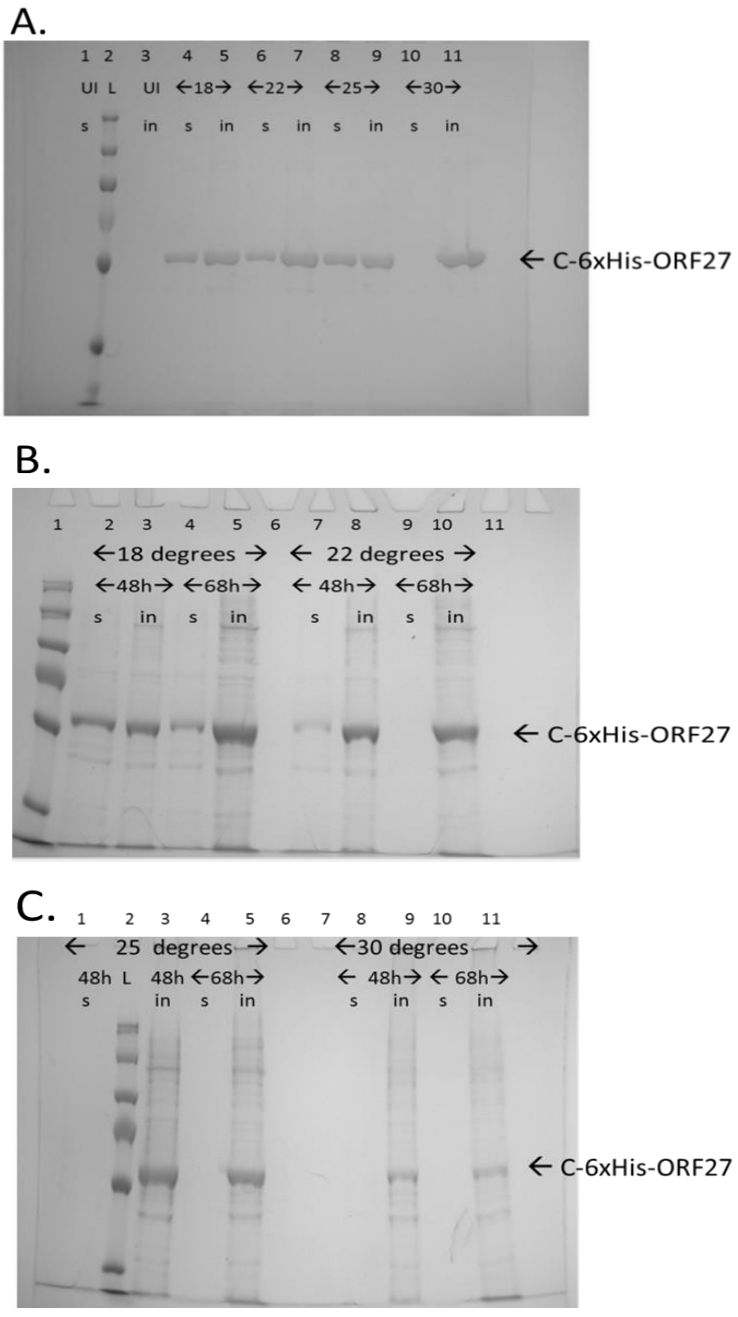


Figure S1.3. Optimization of copy number, temperature and ORF27 solubility for 2-pyrrolidone production from 100 mM GABA. Expression of ORF27 from a low-copy plasmid (pBbS2C-ORF27; strain JZ-298) resulted in lower 2-pyrrolidone production than from a high-copy plasmid (pBbE2C-ORF27; strain JZ-171). Expression of the MBP-ORF27 fusion from a high-copy plasmid (pBbE2C-MBP-ORF27; strain JZ-299) at 25°C gave the optimal 2-pyrrolidone titer. JZ-172 harbors a high-copy plasmid expressing RFP (pBbE2C-RFP).

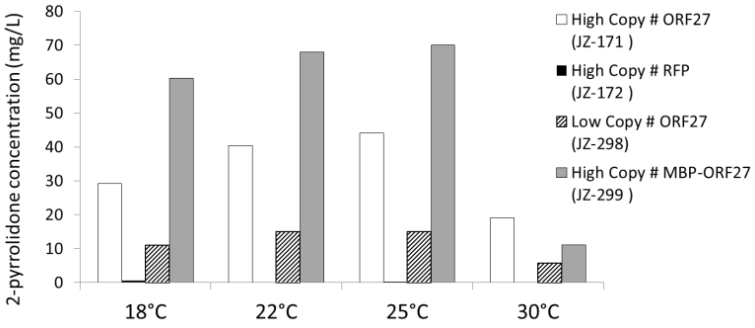
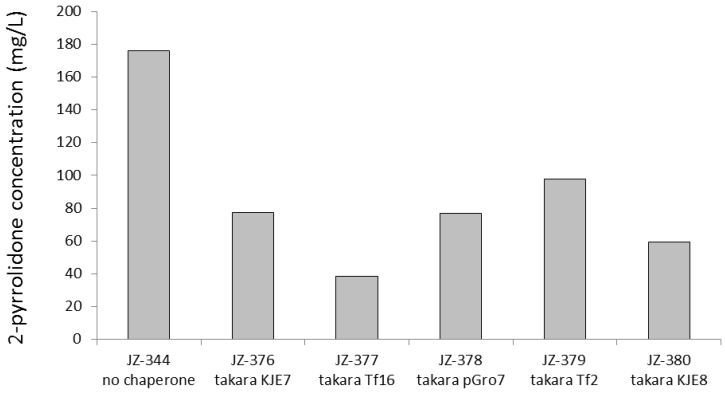


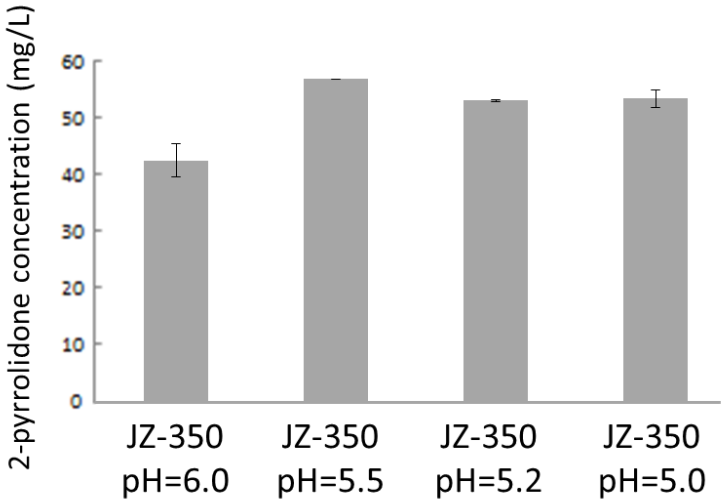
Figure S1.4. Effect of chaperone coexpression on titer when feeding 10 g/L glutamic acid (~68mM) in LB medium. The full set of Takara chaperones were tested (JZ-376 to 380) to determine if chaperone proteins would increase the titer by alleviating MBP-ORF27 aggregation, and in all cases, the titer decreased between 20% to 85% (36, 37).



Utilization of a slow translation ribosome host CH184

CH184 is an *E. coli* strain with slow ribosome translation rate (38). Previous studies showed that it is a good host for expression of multidomain proteins with poor solubilities, presumably a result of co-translational folding of multidomain proteins (38-41). However, its slow growth rate and its inability to improve MBP-ORF27 solubility indicated that CH184 would not be a useful host for 2-pyrrolidone production.

Figure S1.5. pH sensitivity analysis for 2-pyrrolidone production from pH = 5.0 to 6.0.



Cofactor supplementation

PLP is a cofactor for glutamate decarboxylase activity (6, 31, 42); 1 mM, 0.1 mM and 0 mM of PLP supplementation were compared for 2-pyrrolidone production in vivo. PLP supplementation had little impact on 2-pyrrolidone yield; either the first decarboxylation step was not limited by cofactor or glutamate decarboxylation was not the rate limiting step.

Figure S1.6. High density culture 2-pyrrolidone production derived mostly from glutamic acid feeding rather than carbon and nitrogen source in Studier's ZYM-5052 media.

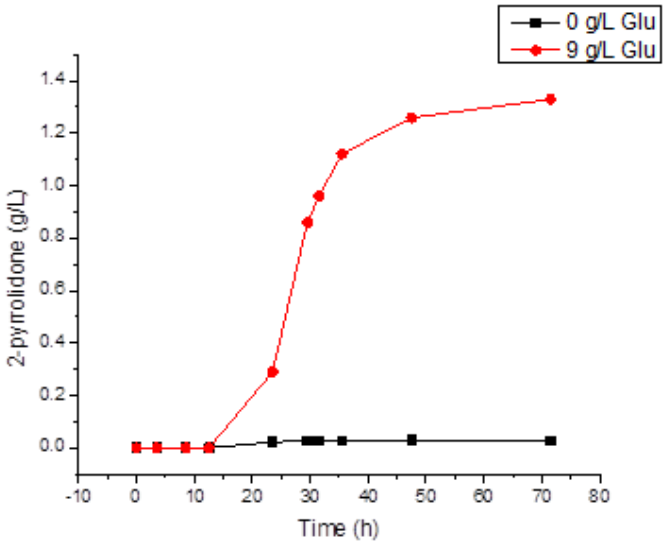
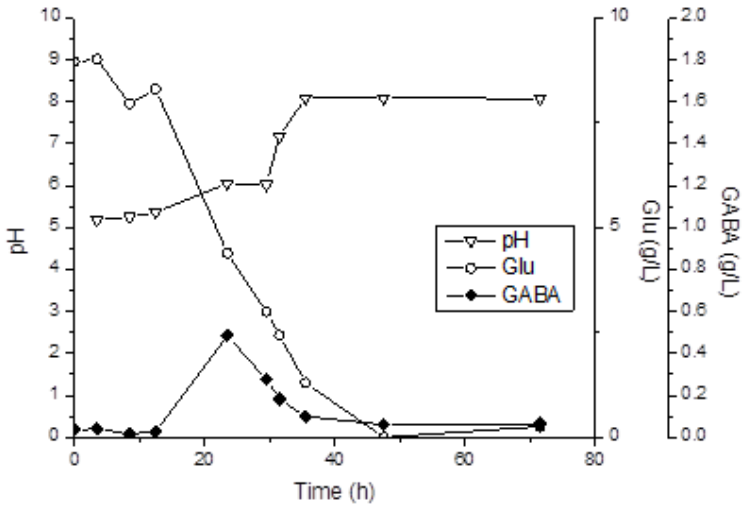


Figure S1.7. pH change (open triangle) during the fermentation process, glutamate consumption (open circle) and GABA intermediate accumulation (closed diamond) .



Supplementary Table 1.1: Some common PKS database websites.

PKS Databases	**website**
NRPS-PKS	http://www.nii.res.in/nrps-pks.html
MAPSI	http://gate.smallsoft.co.kr:8008/pks/mapsidb/
NORINE	http://bioinfo.lifl.fr/norine/
DoBISCUIT	http://www.bio.nite.go.jp/pks/
PKMiner	http://pks.kaist.ac.kr/pkminer/
ClusterMine360	http://www.clustermine360.ca/

Twenty four polyketides with amine starter units were screened using retrobiosynthetic analysis: Vecinistatin, 13-Desmethyl Spirolide C, chejuenolides A and B, oxazolomycins, bacillaene, myxovirescin A, onnamides, virginiamycin M, kirromycin, pederin, etnangien, stellattamide-A, Myxalamids, tolytoxin, salinilactam, MLL-449, hitachimycin, leinamycin, lankacidin, indanomycin, BE-14106, desertomycin A and B, Eco-0501, linearmycin A and Eco-02301 Among them, desertomycin A, linearmycin A and Eco-02301 have a GABA-containing C4 amine starter unit, and linearmycin A has the PKS gene cluster sequence information available (43-46) . Two ORFs in the gene cluster, ORF27 and ORF36, were predicted to be AMP-dependent synthetases (47). ORF36 was highly homologous to a previously identified 5-aminolevulinyl-CoA synthase responsible for 2-amino-3-hydroxycyclopent-2-enone five membered ring formation during Eco-02301 biosynthesis (46). As such, we hypothesized that ORF27 was likely to be involved in activation of the GABA starter unit loading, which made it the best candidate for 2-pyrrolidone synthase.

Supplementary Table 1.2: Other strains, plasmids and oligonucleotides used in the study.

Strains	plasmids	Host
W3110	F- lambda- IN(rrnD-rrnE)1 rph-1	CGSC #4474
CH184	F- lambda- IN(rrnD-rrnE)1 rph-1, rpsL_C256A_C272A	T. Ruusala et al. (1984)
JZ-297	pBbE2C-ORF36	JW2637-4
JZ-316	pBbE2C-MBP-ORF27	W3110
JZ-317	pBbE2C-MBP-ORF27	CH184
JZ-376	takara KJE7(Cm)+ pET28a-MBP-ORF27+ pBbA7a-GadB_ΔHT	BL21 Star (DE3)
JZ-377	takara Tf16(Cm)+ pET28a-MBP-ORF27+ pBbA7a-GadB_ΔHT	BL21 Star (DE3)
JZ-378	takara pGro7 (Cm) + pET28a-MBP-ORF27+ pBbA7a-GadB_ΔHT	BL21 Star (DE3)
JZ-379	takara Tf2 (Cm)+ pET28a-MBP-ORF27+ pBbA7a-GadB_ΔHT	BL21 Star (DE3)
JZ-380	takara KJE8 (Cm) + pET28a-MBP-ORF27+ pBbA7a-GadB_ΔHT	BL21 Star (DE3)

Plasmids		
pDNA2.0-ORF36	pUC ori , KanR ORF36 with Biobrick overhang	DNA 2.0 (Menlo Park, CA)
pG-Tf2	PPzt1, groES-groEL-tig, CmR	Clontech Laboratories, Inc. (Mountain View, CA)
pKJE7	ParaB, dnaK-dnaJ-grpE, CmR	Clontech Laboratories, Inc. (Mountain View, CA)
pTf16	PPzt1, tig, CmR	Clontech Laboratories, Inc. (Mountain View, CA)
pGro7	ParaB, groES-groEL, CmR	Clontech Laboratories, Inc. (Mountain View, CA)
pKJE8	ParaB, dnaK-dnaJ-grpE; PPzt1, groES-groEL, CmR	Clontech Laboratories, Inc. (Mountain View, CA)

Chapter 2 Application of an Acyl-CoA Synthetase from *Streptomyces aizunensis* for Lactam Biosynthesis [#]

2.1 Introduction

Lactams are important compounds used in the manufacture of commercial polymers. ε-Caprolactam (caprolactam) is used in the production of nylon 6 found in fabrics, coatings, plastics of many compositions, lubricants, etc. The worldwide use of nylons requires the production of approximately four million metric tons of caprolactam annually (48). δ-Valerolactam (valerolactam) has been proposed as a monomer for nylon 5 and nylon 6,5 synthesis, addition of which tunes the properties of the resulting polymers (49-51).

Currently, both caprolactam and valerolactam are synthesized from starting materials extracted from petroleum. Caprolactam production starts from cyclohexanone, which is first converted to its oxime. Treatment of this oxime with acid induces the Beckmann rearrangement to give caprolactam (52). Such production involves energy intensive processes and harsh acidic reaction conditions and produces large amount of waste salts. On the other hand, due to the lack of raw five-carbon petrochemical feedstocks, valerolactam is still too costly to achieve wide adoption for nylon synthesis.

Unlike chemical dehydration, enzymatic or whole-cell-catalyzed reactions can be performed at lower temperature and pressure. Although several ω-amino fatty acids have been biosynthesized (50, 53, 54), full biosynthetic pathways to produce lactams are largely unknown. This is due to a lack of enzymes capable of performing the last ring closing step. In terms of polymer chemistry, ring open polymerization of these lactam monomers is preferred over condensation of their corresponding ω-amino fatty acids, because condensation chemistry generates water during each monomer addition, and adversely impacts polymer properties such as molecular weight and polydispersity, and leads to undesirable thermal and mechanical properties. To date, only one enzyme, *Candida antarctica* lipase B (CALB, commercially available asN435), was reported to conduct a reversible aminolysis reaction that can be utilized for valerolactam and caprolactam synthesis (55). However, the reported enzymatic reaction occurs under vacuum over P_2O_5, and requires high temperature and long reaction times to overcome the energy barrier of lactam formation (70% conversion at 90°C and 20% conversion at 55°C over a three-day period). Also, the intermolecular aminolysis reaction results in multiple side products, including macrocyclic dimer and trimer lactams, which are hard to eliminate during product purification. As such, there is no suitable enzyme capable of synthesizing industrially important lactams under microbial fermentation conditions.

Previously, ORF27 from *Streptomyces aizunensis*, was hypothesized to be either a 4-aminobutyryl-CoA synthetase or a 4-guanidinobutyryl-CoA synthase (43, 45, 46, 56), although we demonstrated that expression of ORF27 in vivo resulted in the formation of γ-butyrolactam from 4-aminobutyric acid(57). In this study, we overexpressed ORF27 in *Escherichia coli* and purified it by affinity chromatography. This enabled us to explore the catalytic properties of this enzyme. ORF27 was confirmed to possess 4-guanidinylbutyryl-CoA ligase (4GBL) activity.

[#] Reproduced with permission from Jingwei Zhang, Mehmet Burdu, George Wang, Edward E.K. Baidoob, Jay. D. Keasling. "Application of an Acyl-CoA Synthetase from *Streptomyces aizunensis* for Lactam Biosynthesis" (manuscript in preparation)

ORF27 had broad substrate specificity: it could act on linear or branched acid substrates with positively charged or neutral functional groups on the ω-terminal end, yet the enzyme was selective against negatively charged groups on the substrate's ω-terminal end. Intriguingly, ORF27's activity for ω-amino fatty acids activation led to its application as a general lactam synthase, enabling biosynthesis of caprolactam, valerolactam and butyrolactam. Comparative study of ORF27's reaction products for different ω-amino fatty acid precursors was performed to better understand this enzyme's activity as a lactam synthase.

To apply ORF27 for renewable chemical production, ORF27 was overexpressed in *E. coli*, and both valerlactam and caprolactam were formed in vivo by feeding their respective precursors, 5-aminovaleric acid (5-AVA) and 6-aminohexanoic acid (6-AHA). To achieve renewable production of valerolactam from sugar, we introduced a two-gene pathway into *E. coli* that converts lysine to 5-AVA. The pathway contains an L-lysine monoxygenase (*davB* from *Pseudomonas putida* KT2440) and a 5-aminovaleramide amidohydrolase (*davA* from *Pseudomonas putida* KT2440. Introducing the genes encoding the 5-AVA biosynthetic pathway and a gene encoding a fusion of ORF27 with maltose binding protein, MBP-ORF27, enabled *E. coli* to produce valerolactam from lysine or directly from glucose.

2.2 Materials and Methods

2.2.1 Strains and plasmids

All the strains and plasmids utilized in this study are listed in Table 2.1. The sequences for the plasmids are listed in Supplementary Material S10.

Table 2.1. *E. coli* strains, plasmids and oligonucleotides used.

Name	Relevant genotype	Reference
Host Strains		
DH10B	F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ-rpsL nupG	Life Technologies (Carlsbad, CA)
BL21 Star (DE3)	F− ompT hsdSB (rB− mB−) gal dcm rne131 (DE3)	Life Technologies (Carlsbad, CA)
JW2637-4	F-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ-, ΔgabT743::kan, rph-1, Δ(rhaD-rhaB)568, hsdR514	CGSC #11775

Engineered Strains	plasmids	Host
JZ-171	pBbE2C-ORF27	JW2637-4
JZ-172	pBbE2C-RFP	JW2637-4
JZ-440	pBbA7a-DavB-DavA	BL21(DE3) star
JZ-441	pBbA7a-DavB-DavA + pET28a-MBP-ORF27	BL21(DE3) star

Constructed Plasmid	**Backbone Source (restriction site)**	**Gene Source (direct digestion or PCR)**	**PCR primers**
pBbE2C-ORF27	pBbE2C-RFP (BglII, XhoI)	pDNA2.0-ORF27 (BglII, XhoI)	N/A
pET28b-N-ORF27	pET28b (NdeI, XhoI)	pDNA2.0-ORF27 (NdeI, XhoI)	N/A
pET28b-C-ORF27	pET28b (NcoI, XhoI)	pDNA2.0-ORF27 PCR	JZ_27_C6xHis_f, JZ_27_C6xHis_r
pET28a-MBP-ORF27	pET28a-MBP (NdeI, XhoI)	pDNA2.0-ORF27 (NdeI, XhoI)	N/A
pBbA7a-DavB-DavA	pBbA7a-RFP (BglII, XhoI)	DavB, DavA gBlock	N/A

Oligonucleotides	5'→ 3' Sequence, restriction site underlied, synthesized by Integrated DNA Technologies, Inc (Coralville, IA)	Target Gene
JZ_27_C6xHis_f	GCGCGccatgg gc ATGCGCCCAATGACCGCTAAAATCTTCG	C-6xHis-ORF27
JZ_27_C6xHis_r	GCGCGctcgagTTCTGCCGCCATACGGGTCAGC	C-6xHis-ORF27

2.2.2 ORF27 Protein Expression and Purification

The gene encoding ORF27 (GenBank: AAX98201.1) was purchased from DNA 2.0 (Menlo Park, CA). The synthetic gene was optimized with *E. coli* codon usage and delivered in pDNA2.0-ORF27. For expression and purification of ORF27, the pDNA2.0-ORF27 plasmid was digested with NdeI and XhoI and cloned into pET28b in order to produce ORF27 with an N-terminal 6xHis tag. The resulting plasmid, pET28b-N-ORF27, was transformed into *E. coli* BL21 Star (DE3) for ORF27 overexpression. To place a C-terminal His-tag on ORF27, primers JZ_27_C6xHis_f and JZ_27_C6xHis_r were used to amplify ORF27 from pDNA2.0-ORF27. The resulting PCR product was digested with NcoI and XhoI and cloned into pET28b.
For N-terminal 6xHis ORF27 expression, the overnight culture was inoculated (1:100 v/v) into 1 L LB medium containing 5 μg/ml kanamycin. The culture was grown at 37°C until the O.D. reached 0.6 and cooled on ice for 20 min. 0.5 mM IPTG was added to induce N-6xHis ORF27overexpression for 16 h at 18°C. The cells were harvested by centrifugation (8000 × g, 6 min, 4°C), resuspended in 30 mL of lysis buffer (50 mM HEPES, pH 8.0, 0.5 M NaCl, and 10 mM imidazole), and lysed by sonication on ice. Cellular debris was removed by centrifugation (20,000 × g, 30 min, 4°C). Ni-NTA agarose resin was added to the supernatant (1 mL/L of culture), and the solution was rocked at 4°C for 1 h. The protein resin mixture was loaded onto a gravity flow column, and proteins were washed with washing buffer (50 mM HEPES, pH 8.0, 0.5 M NaCl, and 20 mM imidazole) and eluted with elution buffer (50 mM HEPES, pH 8.0, 0.5 M NaCl, and 250 mM imidazole). Purified proteins (60 mg from 1 L culture) were concentrated to 280 mg/mL and buffer exchanged into storage buffer (50 mM HEPES, pH 8.0, 8% glycerol). The final proteins were aliquoted and flash frozen in liquid nitrogen and stored at -80°C. C-terminal 6xHis ORF27 (65 mg/L LB culture) was produced using BL21 Star (DE3) transformed with pET28b-C-ORF27, purified by Ni-NTA agarose resin and stocked at 220 mg/mL in storage buffer (50 mM HEPES, pH 8.0, 8% glycerol).

Nickel nitrilotriacetic acid agarose (Ni-NTA) resin and SDS-PAGE gels were purchased from Qiagen and Biorad, respectively. Protein samples were concentrated using 10 KDa MMCO Amicon Ultra filters (Millipore). DNA and protein concentrations were determined using a Nanodrop 1000 spectrophotometer (Thermo Scientific). The purified ORF27 was checked on SDS-PAGE gel for purity (Fig. S2.1).

2.2.3. ATP-PPi release Assays for ORF27

The substrate range of ORF27 was determined by ATP-PPi release assays as previously described (46). Without prior knowledge about the reaction pathways, the ATP-PPi assay served a semi-quantitive method to compare substrate induced acceleration of ATP consumption. For kinetic investigation of ORF27 activity with different substrates, the inorganic pyrophosphate

released by enzymatic reaction was measured continuously using the EnzChek Pyrophosphate Assay Kit (Invitrogen). A typical assay contained in a total volume of 150 μL: 5 μM of ORF27, 0-20 mM substrates, 1 mM ATP, 1 mM CoASH and 1 mM $Mg(Cl)_2$ in 100 mM HEPES, pH 7.5. 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) substrate, purine nucleoside phosphorylase and inorganic pyrophosphatase were added according to the protocol. Reactions were initiated by the addition of ATP and monitored at 360 nm with SpectraMax M2 (Molecular Devices, Sunnyvale, CA). Initial velocities were calculated using the standard curve for inorganic pyrophosphate. For each concentration, control reactions were carried out without enzyme or without ATP. The rates of PPi release were converted to observed rates, and the Michaelis-Menten kinetic parameters were obtained.

The acids assayed were glutamic acid, γ-aminobutyrate (GABA), 4-guanidinobutyric acid, (S)-3-hydroxyl-butyric acid, valeric acid, 4-methyl-hexanoic acid, 3-aminobutyric acid, 6-aminocaproic acid, 6-guanidinohexanoic acid, glutaric acid, adipic acid, 2-aminobutyric acid and (Sigma-Aldrich, St. Louis, Missouri). During preparation of substrate stock solution, 6-guanidinohexanoic acid has low solubility under neutral pH, and HCl was added to obtain 100 mM stock solution in pH ~2.0.

2.2.4. 4-Guanidinobutyryl CoA product identification

To confirm ORF27's native activity as a 4-guanidinobutyryl CoA synthetase, the quenched reaction with 4-guanidinobutyric acid was analyzed for CoA products as described previously (58). The mass measurements were carried out in the TOF-Scan monitoring mode for the detection of $[M - H]^-$ ions (4-guanidinobutyrl-CoA, m/z = 893.1825). Negative controls were carried out using no enzyme, no ATP, no substrate or no CoASH.

2.2.5. Product analysis of in vitro ORF27 lactam formation

To compare formation of various lactams by ORF27, a reaction mixture containing 57 μM of ORF27, 5 mM ω-amino fatty acids substrates, 1 mM ATP or ADP, 0.5mM CoASH and 1 mM $Mg(Cl)_2$ in 100 mM HEPES (pH=8) was incubated at 25°C. The reactions were quenched by addition of methanol to a final concentration of 50% (v/v) at multiple time points (0 min, 15 min, 1 h, 2 h, 4 h and 19 h). The resulting quenched reactions were kept at 4°C and filtered through 10 K Amicon Ultra-0.5 mL Centrifugal Filters (Millipore) at 8000 x g for 30min. The filtered solutions were analyzed for lactams and nucleotides using HPLC-MS (see Supplementary Material S2, S3). Control reactions were carried out without enzyme, without substrate, without ATP, without CoASH or without $MgCl_2$ (supplying additional 0.2 mM EDTA to chelate Mg^{2+} ions from purified protein stock). Caprolactam, valerolactam and butyrolactam, AMP, ADP and ATP were purchased as standards from Sigma-Aldrich. The pH dependence of ORF27-catalyzed butrylactam formation was determined using an end-point assay, and the amount of butyrolactam was determined using the HPLC-MS method described above.

2.2.6. Lactam Production in Vivo

E. coli JW2637-4 contains a knockout of *gabT*, which encodes a GABA transaminase (29). This host was initially used to confirm production of valerolactam and caprolactam in vivo. *E. coli*

JZ-171 (ORF27) and JZ-172 (RFP negative control) were grown in LB medium containing 1 mM 5-AVA and 6-AHA. Cultures were propagated in Luria-Bertani (LB) medium, which was prepared from dehydrated powder according to the manufacturer's instructions (BD Biosciences, San Jose, CA). Tl analyze caprolactam production, the culture was pelleted and supernatant was filtered before mixing with 1 volume of MeOH.

2.2.7. In vivo valerolactam biosynthesis

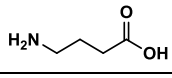
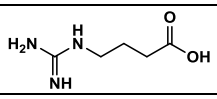
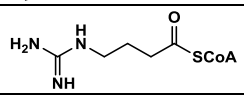
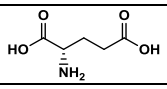
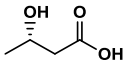
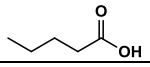
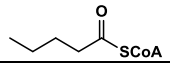
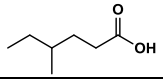
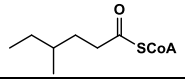
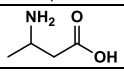
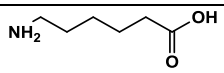
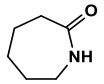
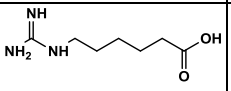
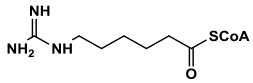
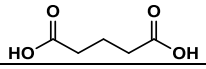
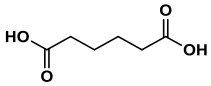
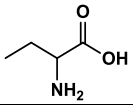
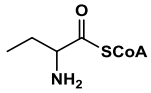
The *davA* and *davB* genes were ordered as gBlocks (Integrated Dna Technologies, Coralville, IA) and cloned into the BglII and XhoI sites on pBbA7, a biobrick vector, to generate plasmid pBbA7a-DavB-DavA(28). To circumvent ORF27's limited solubility during incubation, MBP-ORF27 was utilized(57). For high-density shake flask cultures, Studier's autoinduction ZYM-5052 medium was prepared according to the published protocol (27). Lysine at various concentrations (0 g/L, 1 g/L, 5 g/L and 10 g/L) was included in the ZYM-5052 medium. Kanamycin (20 μg/ml) and ampicillin (100 μg/ml) were added where desired to provide selective pressure for plasmid maintenance.

E. coli strains (JZ-441) harboring plasmids containing genes encoding *davA, davB* and *MBP-ORF27* were inoculated into 10 mL of LB overnight. On day 2, the overnight culture was inoculated 1:100 (v/v) into 25 mL Studier's autoinduction ZYM-5052 medium with various concentrations of lysine (0 g/L to 10 g/L) and appropriate antibiotics (27). The culture was incubated at 37°C. When the O.D. reached around 0.6, the culture was cooled to 25°C. The culture was then placed at 25°C incubator and the valerolactam titer was analyzed at 24 h, 48 h and 72 h. JZ-440, which contains only *davA* and *davB*, served as a negative control.

2.3. Results

2.3.1. ORF27 Biochemistry

The sensitivity of the ATP-PPi release assay (5μM Pi) allowed the initial velocity to be determined within the first 2min of reaction incubation. During this time period, ATP hydrolysis activity by ORF27 was not observed according to the ATP-PPi assay (ATP hydrolysis was observed over longer period of time though, see discussion below). The ATP-PPi release assay revealed that ORF27 has a broad substrate specificity. The K_m and k_{cat} values of various substrates activation are shown in Table 2.2.

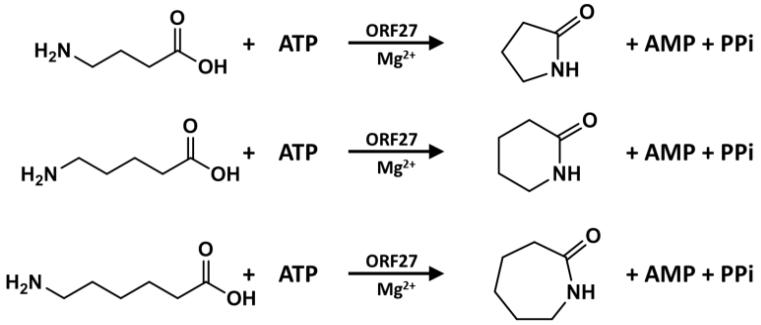
Table 2.2. Steady State Parameters of ORF27[a]					
Substrate	Structure	Product [b]	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m ($M^{-1}min^{-1}$)
γ-Aminobutyrate			17.2	0.35	20.26
4-Guanidinobutyric acid			0.013	0.12	9434
Glutamate [c]		ND	ND	ND	ND
(S)-3-hydroxyl-butyric acid		ND	ND	ND	ND
Valeric acid			52.2	1.70	32.6
4-Methyl-hexanoic acid			12.1	1.3	107.8
3-Aminobutyric acid		ND	ND	ND	ND
6-Aminocaproic acid			0.056	0.17	2962.6
6-Guanidinohexanoic acid			0.11	0.33	2939.7
Glutaric acid		ND	ND	ND	ND
Adipic acid		ND	ND	ND	ND
2-Aminobutyric acid			1.63	0.036	22.2

a The kinetic parameters of ORF27 were determined in a 100 mM HEPES, pH 7.5 at 25°C using the ATP-PPi release Assays
b Other non-product forming pathways could occur such as ATP hydrolysis or acyl-OAMP intermediate hydrolysis.
c ND not determined.

HPLC-MS confirmed that ORF27 catalyzes 4-guanidinobutyryl-CoA formation (Fig. S2.4). The ability of ORF27 to accept a broad range of substrates, especially ω-amino fatty acids, implies an interesting biotechnology application. Once the acid group of ω-amino fatty acids has been activated, amide bond formation through intramolecular cyclization is theoretically thermodynamically favored, suggesting that ORF27 might be utilized as a lactam synthase to

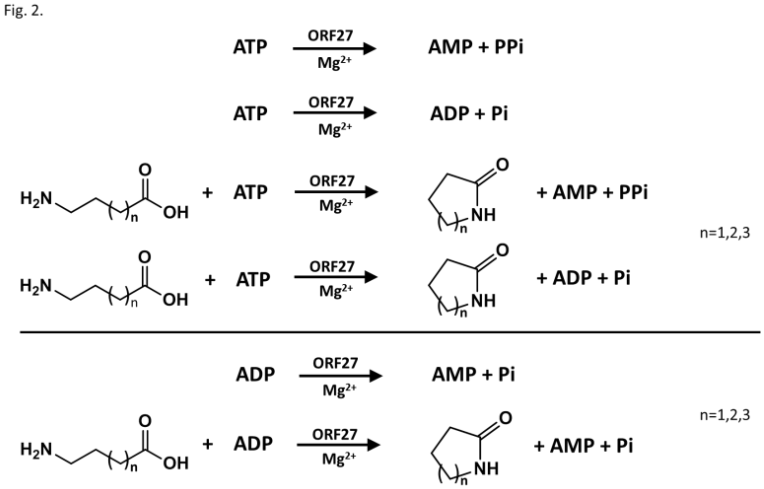
produce industrially important chemicals, such as caprolactam, valerolactam and butyrolactam (Fig. 2.1).

Figure 2.1. Proposed butyrolactam, valerolactam and caprolactam formation from their respective linear substrates catalyzed by ORF27.



The ORF27-catalyzed lactam formation reaction in vitro was monitored for both nucleotide and lactam products, using either ATP or ADP as the energy source. LC-MS analysis of nucleotides (ATP, ADP, AMP) showed that ORF27 catalyzed a series of reactions (Fig. 2.2). In the absence of ORF27, ATP remained stable throughout the reaction (Fig. S2.5A). In the absence of ω-amino fatty acids, ORF27 predominantly hydrolyzed ATP over long term incubation, forming mainly ADP plus Pi and slightly AMP and PPi (Fig. S2.5B). This seeming contradiction with previous negative control ATP-PPi assay result was because: due to the slow kinetics of the in vitro reaction, ATP-PPi assay was not sensitive enough to detect Pi released caused by enzymatic ATP hydrolysis into ADP and Pi during the first 2 minutes(Fig. S2.5B). When ω-amino fatty acids were added into the reaction mixture, they facilitate ATP hydrolysis, as well as routing hydrolysis pathway towards forming AMP and PPi rather than ADP plus Pi. Therefore, the obsevered signal increase in the presence of substrate for the ATP-PPi assay was due to both increased ATP consumption, and having more PPi release instead of Pi release without ω-amino fatty acid substrates.

Figure 2.2. ORF27 catalyzes multiple reaction pathways.



Although ORF27 was predicted to be an acyl-CoA synthetase, CoASH was not required for lactam formation (Fig. S2.5C-S2.5H). The minimal lactam formation system constituted ω-amino fatty acids, ORF27, ATP and Mg^{2+} (Fig. S2.6). Lactam formation is most rapid for valerolactam, followed by butyrolactam and caprolactam. ADP appeared earlier in in reactions containing ω-amino fatty acids but not in control reactions without the substrates , suggesting ω-amino fatty acids dependent ATP hydrolysis into ADP and Pi also occurred during the reaction. When CoASH was added to the corresponding reactions, no 4-aminobutyryl-CoA or 5-aminovaleryl-CoA was detected and barely distinguishable 6-aminohexanoyl-CoA mass ions, corresponding to 0.6% of CoASH signal abundance, were detected (data not shown). Mass ions corresponding to other off-pathway products, such as the ω-amino fatty acids dimers or trimers, were not observed.

In additional to using ATP, ORF27 could also utilize ADP to activate ω-amino fatty acids and catalyze lactam formation, although the reaction occurred at only 20-50% of the rate of the similar reaction when ATP was the substrate (Fig. S2.5I-S2.5K).

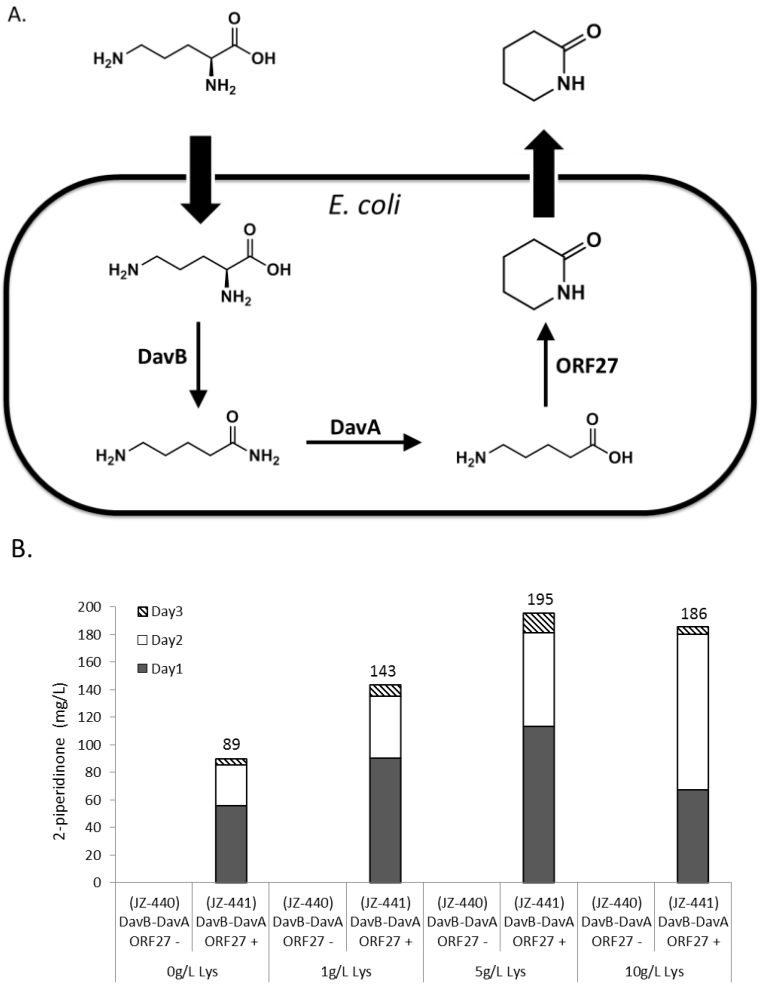
The pH profile of ORF27 activity was determined for butyrolactam, and the enzyme had a pH optimum of 8.0. The enzyme precipitated and became inactive when the pH dropped below 6.0 (Fig. S2.7).

2.3.2. Valerolactam and caprolactam production in vivo

To demonstrate that valerolactam or caprolactam can be produced using whole-cell catalysts, we fed the respective precursors, 5-AVA or 6-AHA, to *E. coli* JZ-171, which overexpressed ORF27. Valerolactam (5-AVA fed) or caprolactam (6-AHA fed) was observed in the medium upon ORF27 expression (Fig. S2.8).

To demonstrate that valerolactam could be produced directly from glucose with no ω-amino fatty acid feeding, we introduced a pathway for 5-AVA biosynthesis into *E. coli* BL21(DE3)star, resulting in strain JZ-441(49-51). Valerolactam was produced directly from lysine when *davA* and *davB* were introduced into *E. coli* (Fig. 2.3A). Cells growing in medium supplemented with 5 g/L of lysine produced a maximum of 195 mg/L (~2 mM) valerolactam. Increasing lysine supplementation to 10 g/L caused a reduction in bacterial growth and caused a delayed and decreased production of valerolactam. At 0 g/L lysine feeding, JZ-441 produced 89 mg/L of valerolactam (Fig. 2.3B). JZ-440, which lacks MBP-ORF27, produced no measurable valerolactam.

Figure 2.3. Valerolactam production in recombinant *E. coli*. A) Valerolactam biosynthesis pathway. B) Production of valerolactam via fermentation. Titer gain by day 1-3 after induction in strains JZ-440 (ORF27 negative control) and JZ-441 from 0g/L, 1g/L, 5g/L and 10g/L lysine feeding.



2.4. Discussion

Natural product biosynthesis continues to be a rich source of enzyme candidates with novel activities. ORF27, an enzyme in the *Streptomyces aizunensis* ECO-02301 biosynthetic cluster, was identified to be a 4-guanidinobutyryl-CoA synthetase, which has interesting implications for the ECO-02301 loading mechanism (Supplementary material S9) (46, 59, 60).

ORF27 accepted a wide range of short chain fatty acid substrates and their functionalized analogs. Both linear and branched fatty acid substrates were accepted as substrates. However, the enzyme seems to have little tolerance of polar groups at C3 position. Positively charged substrates, such as the ω-amino fatty acids or the ω-guanidino fatty acids are well tolerated, despite the fact that some substrates such as 6-guanidinohexanoic acid have two more carbons in their backbone than the native substrate. However, enzyme activity on substrates that have a

negatively charged group on the ω-terminal end, such as a carboxylic acid group, was not observed, even though glutaric acid and adipic acid have similar steric hindrance as 6-guanidinohexanoic acid. This suggests that ORF27 has strict substrate selection residues around it binding pocket and preferentially favors substrates with a positively charged group on the ω-terminal end.

The ability of ORF27 to accept various ω-amino fatty acids was explored for lactam biosynthesis. For lactam formation, ORF27 does not require CoASH as substrate and utilizes multiple reaction pathways. The activation of ω-amino fatty acids by ATP facilitates cyclization. This enables five-membered, six-membered and even seven-membered ring formation at mild temperatures, resulting in the production of important industrial lactams such as valerolactam and caprolactam via fermentation. Unlike the reversible aminolysis enzyme CALB that was previously described, ORF27 performs the reaction under milder condition in an irreversible fashion. The lactam product observed is exclusively cyclized monomers, without dimer or trimer contaminants. These novel features make ORF27 an ideal candidate for lactam biosynthesis. However, the enzyme catalyzed significant ATP and ADP hydrolysis during the reaction. Directed evolution of ORF27 towards less futile ATP and ADP hydrolysis could potentially improve ORF27 as a lactam synthase.

By introducing *davB* and *davA* into *E. coli*, we demonstrated valerolactam biosynthesis from lysine. Valerolactam production in autoinduction medium was lysine dependent. At concentrations of lysine below 5 g/L, valerolactam production correlated with lysine concentration. However, high concentrations of lysine (e.g., 10 g/L) inhibited bacterial growth and resulted in decreased valerolactam production. Even without lysine feeding, direct lysine production from glucose was sufficient to produce 50% of the valerolactam produced when 5 g/L lysine was fed to the culture. Interestingly, in the absence of ORF27, a slight amount of valerolactam was observed. Since 5-AVA does not cyclize when incubated in medium at 25°C (data not shown), he observed valerolactam could result from catalysis by innate *E. coli* enzymes. Caprolactam was also biosynthesized by feeding its precursor 6-aminohexanoic acid. With the recent publication of 6-aminohexanoic acid biosynthetic pathway in *E. coli*, ORF27 holds the promise to enable full caprolactam biosynthesis in a microbial host (54). This opens the door to substitute the current petrochemical process with a renewable bioprocess for a bulk chemical with four million ton annual production, potentially could make huge positive impact on our environment.

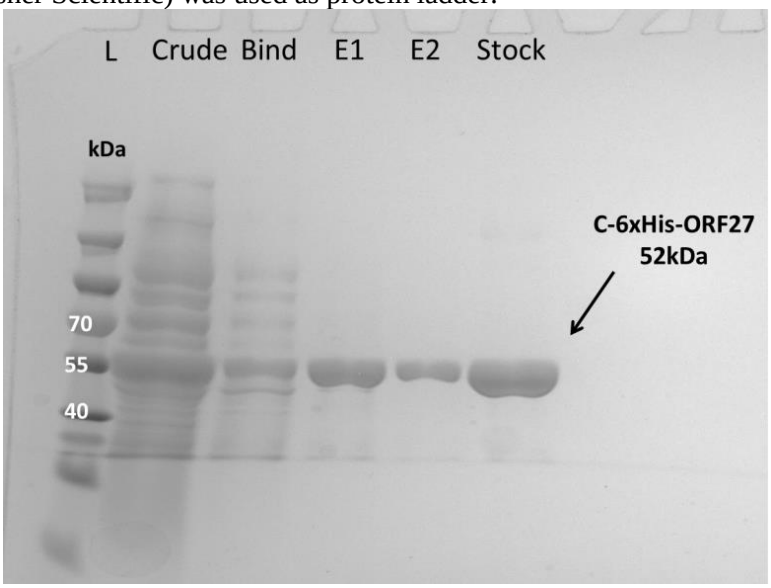
2.5. Conclusion

ORF27 is an enzyme in the ECO-02301 biosynthetic cluster. It was identified to catalyze 4-guanidinobutyryl-CoA formation, and was able to activate a wide range of substrates. While acting on ω-amino fatty acids, ORF27 can catalyze the ring closing reaction to produce lactams. Biosynthesis of polymer precursors such as caprolactam, valerolactam and butyrolactam were demonstrated from their respective ω-amino fatty acids. For in vivo production of valerolactam from lysine or glucose, the lysine to 5-AVA pathway was introduced into an *E. coli* strain harboring the gene encoding ORF27.

Supplemental Material

Supplementary Figures

Figure S2.1. SDS-PAGE of ORF27 purified from *E. coli*. Ready gel for Tris-Glycine Gel (10% precast, Bio-Rad) was used. For molecular weight determination, PageRuler™ Prestained Protein Ladder (10 to 180 kDa, ThermoFisher Scientific) was used as protein ladder.



S2.2. Caprolactam, valerolactam and butyrolactam analytical method

Liquid chromatography (LC) separation of lactams were conducted at 55°C with an Inertsil ODS-3 reverse-phase C18 column (250 mm length, 2.1 mm internal diameter, 3 μM particle size; GL Sciences) using a 1100 series high-performance LC system (Agilent Technologies). The mobile phase was composed of 0.1% formic acid in H_2O (solvent A) and 0.1% formic acid in MeOH (solvent B). Butyrolactam was separated with the following gradient: 40% to 60% B for 4.5 min, 60% to 100% B for 0.5 min, 100% to 40% B for 0.5 min, held at 10% B for 8.5 min. A flow rate of 0.18 mL/min was used throughout.

Time-of-Flight Mass Spectrometry method for lactam accurate mass monitoring

The LC system was coupled to an Agilent Technologies 6210 electrospray time-of-flight (TOF) mass spectrometer. 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 11 L/min and 25 psig, respectively, and a drying gas temperature of 320°C was used throughout. ESI was conducted in the positive-ion mode with a capillary voltage of 3.5 kV. Mass measurements were carried out in the TOF-Scan monitoring mode for the detection of $[M + H]^+$ ions (2-pyrrolidone, m/z=86.0600; 2-piperidinone, m/z= 100.07569; caprolactam, m/z= 114.09134). The instrument was tuned for a range of m/z 70 to 300. Data acquisition and processing were performed using MassHunter Workstation (Agilent Technologies).

S2.3. ATP, ADP, AMP analytical method

Nucleotide product analysis studies was analyzed using an expedited modification of the HILIC method previously described(61). Liquid chromatography (LC) separation of lactams were conducted at 40°C

with an SeQuant Zic-pHILIC column (150 mm length, 2.1 mm internal diameter, 5 µM particle size; GL Sciences) using a 1100 series high-performance LC system (Agilent Technologies). The mobile phase was composed of 50mM $(NH_4)_2CO_3$ in H_2O (solvent A) and acetonitrile (solvent B). ATP, ADP, AMP and CoAs were separated with the following gradient: 73% to 43% B for 6 min (flow rate 0.25mL/min), 43% to 73% B for 0.2 min(flow rate 0.25mL/min), hold at 73% B for 1 min(flow rate 0.3mL/min), hold at 73% 0.2 min(flow rate 0.3mL/min), held at 73% B for 5.1 min (flow rate of 0.38mL/min). The HPLC system was coupled to TOF MS. ESI was conducted in the negative ion mode, and a capillary voltage of 3.5 kV was utilized. Fragmentor, skimmer, and OCT1 RF voltages were set to 200 V, 65 V, and 300 V, respectively. Mass measurements were carried out in the TOF-Scan monitoring mode for the detection of $[M - H]^-$ ions (ATP, m/z=505.9885; ADP, m/z= 426.0221; AMP, m/z= 346.0558; CoASH, m/z=766.1079). MS experiments were carried out in the full-scan mode (m/z 100 to 1000).

Figure S2.4. LC–MS analysis of 4-guanidinobutyrl-CoA formation catalyzed by ORF27. (A) 4-guanidinobutyrl-CoA synthetase reaction. (B) LC-MS confirmed that the reaction system requires ORF27, 4-guanidino butyric acid (4-GBA) and CoASH for 4-guanidinobutyryl-CoA (4GB-CoA) product formation.

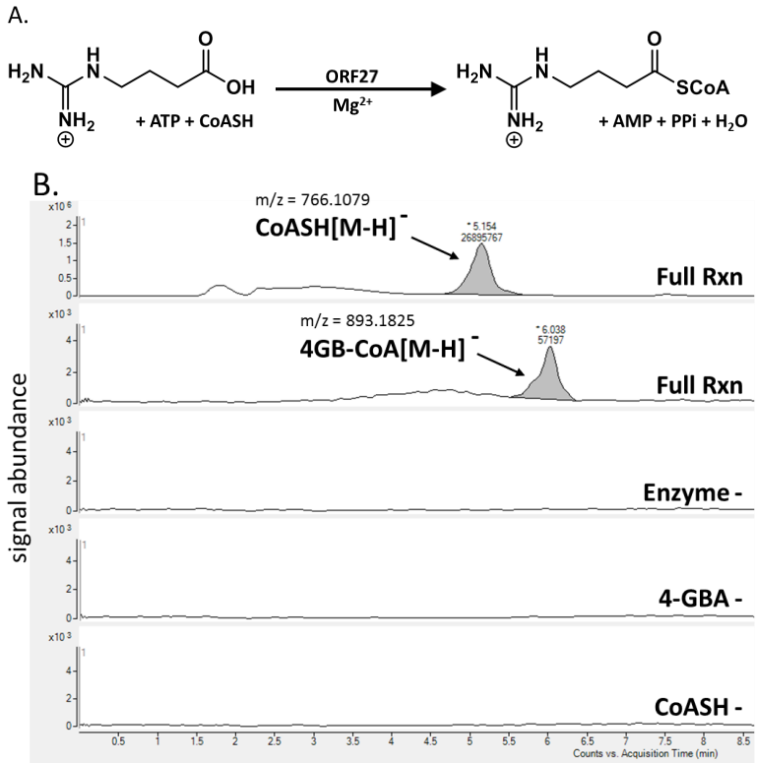
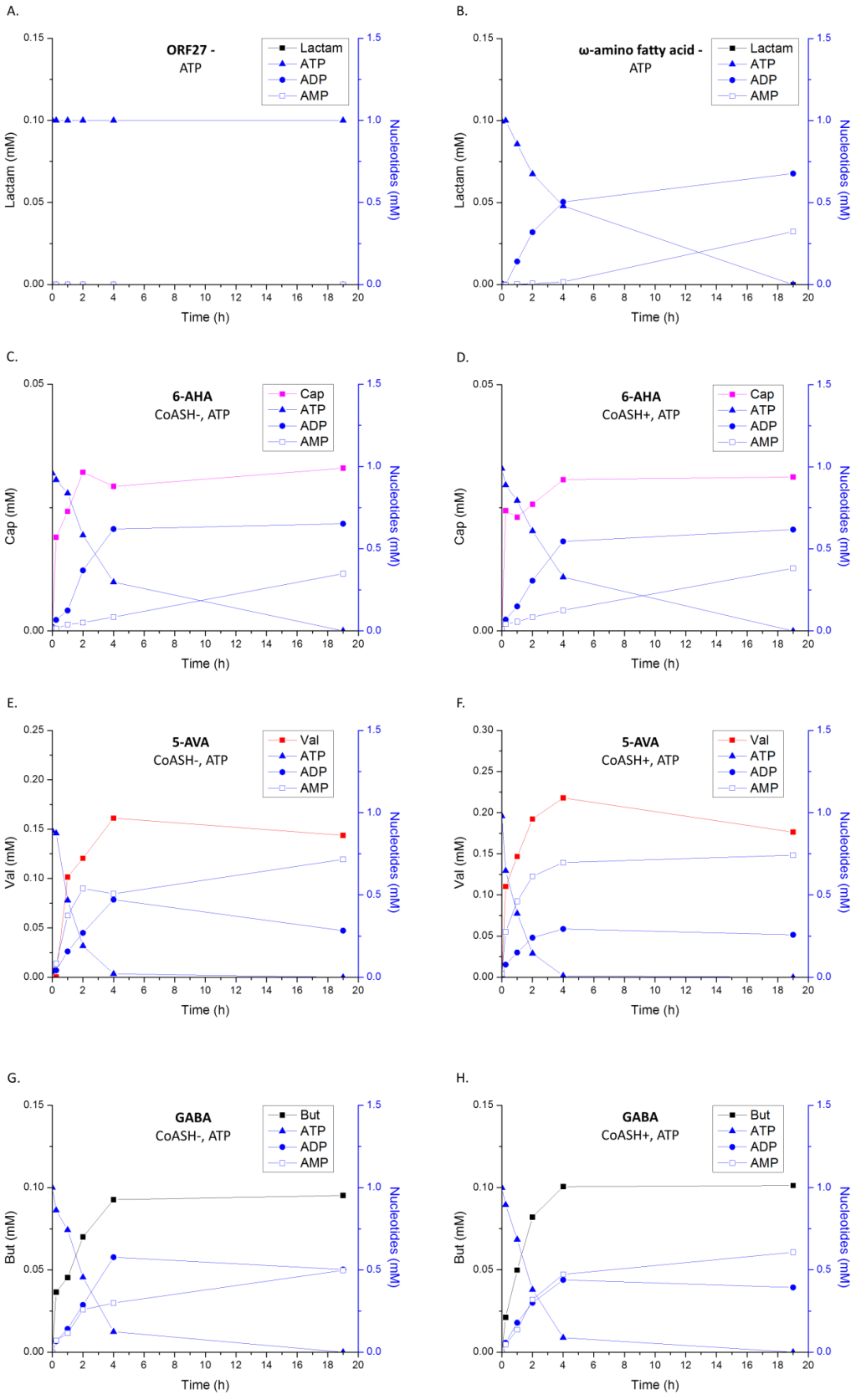


Figure S2.5. Nucleotide and lactam product analysis of ORF27 catalyzed lactam formation. A) ORF27 minus, ATP for activation, CoASH plus. B) ω-amino fatty acid minus, ATP, CoASH plus. C) 6-AHA, ATP, CoASH minus. D) 6-AHA, ATP, CoASH plus. E) 5-AVA, ATP, CoASH minus. F) 5-AVA, ATP, CoASH plus. G) GABA, ATP, CoASH minus. H) GABA, ATP, CoASH plus. I) 6-AHA, ADP, CoASH plus. J) 5-AVA, ADP, CoASH plus. K) GABA, ADP, CoASH plus.


A.
ORF27 -
ATP
Lactam
ATP
ADP
AMP
Lactam (mM)
Nucleotides (mM)
Time (h)
B.
ω-amino fatty acid -
ATP
Lactam
ATP
ADP
AMP
Lactam (mM)
Nucleotides (mM)
Time (h)
C.
6-AHA
CoASH-, ATP
Cap
ATP
ADP
AMP
Cap (mM)
Nucleotides (mM)
Time (h)
D.
6-AHA
CoASH+, ATP
Cap
ATP
ADP
AMP
Cap (mM)
Nucleotides (mM)
Time (h)
E.
5-AVA
CoASH-, ATP
Val
ATP
ADP
AMP
Val (mM)
Nucleotides (mM)
Time (h)
F.
5-AVA
CoASH+, ATP
Val
ATP
ADP
AMP
Val (mM)
Nucleotides (mM)
Time (h)
G.
GABA
CoASH-, ATP
But
ATP
ADP
AMP
But (mM)
Nucleotides (mM)
Time (h)
H.
GABA
CoASH+, ATP
But
ATP
ADP
AMP
But (mM)
Nucleotides (mM)
Time (h)

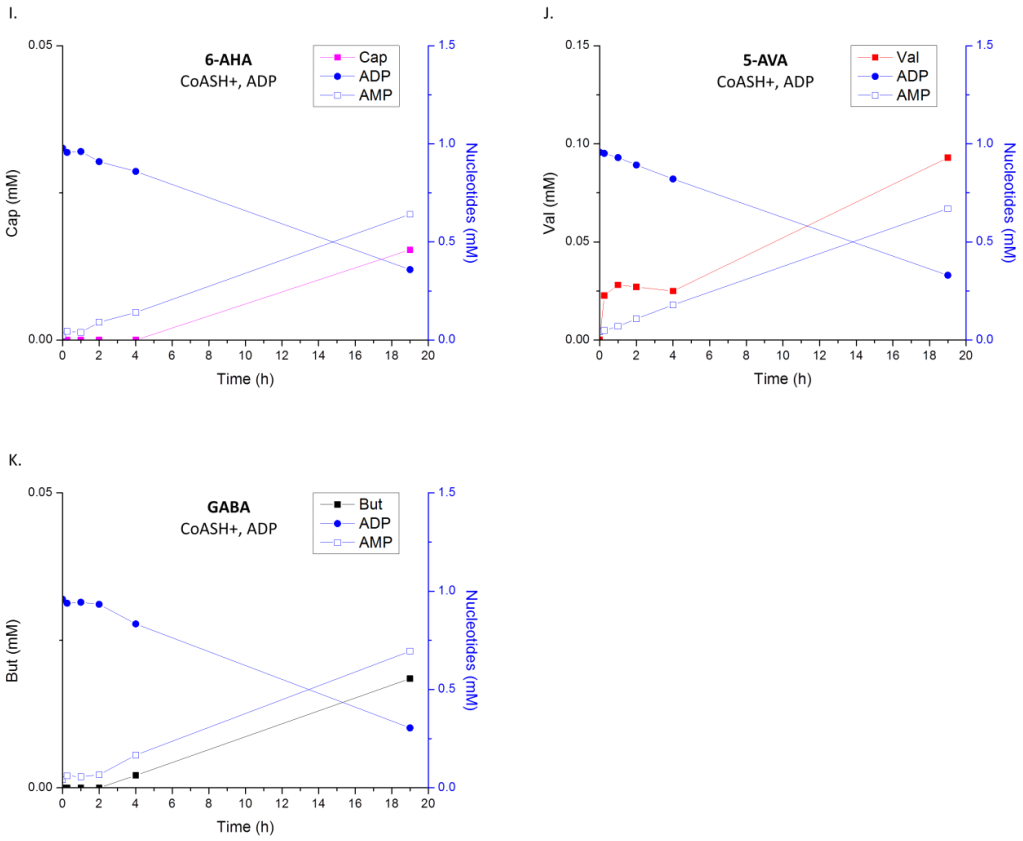
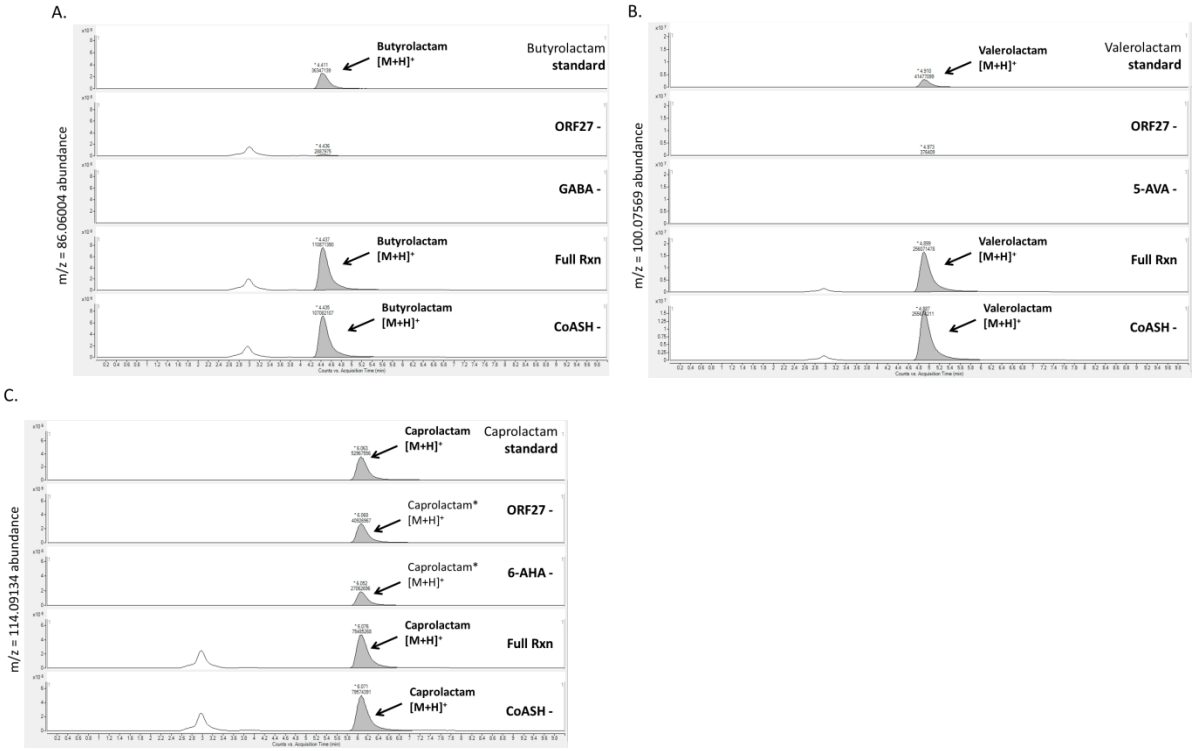


Figure S2.6. ORF27 catalyzed lactam formation. A) Butyrolactam formation. B) Valerolactam formation. C) Caprolactam formation. A reaction mixture containing 57 µM of ORF27, 5 mM ω-amino fatty acids substrates, 1 mM ATP or ADP, 0.5mM CoASH and 1 mM $Mg(Cl)_2$ in 100 mM HEPES (pH=8) was incubated at 25°C for 19h and quenched with methanol. The quenched reaction was filtered to get rid of protein aggregates before loading onto LC-MS.



* It was noticed that various types of filters tested (based on nylon, polypropylene or cellulose) all contain trace amount of caprolactam if organic solvents such as urea, MeOH were used to quench the reaction before filtering. Heat inactivation or acid, base based quenching compromised caprolactam stability. To determining caprolactam in production medium, culture could be filtered across 10kDa membrane before mixing with 1 volume of MeOH to prevent leaching of caprolactam from the filter membrane when preparing the samples for LC-MS analysis (Fig. S2.8).

Figure S2.7. pH profile of ORF27 catalyzed butyrolactam formation. A reaction mixture containing 5 μM ORF27, 1 mM GABA, 1 mM ATP, 1mM CoASH and 1 mM $Mg(Cl)_2$ in 100 mM HEPES (pH 7.5) was incubated at 25°C for 30 min. The reaction was quenched by equal volume of methanol, and the filtered solution was analyzed by analytical method described above.

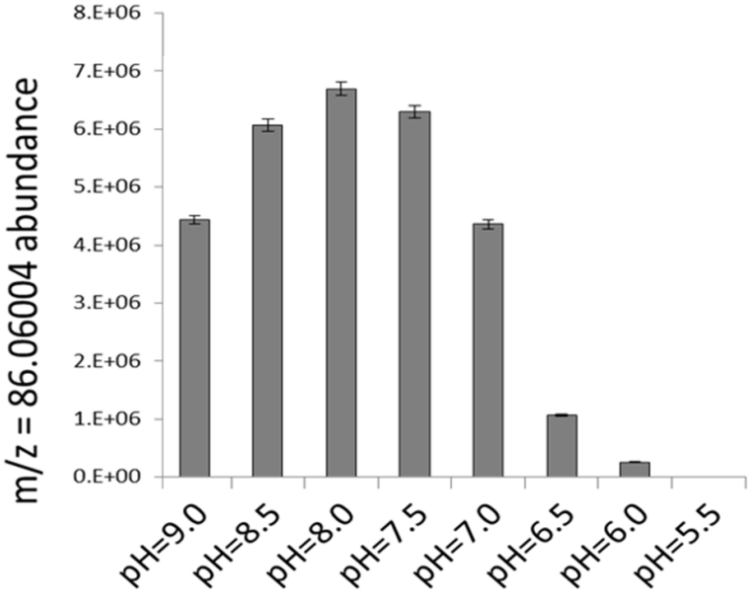
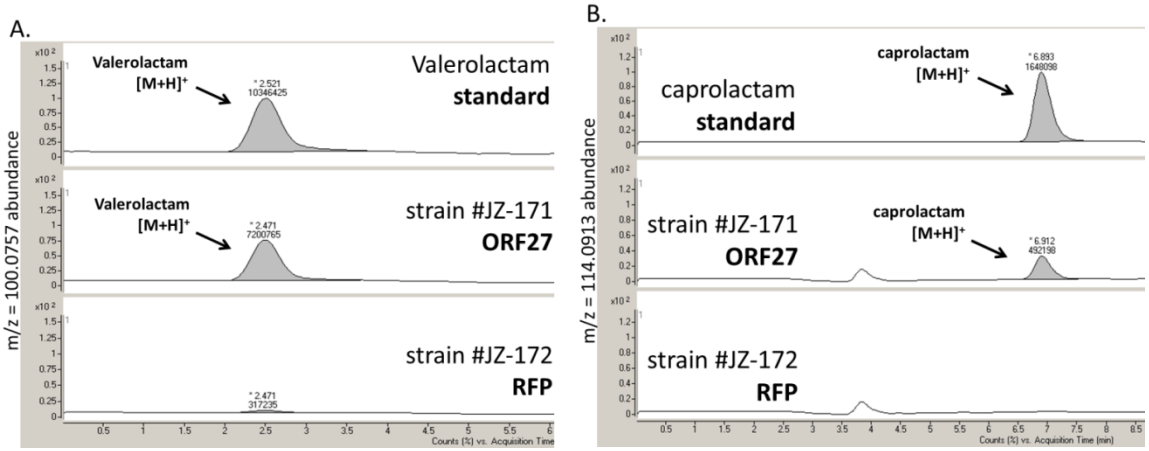


Figure S2.8. Production of industrial lactam via biosynthesis. (A) LC-MS analysis of valerolactam production from recombinant *E. coli* expressing ORF27 with 1mM 5-AVA feeding. (B) caprolactam production.



S2.9. ORF27 Biochemistry and Implications on ECO-02301 Loading Mechanism

By exploring its substrate specificity, 4-guanidinobutyric acid is confirmed to be its natural substrate. This is consistent with previous hypothesis that ECO-02301 biosynthesis uses 4-guanidinobutyric acid as a starter unit. The inability to isolate activated 4-aminobutyryl esters in the enzymatic reaction showed the transient nature of the intermediate. This suggests that the proposed amidinohydrolase (ORF33) in the gene cluster may hydrolyze the ureido group later during ECO-02301 biosynthesis rather than before loading onto the first ACP domain.

Chapter 3 Develoapment of a Transcription Factor Based Lactam Biosensor [§]

3.1. Introduction

3.1.1. Lactam Biosynthesis

Due to environmental concerns of traditional chemical synthesis routes, there is a great interest in producing materials from renewable biomass through biorefineries(49, 50, 62-68). Lactams are important compounds used in the manufacture of polyamides(69). With an annual production of four million metric tons, ε-Caprolactam (caprolactam) feeds into the production of nylon 6 found in fabrics, coatings, plastics, lubricants, etc.(48). δ-Valerolactam (valerolactam) is also a monomer for nylon 5 and nylon 6,5 synthesis, addition of which tunes the properties of the resulting polymers(49-51). And butyrolactam was identified by the US Department of Energy as an important C4 "Top Value-Added Chemical from Biomass" (1). It is currently used as the precursor for the production of N-vinylpyrrolidone, etc (2), and is also proposed to be monomer for nylon-4, a fiber material with the highest hydrophilicity in the nylon family of materials (3). With the recent discovery of lactam synthases, renewable production of such as ε-caprolactam, δ-valerolactam and butyrolactam in *E. coli* is no possible(57) (caprolactam and valerolactam biosynthesis were unpublished data). To produce lactams via fermentation at industrial scale, pathway yields and productivities need to be significantly improved.

3.1.2. High Throughput Screening

Directed evolution has been applied successfully to evolve enzymes and engineer strains(32, 70-77). A successful directed evolution procedure involves two steps: first, generating diverse genotypic libraries; and second, identifying the desired phenotype from the library(78). Our ability to generate genotypic diversity is enabled by a set of technologies, including introducing either random or targeted mutations during DNA synthesis, incorporating combinatorial promoter and ribosomal binding site (RBS) during multi-gene pathway assembly(79, 80), genome engineering techniques, such as multiplex genome engineering (MAGE)(77) and CRISPR(81-94), global transcription machinery engineering (gTEM)(95), multiscale analysis of library enrichment (SCALEs)(96), and random mutagenesis by chemical mutagens or uv radiation. However, our ability to screen through the resulting library lags behind(72). Although gas and liquid chromatography techniques allow lactam product confirmation and quantification, these analytical methods have limited throughput (10^2-10^3 variants per machine per day), which constrains the number of strains and conditions that can be tested(72). Until now, no lactam-responsive biosensor has been developed. Therefore, to construct a lactam-responsive biosensor we first sought to identify a naturally occurring transcription factor-promoter pair.

[§] Reproduced with permission from Jingwei Zhang, Mehmet Burdu, Bryton Dias, Jay. D. Keasling. "Development of a Transcription Factor Based Lactam Biosensor" (submitted)

3.1.3. Molecular Shape

Molecular shape is a very important concept in computer-aided drug design, and high throughput virtual screening based on molecular shape similarity has been widely used in drug discovery(97). However, the molecular shape similarity algorithms for drug discovery have not been widely used for metabolic engineering, and the similarity assessment was often too stringent for molecules of interest in metabolic engineering. For instance, even at 60% Tanimoto similarity cutoff computed with ROCS (Rapid Overlay of Chemical Structures)(98), the current gold standard for shape comparison, the output molecule analogs of valerolactam all contain amide bonds, which for transcription factor binding may not be necessary, since molecule analogs with ester or even α-methylene carbonyl groups can potentially be recognized. Another issue in utilizing molecular shape-based drug screening algorithms is that the compound library in medicinal chemistry has little relevance to microbial catabolism because most drug candidates have not been produced in large quantities and become environmental contaminants such that catabolic pathways would have evolved or been explored. Therefore, both the chemical library and the similarity search algorithms need to be adapted in order to apply molecular shape similarity search to metabolic engineering.

We employed a targeted strategy to identify appropriate small molecule sensor gene candidates, named "Analog Generation Towards Catabolizable Chemicals" ($AGTC^2$). The approach is based on two principles: 1) catabolic gene clusters offer a rich source of transcription factor candidates; 2) metabolic engineering targets with molecular shapes similar to catabolizable chemicals are likely to be recognized by the same transcription factor.

3.2. Material and Methods

3.2.1. Chemical Space of Bacteria Biosensor Inducers

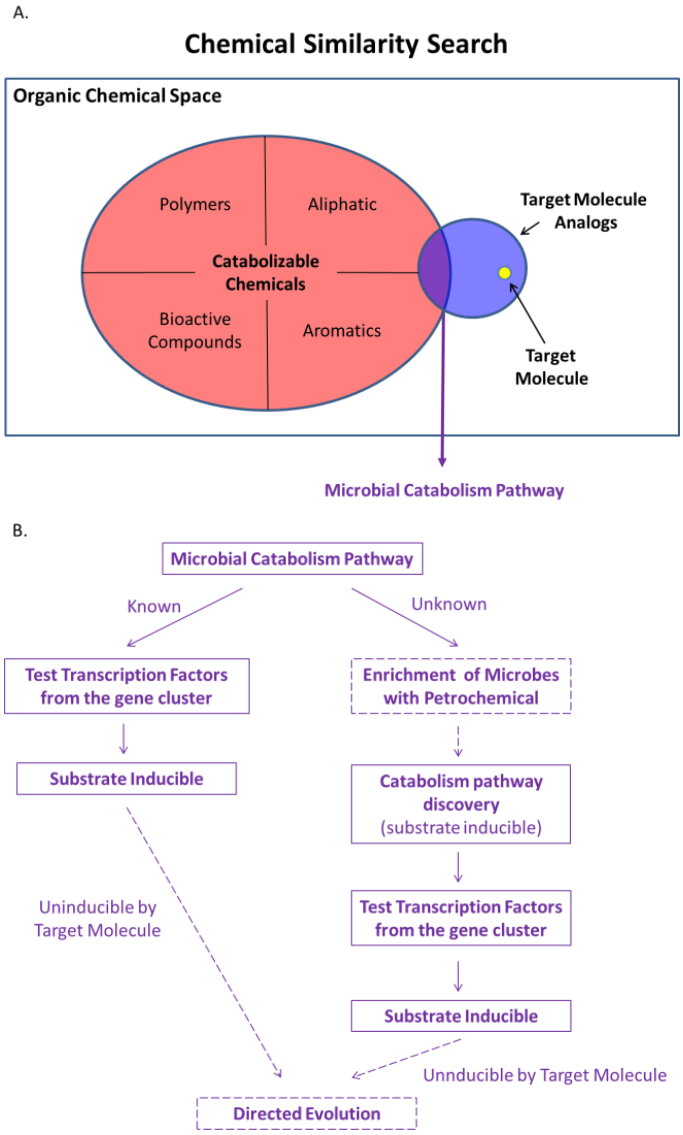
Bacteria usually use small-molecule-inducible transcription factors for substrate utilization(99), communication(100) and toxicity resistance(101). Molecules for bacterial communication cover a very limited chemical space (Table S3.4). Because toxicity resistance mechanisms involve a single pump protein(102) or a global network(95), it is difficult to identify genes whose protein products will be capable of recognizing a desired chemical. In contrast, bacteria catabolism covers a very large chemical space, and the pathways involve several genes usually present in operons on plasmids or the bacterial genome under the control of inducible promoters(99, 103). Because these pathways are only induced in the presence of the desired chemical, they are excellent candidates for transcription factors that recognize the target molecule.

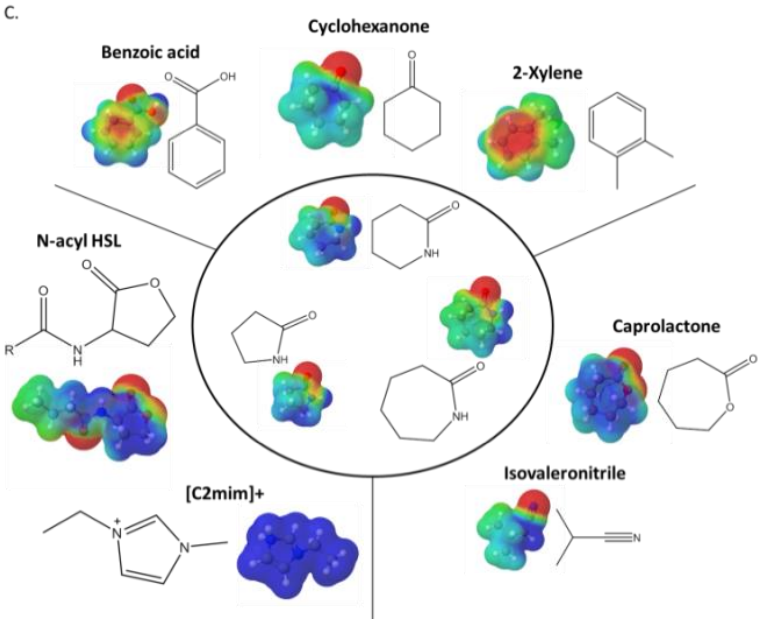
3.2.2. Analog Generation towards Catabolizable Chemicals

The search begins with the metabolic engineering target molecule. A list of molecules with various degrees of chemical similarity to the target is generated (Fig. 3.1A). Some rules of thumb are 1) maintain general molecular shape, either open chain or ring structure; 2) maintain charge distribution and polarity within a molecule, for example, C-N and C-O bonds are interchangeable, as well as nitrile groups with carbonyl groups; 3) maintain hydrophilicity and hydrophobicity distribution of molecules; and 4) in terms of similarity ranking, the further the modification is

away from the core functional group, the higher such modification ranks. For instance, γ-methyl-caprolactam is more similar to caprolactam than α-methyl-caprolactam.
The similarity search generates analogous chemicals known to be catabolized by microbes. The catabolism reactions are well curated in several databases, including the "BRENDA substructure search" (http://www.brenda-enzymes.org/index.php) (104), and many more are continually being discovered through large-scale megagenomics sequencing efforts(43, 45). Once the catabolic pathway is identified, the gene cluster encoding the catabolic pathway usually contains regulator candidates inducible by either the chemical itself or its downstream catabolic intermediates (Fig. 3.1B). When choosing from multiple candidates, higher priority is given to organisms that are closer in phylogeny to the host into which this transcription factor is planned to be engineered.

Figure 3.1. Analog Generation towards Catabolizable Chemicals (AGTC[2]). A) Chemical similarity search towards chemicals catabolizable by microbes; B) Search for transcription factor candidates in catabolism pathway; C) Lactam analogs in the catabolizable chemical space (excluding $[C2mim]^+Cl^-$).





3.2.3. Biosensing protocol

E. coli JZ-359 and JZ-439 were grown overnight in LB medium at 30°C. The overnight culture was inoculated 1:100 into fresh LB medium containing antibiotics (25 µg/mL chloramphenicol, 20 µg/mL kanamycin and 100 µg/mL ampicillin) to create a culture master stock. Stock chemicals were prepared at various concentrations 50 times the final concentration. A master culture stock (147 µL) plus 3 µL of chemical stock were transferred into polystyrene 96-well Costar Assay Plate (black with clear flat bottom, Corning Inc., New York, NY). Oxygen permeable Adhesive Plate Seals (Catalog# AB-0580, Thermo Scientific, Waltham, MA) were used to seal the plate to prevent evaporation. Due to the temperature difference between the culture after plate preparation and the 30°C incubator, condensation on the inner surface of the seal would interfere with optical signal. Therefore, after liquid was transferred, the 96-well microplate and the seals were both kept at 30°C room for 5 min before sealing the plate. The plate was monitored every 15 min for $O.D._{600nm}$ absorbance and mCherry fluorescence (λ_{ex}= 575nm, λ_{em}=620nm) in Infinite F200 multimode reader (TECAN, San Jose, CA) for 10h at 30°C.

3.2.4. Biosensor Performance Calculations

The following Hill equation was used to fit the transfer function to derive the biosensor performance features:

$$\text{mCherry signal} = \text{mCherry}_0 + (\text{mCherry}_{chemical} - \text{mCherry}_0)\cdot\frac{[Lactam]^n}{{K_m}^n + [Lactam]^n}$$

where mCherry_0 is the fluorescence in the absence of chemical, $\text{mCherry}_{\text{chemical}}$ is the fluorescence in the presence of the chemical, [Lactam] is the ligand concentration, K_m is the ligand concentration that results in half-maximal signal, and n is a measure of the biosensor sensitivity (Hill coefficient)

Data were fit with the Hill Equation to mathematically describe biosensor transfer functions and derive biosensor performance features (**Table 3.2**).

Table 3.2. Biosensor performance features[a]

Compound	$mCherry_{max}$ (AU)	Transfer Function	Dynamic Range	K_m (mM)	Linear range (mM)
Butyrolactam	11781	$mCherry = 6461 + 15932 \cdot \frac{[But]^{1.17}}{216^{1.17}+[But]^{1.17}}$	1.8	216	1 – 120
Valerolactam	23292	$mCherry = 5533 + 30905 \cdot \frac{[Val]^{0.85}}{85^{0.85}+[Val]^{0.85}}$	3.5	85	1 – 120
Caprolactam	21497	$mCherry = 6636 + 23082 \cdot \frac{[Cap]^{1.28}}{33^{1.28}+[Cap]^{1.28}}$	3.1	33	1 – 50

[a] Dose response curves were fitted to the Hill equation to derive the biosensor transfer functions. Dynamic range is the difference in output signal between high and low states. K_m, the inducer concentration resulting in half-maximal induction. Linear range is the series of chemical concentrations for which a change in signal can be detected.

3.3. Results

3.3.1. Lactam Analog Search

For lactam, we conducted analog generation towards catabolizable chemicals for butyrolactam, valerolactam and caprolactam (Fig. 3.1C). By comparing their molecular shape, cyclohexanone, caprolactone and isovaleronitrile stood out. Among them, caprolactone catabolism is likely inducible by its downstream products upon metabolism. Isovaleronitrile serves as the inducer for NitR/pNitA transcription factor/promoter pair from the *R. rhodochrous* J1(105-108) isovaleronitrile catabolism pathway (Fig. S3.1A). Cyclohexanone serves as the inducer for ChnR/Pb transcription factor/promoter pair originating from the *Acinetobacter sp.* SE19 or NCIMB 9871(109-113) cyclohexanol catabolic pathway (Fig. S3.1B). 1-Ethyl-3-methylimidazolium chloride ($[C2mim]^+Cl^-$) served as a negative example. Although it also has five-membered ring structure, the charge distribution is significantly different from butyrolactam. A gel shift assay using purified EilR(114), a transcription factor inducible by $[C2mim]^+Cl^-$, did not show lactam inducibility (data not shown).

3.3.2. Isovaleronitrile-inducible NitR/pNitA promoter pair

Although the NitR/pNitA system was previously shown to be inducible by caprolactam when heterologously expressed in Gram positive *Streptomyces(105, 115)* and *Mycobacteria* species(107), attempts to transfer this system into Gram negative *E. coli* failed due to inability to heterologously express NitR solubly in *E. coli* (Fig. S3.2).

3.3.3. Cyclohexanone inducible ChnR/Pb promoter pair

3.3.3.1. Design and Confirmation of a Bi-plasmid Lactam Biosensor System.

The bacteria biosensor strains are listed in Tables 3.1 and S3.1-S3.3. The ChnR-Pb transcription factor/promoter pair is involved in the catabolism of cyclohexanone. It was first tested as a bi-plasmid system in *E. coli* JZ-359. One plasmid, pBbA8a-ChnR, has *chnR* under control of the arabinose-inducible *araBAD* promoter (P_{BAD}), and the second plasmid, JZ-BCS-2, harbors the *mCherry* gene encoding for mCherry fluorescent protein (mCherry) under control of the *Acinetobacter sp. chnB* promoter (Pb). Biosensor constructs were co-transformed into *E. coli* DP10 containing a Δ*araE* deletion to enable titratable induction of ChnR, which is important for tuning both protein toxicity and transcription factor/small molecule complex formation.

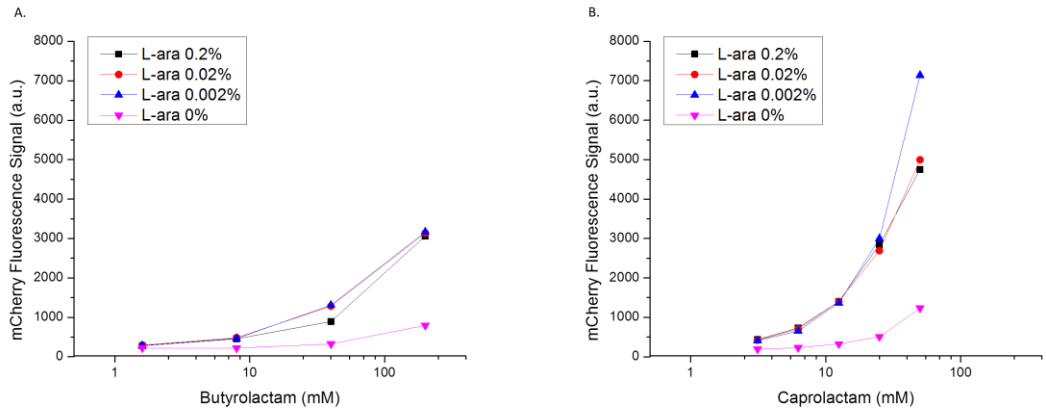
Table 3.1. *E. coli* strains.

Name	Relevant genotype	Reference
Strains		
DH10B	F⁻ mcrA crmrr-hsdRMS-mcrBC) r-hsdRMS-mcrBC) and oligonucleotide139 Δ(ara, leu)7697 galU galK alrpsL nupG	Life Technologies (Carlsbad, CA)
DP10	DaraFGH DP-araE & PCP18-araE	

Strains	plasmids	Host
JZ-359	JZ-BCS-2 (Kan) + pBbA8a-ChnR	DP10
JZ-439	pBbSLactamC-mCherry	DH10B

Biosensor response was observed with either exogenously added butyrolactam or caprolactam when cultured at 30°C (Fig. 3.2A-3.2B). Additionally, the biosensor response was shown to be slightly negatively correlated with arabinose induction, suggesting some levels of ChnR toxicity or competition with mCherry production. In the absence of L-arabinose, lactam dose dependent induction was observed, suggesting that leaky expression of ChnR is sufficient for lactam inducible response.

Figure 3.2. Lactam-inducible ChnR/Pb pair. A) Inducible and dose dependent mCherry expression by butyrolactam; B) Inducible and dose dependent mCherry expression by caprolactam.



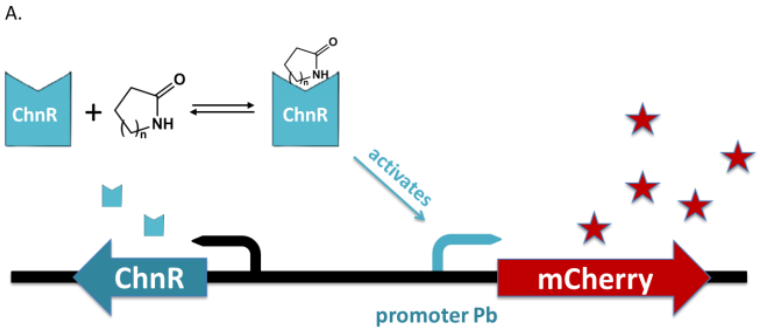
3.3.3.2. Single Plasmid Lactam Biosensor System for Metabolic Engineering

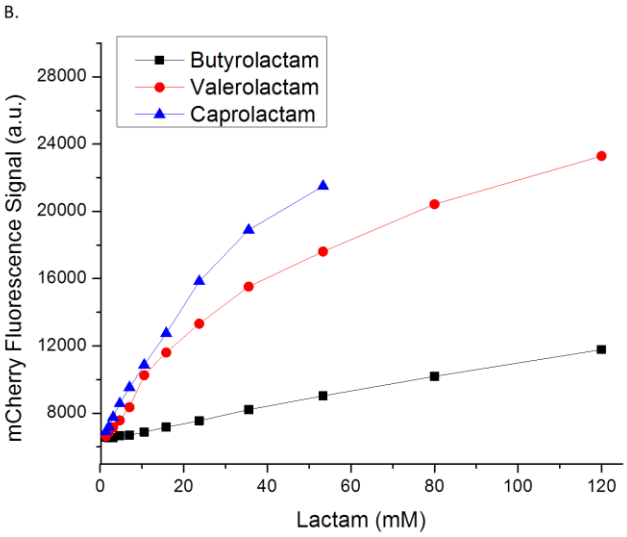
The initial biosensor system had limited biotechnology usefulness, because the ChnR was controlled by external inducer, and the two-plasmid system would have significant burden for the production strains. Further improvements were made to create a single plasmid biosensor JZ-439 for high throughput metabolic engineering with more desirable features: 1) ChnR was placed under the control of a constitutive promoter; 2) a low copy number plasmid was used to minimize the metabolic burden of plasmid replication; 3) The biosensor's response against pathway intermediates was determined (Fig. S3.4); and 4) the biosensor's dynamic range was matched with the anticipated future lactam titer improvement range.

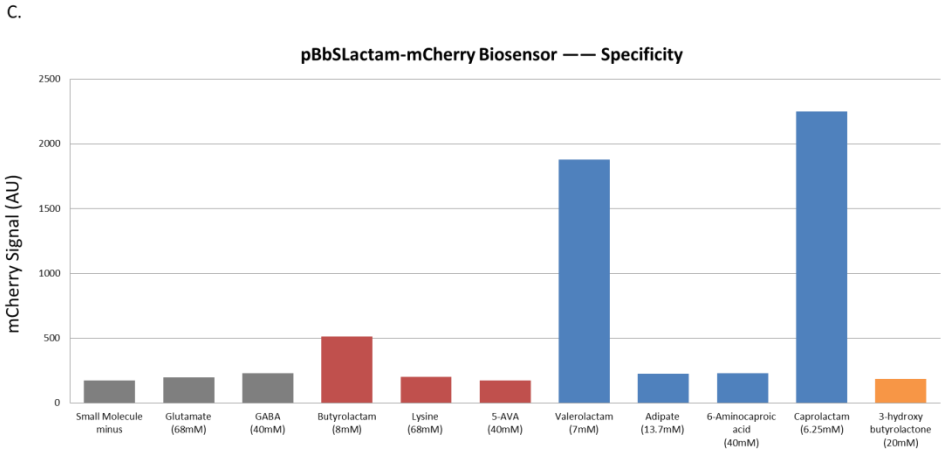
An optimized biosensor design (pBbSLactamC-mCherry) was constructed by incorporating a weak RBS 5' of ChnR, and the SC101 origin of replication was chosen to minimize biosensor footprint (Fig.3.3A). *E. coli* DH10B transformed with pBbSLactamC-mCherry was used for characterization of biosensor response to exogenously added butyrolactam, valerolactam and caprolactam (Fig.3.3B). Butyrolactam and valerolactam did not affect growth (Fig. S3.3A-S3.3B). Because caprolactam is toxic beyond 50 mM, which results in a decrease in normalized mCherry fluorescence (Fig. S3.3C), caprolactam biosensor data were collected only up to the toxicity limit.

Transfer functions for the valerolactam and caprolactam had the largest dynamic range; in the case of all lactams, the linear range of detection range was between one to two orders-of-magnitude (Table 3.2.). In addition to a broad linear range of detection and large dynamic range for lactams, the biosensor was also highly selective against lactam biosynthetic intermediates. Even at significantly high concentration, lactam biosynthetic pathway intermediates did not trigger mCherry expression above baseline. S-3-hydroxy butyrolactone, another "Top Value-Added Chemical from Biomass", was not able to induce mCherry expression(Fig. 3.3C).

Figure 3.3. Single Plasmid Biosensor System. A) The lactam-inducible mCherry detection system; B) Inducible and dose dependent mCherry expression by butyrolactam, valerolactam and caprolactam; C) ChnR was inducible by 8mM butyrolactam, 7mM valerolactam, 6.25mM caprolactam, yet not inducible by lactam biosynthetic pathway intermediates, such as 68mM glutamate, 40mM GABA, 68mM Lysine, 40mM 5-AVA, 13.7mM adipate, 40mM 6-aminocaproic acid or butyrolactam anaolog 3-hydroxy-butyrolactone, demonstrating great specificity.







Currently, engineered *E. coli* can produce butyrolactam at >10 mM titer, and valerolactam at approximately 2 mM (unpublished data). The the linear range (1-120mM) of the constructed biosensor strain JZ-439 falls within a reasonable range to anticipate future improvements in titer. In case when titer reaching beyond the linear range of detection, appropriate dilution of medium could bring the concentration of lactams back to linear range.

From these results, we concluded the biosensor is well-suited for high-throughput lactam screening.

3.4. Discussion

Developing higher throughput screening assays for microbial lactam production is challenging due to several features of this class of molecules: 1) they lack differential spectroscopic features that would allow them to be monitored using high throughput detection methods; 2) the amide bond is relatively inert, and lactams have few activated functional groups to enable specific chemical derivatization or enzymatic transformation; 3) it is completely mixable with both water and organic solvents, making it hard to extract or enrich; 4) DNA/RNA sensors for lactams are

unlikely to be developed due to their small molecular size, relative high hydrophobicity and lack of distinct hydrogen bond donors and acceptors; and 5) Chemical or protein sensors based on host-guest supramolecular interaction are the only likely sources of candidates. Protein sensors have the advantage of low price during implementation since it's genetically coded. The protein residues provide a much larger interaction network, potentially leading to greater specificity.

Nature does not generally develop transcriptional control for butyrolactam, valerolactam and caprolactam though. These are artificial petrochemicals, and do not participate in metabolism. Their low toxicity for cell growth does not justify the need for evolving an inducible transporter system that pumps lactams outside the cell. Even in the cases where lactams is utilized by microbes as sole carbon and nitrogen source, inducible catabolism of lactams are usually achieved by a constitutive expression of a hydrolase, followed by transcriptional control by the downstream open chain ω-amino fatty acids upon hydrolysis. Therefore, we explored chemical analogs with similar molecular shape as the lactam, and seek their catabolism pathways.

The "Analog Generation towards Catabolizable Chemicals" ($AGTC^2$) strategy is generally applicable because: First, it is believed that microbes can utilize virtually any organic compounds for growth given long term exposure in the environment. There is a tremendous set of literature on microbial substrate utilization with well-characterized catabolic pathways and their regulatory mechanisms. Studies on microbial growth in pollutant-contaminated environment greatly expanded the chemical space of catabolizable chemicals. Large-scale meganomics studies will make more microbial genome information available. In case where analogs are not documented in the literature, it is still convenient to look for candidates because pollution or wastewater treatment already created enriched bacterial candidates. Second, transcription factor discovery is naturally amenable to high throughput screening. Even if the transcription factor candidates were discovered to be uninducible by the metabolic engineering target molecule, directed evolution of the transcription factor candidate to recognize molecules that are similar to its native ligand has a high probability of success.

A number of chemical informatics algorithms have been developed for drug discovery. By incorporating features relevant to microbial sensing, curating a chemical space specific for catabolizable chemicals, as well as applying filters in the context of metabolic engineering, these algorithms can be further tuned to automate and expedite transcription factor discovery for small molecule inducibility. Apart from small molecule inducibility, additional filters such as phylogenetic distance from host organism, specificity against metabolic pathway intermediates and relevant dynamic range are also important factors to achieve protein functional heterologous expression and to be utilized for high throughput screening.

The ChnR/Pb transcription promoter pair had Hill coefficient of close to 1. The Hill coefficient may suggest the ChnR protein functioned as a monomer. Since multimeric proteins normally recognize DNA sequences with C2 or semi-C2 symmetry, during transcription factor screening, additional filter such as prioritizing promoter sequences without palindromic sequences may further facilitate the discovery of biosensor with great linear range.

3.5. Conclusion

The discovered ChnR/Pb transcription factor/promoter pair is a proof of concept demonstration of transcription factor scouting using the "Analog Generation towards Catabolizable Chemicals" ($AGTC^2$) strategy. Plasmid origin of replication, promoter strength and RBS 5' of either the transcription factor or reporter protein are parameters that can fine-tune the properties of the biosensors. The ChnR/Pb biosensor system can potentially enable high throughput screening of lactam titer improvement over 1-2 orders of magnitude, depending on specific lactam targets.

3.6. Acknowledgement

Plasmid pMS119 was a kind gift from provided by Prof. Svein Valla. We thank Dr. Magnus Steigedal for discussion related to discussion on transcription factors. Plasmid pNIT-1 was a kind gift from provided by Prof. Christopher M. Sassetti. We thank Dr. Thomas L. Ruegg for providing purified EilR protein and conducting the EilR/lactam gel shift assays.

Supplemental Material

Materials and Methods

Vector construction

ChnR was amplified from pMS119 using primers JZ_s8CChnR_f / JZ_s8CChnR_r and cloned under P_{BAD} in pBbA8a-RFP at BglII/XhoI to create pBbA8a-ChnR.
JZ-BCS-2 was constructed by Gibson reaction of 2 DNA parts: 1. the vector backbone of WH32-5 was digested at SalI/KpnI restriction site; 2. Pb promoter was amplified from pMS119 using primers JZ_Pbinsert_F/ JZ_Pbinsert_R.

pBbSLactamC-mCherry was constructed by Gibson reaction of 4 DNA parts: 1. the vector backbone of pBbS5C-RFP was amplified using primers JZ_pBbS5C_backbone_f/ JZ_pBbS5C_backbone_r; 2. the constitutive expression of ChnR was achieved by amplifying the constitutive promoter in front of LacI from pBbS5C-RFP using primers JZ_lacuv5space_f/ JZ_lacuv5space_r; 3. The mCherry was amplified from WH32-5 using primers JZ_Pb_mcherry_f/JZ_Pb_mcherry_r; 4. The ChnR was amplified from pBbA8a-ChnR using primers JZ_ChnR_Gib_f/ JZ_ChnR_Gib_r

NitR was amplified from pNIT-1 using primers JZ_NitR_N6xHis_f/ JZ_NitR_N6xHis_r and cloned into pSKB3 vector at NdeI/XhoI to create pSKB3-N-6xHis-NitR. NitR was amplified from pNIT-1 using primers JZ_NitR_C6xHis_f/ JZ_NitR_C6xHis_r and cloned into pSKB3 vector at NcoI/EcoRI to create pSKB3-C-6xHis-NitR.

Strain construction

Standard molecular biology techniques were used to construct the two biosensor strains. Electrocompetent DP10 strain was prepared and double transformed with pBbA8a-ChnR+JZ-BCS-2 to create strain JZ-359. JZ-439 was prepared by transforming pBbSLactamC-mCherry into chemically competent DH10B strain. JZ-243 and JZ-244 were prepared by transforming pSKB3-N-6xHis-NitR or pSKB3-N-6xHis-NitR into electrocompetent BL21(DE3)star strain.

Solubility Check for NitR Protein Expression

For either N-terminal 6xHis NitR or C-terminal 6xHis NitR expression, the overnight NitR overexpression strain culture was inoculated (1:100 v/v) into 10mL LB medium containing 5 μg/ml kanamycin. The culture was grown at 37°C until the O.D. reached 0.6 and cooled on ice for 20 min. 1mM IPTG was added to induce NitR overexpression for 16 h at either 18°C or 13°C. 1.5mL of the culture was harvested by centrifugation (8000 × g, 2 min, 4°C), resuspended in 750μL of lysis buffer (50 mM HEPES, pH 8.0, 0.5 M NaCl), and lysed by sonication on ice. The solube and insoluble fractions were seperated by centrifugation (20,000 × g, 40 min, 4°C). Supernatant containing soluble protein was taken out and the pellet was resuspended in 750 μL lysis buffer. 14μL of the soluble supernatant sample or the insoluble pellet suspension was taken, mixed with 1μL 2mM DTT plus 5 μL 4X NuPAGE LDS Sample Buffer (ThermoFisher

Scientific, Waltham, MA). The mixture was heated at 70°C for 5min and loaded onto Ready gel for Tris-Glycine Gel (10% precast, Bio-Rad).

Supplementary Table 3.1. Table of strains

Strains		Host
JZ-243	pSKB3-N-6xHis-NitR	BL21(DE3)star
JZ-244	pSKB3-C-6xHis-NitR	BL21(DE3)star

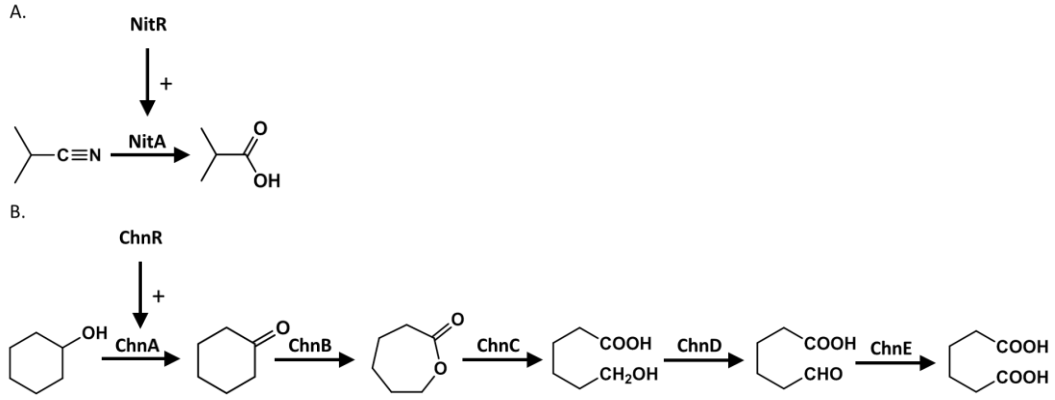
Supplementary Table 3.2. Table of plasmids

Plasmids		Reference
pBbA8a-RFP	p15A *ori*, Amp^R, *araC*, RFP	(28)
pBbS5C-RFP	SC101 *ori*, Cm^R, *lacI*, RFP	(28)
pMS119	ColE *ori*, Amp^R, *LacO*,ChnR, ChnE, ChnB	(113)
pBbA8a-ChnR	SC101 *ori*, Cm^R, *araC*,ChnR	This study
WH32-5	ColE *ori*, Kan^R, ZF repressible promoter, mCherry	This study (courtesy of Dr. Will Holtz)
JZ-BCS-2	ColE *ori*, Kan^R, Pb promoter, mCherry	This study
pBbSLactamC-mCherry	SC101 *ori*, Cm^R, *chnR*, mCherry	This study
pNIT-1	OriM *ori*, OriE *ori*, Kan^R, pNitA, GFP	(107, 109)
pSKB3-N-6xHis-NitR	ColE *ori*, Kan^R, LacI, N-6xHis-NitR	This study
pSKB3-C-6xHis-NitR	ColE *ori*, Kan^R, LacI, N-6xHis-NitR	This study

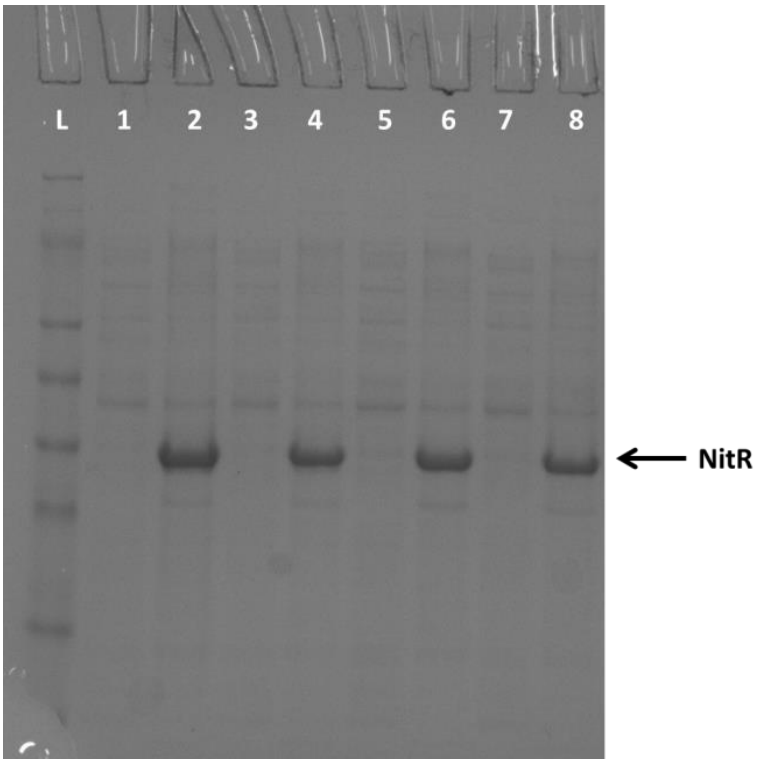
Supplementary Table 3.3. Table of primers

Oligonucleotides	5'→ 3' Sequence Integrated DNA Technologies, Inc (Coralville, IA)	Target Gene
JZ_s8CChnR_f	GCGCGagatcttttaagaaggagatatacatatgagcacagacaaagcaaatacgc	ChnR
JZ_s8CChnR_r	GCGCGctcgagtttggatcctcaaaaaacaatagaggagactgaattttcagac	ChnR
JZ_Pbinsert_F	ttcatcccaatccacacgtccaac	Pb promoter
JZ_Pbinsert_R	gatagccatgttatcttcttctcctttacgc	Pb promoter
JZ_ChnR_Gib_f	gggaaaaagtggaagcggcgatgagcacagacaaagcaaatacgc	ChnR
JZ_ChnR_Gib_r	ggaagagagtcaattcagggtggtgaatatgagcacagacaaagcaaatacgc	ChnR
JZ_lacuv5space_f	ctagtgatatcccgcggccactcgatcctctacgccggac	LacI promoter
JZ_lacuv5space_r	gcgtatttgctttgtctgtgctcatattcaccaccctgaattgactctcttcc	LacI promoter
JZ_Pb_mcherry_f	ggagatccttactcgagtttggatccttatttgtatagttcatccatgccaccgg	mCherry
JZ_Pb_mcherry_r	gtccggcgtagaggatcgagtggccgcgggatatcactag	mCherry
JZ_pBbS5C_backbone_f	gtctgaaaattcagtctcctctattgtttttgagcgcaacgcaattaatgtaagttagc	pBbS5C vector backbone
JZ_pBbS5C_backbone_r	ccggtggcatggatgaactatacaaataaggatccaaactcgagtaaggatctcc	pBbS5C vector backbone
JZ_NitR_N6xHis_f	gcgcg CAT atgaacactttcttctcctcagac	NitR
JZ_NitR_N6xHis_r	gcgcg GGATCCctacgaaacctccgtcggtg	NitR
JZ_NitR_C6xHis_f	GCGCG ccatgggc atgaacactttcttctcctcagac	NitR
JZ_NitR_C6xHis_r	gcgcgCGAATTCGGcgaaacctccgtcggtg	NitR

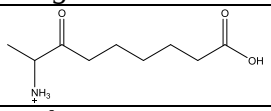
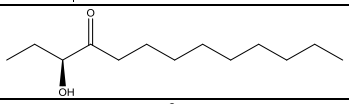
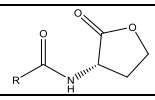
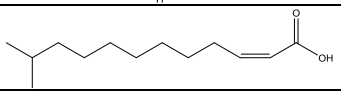
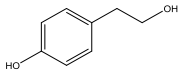
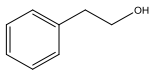
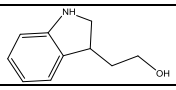
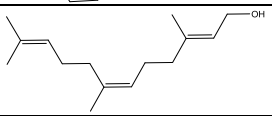
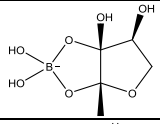
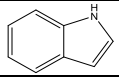
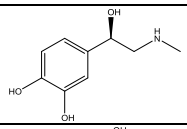
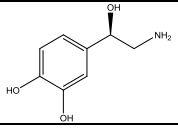
Supplementary Figure 3.1. Catabolic pathways and inducible transcription factors. A) Isovaleronitrile catabolism pathway in *R. rhodochrous* J1, NitR activated by isovaleronitrile; B) Cyclohexanol catabolism pathway in *Acinetobacter sp.*, ChnR activated by cyclohexanone.



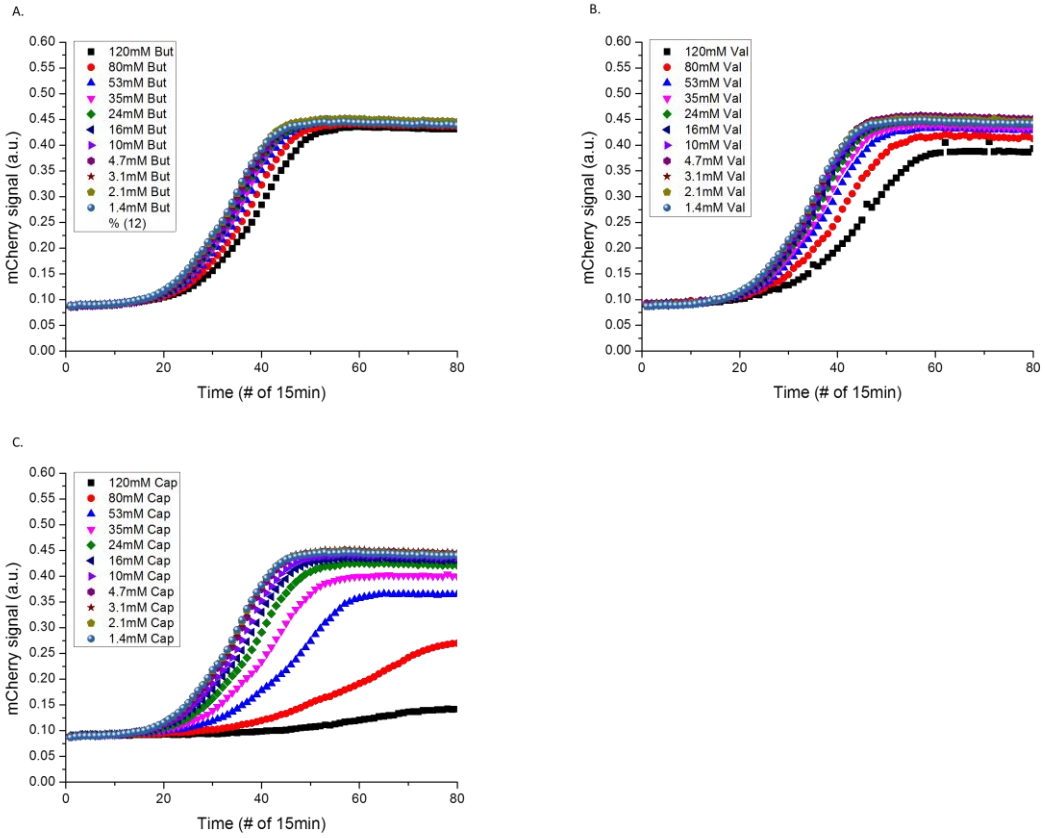
Supplementary Figure 3.2. Solubility test of NitR in E. coli. SDS-PAGE of NitR expressed at various temperature in *E. coli*. For molecular weight determination, PageRuler™ Prestained Protein Ladder (10 to 180 kDa, ThermoFisher Scientific, Waltham, MA) was used as protein ladder. L: Protein ladder; Lane 1: N-6xHis-NitR induction at 18°C (soluble); Lane 2: N-6xHis-NitR induction at 18°C (insoluble); Lane 3: N-6xHis-NitR induction at 13°C (soluble); Lane 4: N-6xHis-NitR induction at 13°C (insoluble); Lane 5: C-6xHis-NitR induction at 18°C (soluble); Lane 6: C-6xHis-NitR induction at 18°C (insoluble); Lane 7: C-6xHis-NitR induction at 13°C (soluble); Lane 8: C-6xHis-NitR induction at 13°C (insoluble).



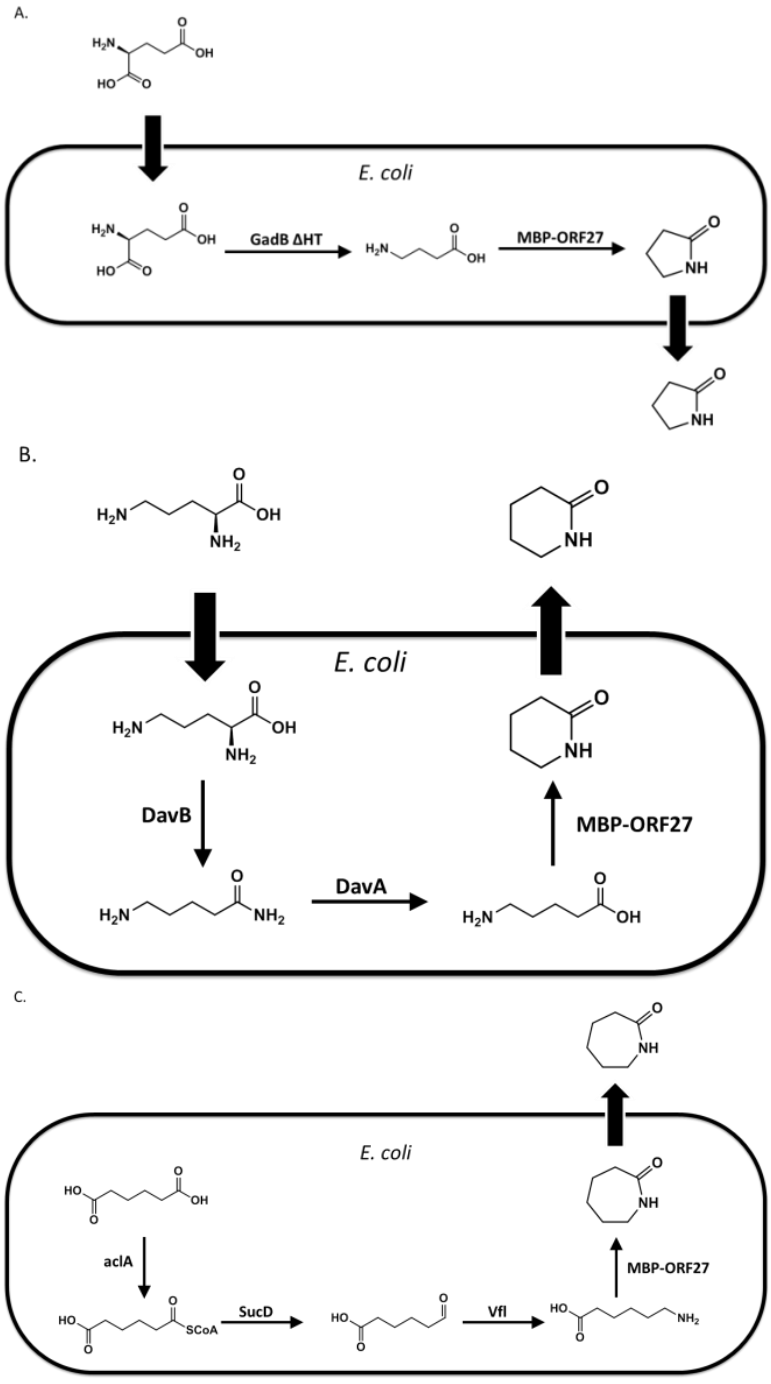
Supplementary Table 3.4. Small molecule inducible transcription factors for *cell-cell communication.*

Organism	*Ligand*	*Ligand Structure*	*Sensor*	*Ref.*
L. pneumophila	8-amino-7-oxononanoate		LqsS	**(116)**
V. cholerae	*(S)-3-hydroxytridecan-4-one*		CqsS	**(117)**
V. fischeri	*N-acyl HSL*		LuxR	**(118)**
X. campestris	*cis*-11-methyl-dodecenoic acid		RpfC	*(119)*
C. albicans S. cerevisiae	*tyrosol*		*unknown*	*(120)*
	phenylethanol			
	tryptophol			
P. aeruginosa	Farnesol (inhibitory)		PqsR (MvfR)	*(121)* *(122)*
V. cholerae	AI-2		LuxPQ	*(117)*
E. coli	Indole		SdiA	(123)
E. coli	AI-3	Uncharacterized	QseBC	(124)
	Epinephrine			
	norepinephrine			

Supplementary Figure 3.3. Lactam induced toxicity. A) *E. coli* strain JZ-439 growth curve under various butyrolactam concentration, the O.D.$_{600nm}$ was monitored every 15min; B) growth under various valerolactam concentration; C) growth under various caprolactam concentration.



Supplementary Figure 3.4. Intermediates for engineered biosynthetic pathways of various lactams. A) Butyrolactam biosynthesis; B) Valerolactam biosynthesis; C) Caprolactam biosynthesis.



Chapter 4. Future Directions

4.1. 2-pyrrolidone biosynthesis in *Corynebacterium glutamicum*

Glutamic acid is considered as one of the 12 potential platform chemicals that could be derived from sugar. It is possible to convert glutamic acid through bio-catalysis process to high value-added chemicals for functional food, medicine and materials, such as γ-aminobutyric acid, α-ketoglutaric acid, and 2-pyrrolidone etc. Among these derivatives, 2-pyrrolidone is an important solvent and organic chemical intermediate that is widely used in analysis equipment washing, biorefining, and production of synthetic resins, chemical, pharmaceutical, pesticides, dyes and light industry. It is the important feedstock for synthesizing N-vinyl pyrrolidone, the monomer of polyvinyl pyrrolidone, which is a non-ionic polymer compound used widely in medicine, food, household chemicals, textile etc. Besides, ring-opening polymerization of 2-pyrrolidone forms nylon-4, a translucent or milky white thermoplastic resin. Compared with other nylon materials, nylon-4 has better thermal stability and the highest hydrophilicity. The synthetic leather made from Nylon-4 is flexible, porous, and anti-static. Nylon-4 has similar hydrophilicity as cotton and silk, and is considered as an ideal substitution of cotton fibre.

Nowadays, 2- pyrrolidone is synthesized by chemical process using fossil-based feedstock. The price of 2-pyrrolidone ranges from $3000 to $7000 per ton, which is much higher than $1300 per ton of glutamic acid. Thus, it is economically attractive to develop biosynthesis process of producing 2-pyrrolidone from glutamic acid or even directly from glucose by engineering an industrial glutamic acid producer.

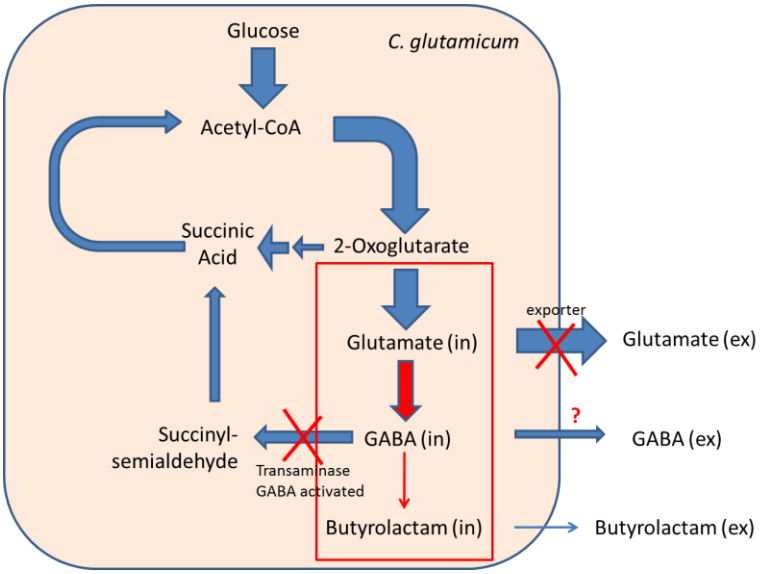
The scope of the project is to developing new strains that could be used for converting glucose to 2-pyrrolidone by fermentation:

(1) Screening and cloning new genes of glutamic acid decarboxylase and pyrrolidone synthase. Obtaining new glutamic acid decarboxylase and pyrrolidone synthase genes by using bioinformatic tools for gene mining. Optimizing DNA sequences. Overexpressing the genes in *Escherichia coli* or *Corynebacterium glutamicum*. Purifying and characterizing the enzymes.

(2) Constructing and optimizing 2-pyrrolidone synthesis pathway in *C. glutamicum* using glucose as substrate. Optimizing codon usage of the high activity and stability glutamic acid decarboxylase and pyrrolidone synthase candidates for *C. glutamicum*. Installing these enzymes into industrial glutamic acid strain. Determining bottle neck steps and metabolic flux by monitoring proteomics and metabolomics during fermentation. Engineering the bottle neck steps to increase the reaction rate and reducing intermediates by increasing copy number, changing promoter strength, enzyme engineering etc.

Several *C. glutamicum* gene candidates are likely involved in consumption of glutamate or GABA. Knock out of Ncgl0462 (GABA transaminase), Ncgl1221 (glutamate exporter) or double knockout of both can potentially direct flux towards 2-pyrrolidone (Fig. 4.1).

Figure 4.1. Host engineering candidates for 2-pyrrolidone production in *C. glutamicum*.



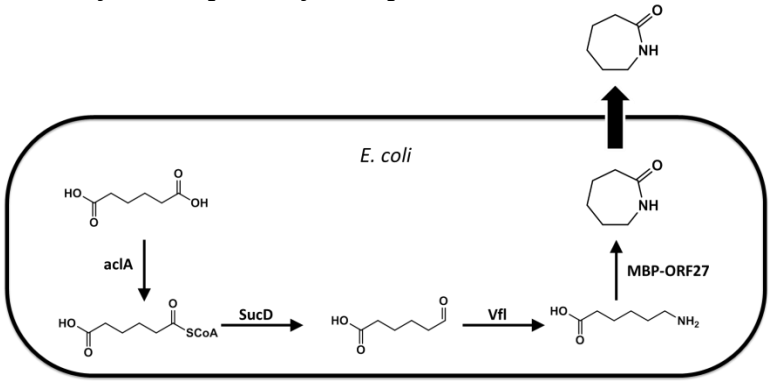
(3) Optimizing fermentation conditions for 2-pyrrolidone overproduction in *C. glutamicum*. Fermentation conditions will be further tuned to optimize the production of 2-pyrrolidone for the engineered *C. glutamicum*: including growth and fermentation condition, temperature, pH, etc.

(4) Optimizing fermentation conditions for 2-pyrrolidone overproduction by whole cell catalysis. Conditions will be further tuned to optimize the production of 2-pyrrolidone for the engineered whole cell (*E. coli* or *C. glutamicum*, etc.): including cell growth condition such as media, temperature or pH, protein overexpression condition, as well as cell stability, etc.

4.2. Construction of Caprolactam Producer Strains

Caprolactam (caprolactam) is used in the production of nylon 6 found in fabrics, coatings, plastics of many compositions, lubricants, etc. The worldwide use of nylons requires the production of approximately four million metric tons of caprolactam annually (48). With the discovery of ORF27 as a general lactam synthase, it would be possible to integrate ORF27 with 6-aminohexanoic acid pathway to enable complete biosynthesis of caprolactam (Fig. 4.2).

Figure 4.2. Proposed biosynthetic pathway of caprolactam



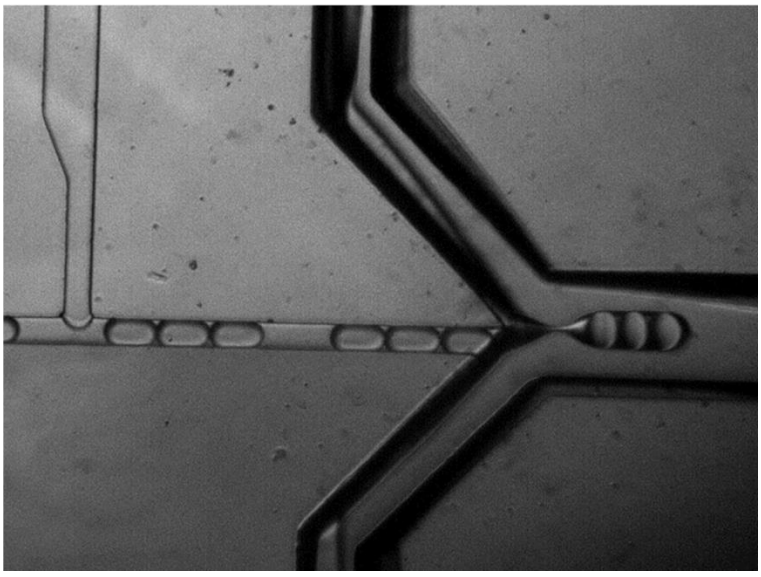
4.3. High Throughput Screening of Lactam Production Strains

The integration of biosensor with the lactam producers would enable high throughput screening of lactam production strains with higher titer.

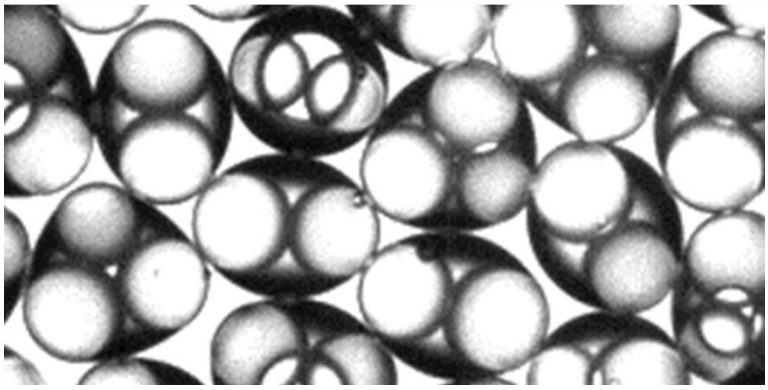
Droplet microfluidics offers an interesting high throughput system for screening millions of strain variants. A double emulsion device was constructed, courtesy of Tuan Tran at Adam Abate lab, UCSF (Fig. 4.3). The device can achieve co-localization of production strains and biosensing strain in proximity at a frequency of around 10,000/sec, allowing extremely high throughput screening of strains. Lactam was shown to be diffusible across the droplet boundary.

Figure 4.3. Droplet Microfluidic System for High Throughput Screening of Lactam Producers. A) Droplet system to form triplet emulsion, plugging of one channel could lead to double emulsion formation devices; B) Double emulsion formed from microfluidic device; C) Schematic representation of encapsulating single production strain variant in droplet and sensing strain in the nearby double emulsion droplets. The single cell encapsulation was achieved by cell dilution (1/10 droplets have cells)

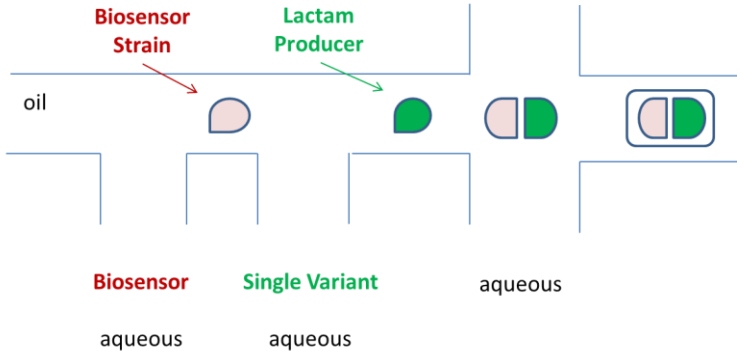
A.



B.



C.


Biosensor Strain
Lactam Producer
oil
Biosensor
Single Variant
aqueous
aqueous
aqueous

Reference

1. **Werpy P.** 2004. Top Value Added chemicals from Biomass. **Volume I - results of screening for potential candidates from sugars and synthesis gas**.
2. **BASF.** 2015. 2-Pyrrolidone. http://www.basf.ca/group/corporate/ca/en/brand/2_PYRROLIDONE. Accessed 03.05.
3. **Park.** 2013. Synthesis of nylon 4 from gamma-aminobutyrate (GABA) produced by recombinant Escherichia coli. Bioprocess Biosyst Eng **36:**885-892.
4. **Harreus RB, J.-O. Eichler, R. Feuerhake, C. Jakel, U. Mahn, R. Pinkos, R. Vogelsang.** 2011. 2-pyrrolidone, p *In* (ed), Ullmann's Encyclopedia of Industrial Chemistry, ed vol Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim,
5. **Shi.** 2013. Enhancement of gamma-aminobutyric acid production in recombinant Corynebacterium glutamicum by co-expressing two glutamate decarboxylase genes from Lactobacillus brevis. J Ind Microbiol Biotechnol **40:**1285-1296.
6. **Takahashi.** 2012. Robust production of gamma-amino butyric acid using recombinant Corynebacterium glutamicum expressing glutamate decarboxylase from Escherichia coli. Enzyme Microb Technol **51:**171-176.
7. **Ma.** 2012. Structure and mechanism of a glutamate-GABA antiporter. Nature **483:**632-636.
8. **Vo L.** 2012. Effects of glutamate decarboxylase and gamma-aminobutyric acid (GABA) transporter on the bioconversion of GABA in engineered Escherichia coli. Bioprocess Biosyst Eng **35:**645-650.
9. **Stavila.** 2013. Synthesis of lactams using enzyme-catalyzed aminolysis. Tetrahedron Letters **54:**370-372.
10. **Du.** 2001. Hybrid peptide-polyketide natural products: biosynthesis and prospects toward engineering novel molecules. Metab Eng **3:**78-95.
11. **Menzella.** 2005. Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. Nat Biotechnol **23:**1171-1176.
12. **Weissman.** 2008. Protein-protein interactions in multienzyme megasynthetases. Chembiochem **9:**826-848.
13. **Khosla.** 2009. Revisiting the modularity of modular polyketide synthases. Curr Opin Chem Biol **13:**135-143.
14. **Li.** 2009. Analysis of the indanomycin biosynthetic gene cluster from Streptomyces antibioticus NRRL 8167. Chembiochem **10:**1064-1072.
15. **Yadav.** 2009. Towards prediction of metabolic products of polyketide synthases: an in silico analysis. PLoS Comput Biol **5:**e1000351.
16. **Du.** 2010. PKS and NRPS release mechanisms. Nat Prod Rep **27:**255-278.
17. **Tran.** 2010. Insights into protein-protein and enzyme-substrate interactions in modular polyketide synthases. Chem Biol **17:**705-716.
18. **Wong.** 2010. Protein-protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases. Biochemistry **49:**95-102.
19. **Dutta.** 2014. Structure of a modular polyketide synthase. Nature **510:**512-517.
20. **Moorea.** 2002. Biosynthesis and attachment of novel bacterial polyketide synthase starter units. Natural Product Reports **19:**70-99.
21. **Ogasawara.** 2004. Cloning, sequencing, and functional analysis of the biosynthetic gene cluster of macrolactam antibiotic vicenistatin in Streptomyces halstedii. Chem Biol **11:**79-86.
22. **Li.** 2005. Biosynthesis of the unique amino acid side chain of butirosin: possible protective-group chemistry in an acyl carrier protein-mediated pathway. Chem Biol **12:**665-675.

23. **Simunovic.** 2006. Myxovirescin A biosynthesis is directed by hybrid polyketide synthases/nonribosomal peptide synthetase, 3-hydroxy-3-methylglutaryl-CoA synthases, and trans-acting acyltransferases. Chembiochem **7:**1206-1220.
24. **Llewellyn.** 2007. Biosynthesis of butirosin: transfer and deprotection of the unique amino acid side chain. Chem Biol **14:**379-386.
25. **Kalaitzis.** 2009. In vitro biosynthesis of unnatural enterocin and wailupemycin polyketides. J Nat Prod **72:**469-472.
26. **Shinohara.** 2011. A natural protecting group strategy to carry an amino acid starter unit in the biosynthesis of macrolactam polyketide antibiotics. J Am Chem Soc **133:**18134-18137.
27. **Studier W.** 2005. Protein production by auto-induction in high-density shaking cultures. Protein Expression and Purification **41:**207-234.
28. **Lee TS, Krupa RA, Zhang F, Hajimorad M, Holtz WJ, Prasad N, Lee SK, Keasling JD.** 2011. BglBrick vectors and datasheets: A synthetic biology platform for gene expression. J Biol Eng **5:**12.
29. **Tomoya Baba TA, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A Datsenko, Masaru Tomita, Barry L Wanner, Hirotada Mori.** 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Molecular Systems Biology **2**.
30. **Bernal.** 2008. Role of betaine:CoA ligase (CaiC) in the activation of betaines and the transfer of coenzyme A in Escherichia coli. J Appl Microbiol **105:**42-50.
31. **Pennacchietti.** 2009. Mutation of His465 alters the pH-dependent spectroscopic properties of Escherichia coli glutamate decarboxylase and broadens the range of its activity toward more alkaline pH. J Biol Chem **284:**31587-31596.
32. **Dougherty.** 2009. Directed evolution: new parts and optimized function. Curr Opin Biotechnol **20:**486-491.
33. **Medema.** 2011. Synthetic biology in Streptomyces bacteria. Methods Enzymol **497:**485-502.
34. **Choi.** 2015. Enhanced production of gamma-aminobutyrate (GABA) in recombinant Corynebacterium glutamicum by expressing glutamate decarboxylase active in expanded pH range. Microbial Cell Factories **14:**21.
35. **Lothar.** 2005. Handbook of Corynebacterium glutamicum. CPC Press, Taylor & Francis Group.
36. **Nishihara.** 1998. Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GrpE and GroEL-GroES in Assisting Folding of an Allergen of Japanese Cedar Pollen, Cryj2, in Escherichia coli. APPLIED AND ENVIRONMENTAL MICROBIOLOGY **64:**1694–1699.
37. **Nishihara.** 2000. Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in Escherichia coli. APPLIED AND ENVIRONMENTAL MICROBIOLOGY **66:**884–889.
38. **Ruusala.** 1984. Hyper-accurate ribososomes inhibit growth. The EMBO Journal **3:**.2575-2580.
39. **Fluitt.** 2007. Ribosome kinetics and aa-tRNA competition determine rate and fidelity of peptide synthesis. Comput Biol Chem **31:**335-346.
40. **Proshkin.** 2010. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science **328:**504-508.
41. **Siller.** 2010. Slowing bacterial translation speed enhances eukaryotic protein folding efficiency. J Mol Biol **396:**1310-1318.
42. **Capitani.** 2003. Crystal structure and functional analysis of Escherichia coli glutamate decarboxylase. EMBO J **22:**4027-4037.
43. **McAlpine.** 2005. Microbial Genomics as a Guide to Drug Discovery and Structural Elucidation: ECO-02301, a Novel Antifungal Agent, as an Example. J Nat Prod**:**493-496.
44. **Sakuda.** 1996. Novel linear polyene antibiotics: linearmycins. Journal of the Chemical Society, Perkin Transactions 1 doi:10.1039/p19960002315**:**2315.

45. **Zazopoulos.** 2003. A genomics-guided approach for discovering and expressing cryptic metabolic pathways. Nat Biotechnol **21:**187-190.
46. **Zhang.** 2010. A Three Enzyme Pathway for 2-Amino-3-hydroxycyclopent-2-enone Formation and Incorporation in Natural Product Biosynthesis. J AM CHEM SOC **132:**6402-6411.
47. **Bachmann.** 2008. Polyene polyketides, processes for their production and their use as a pharmaceutical.
48. **PCI-Nylon.** 2011. Polyamide 6. http://pcinylon.com/index.php/markets-covered/polyamide-6. Accessed 2015-10-08.
49. **Liu P, Zhang H, Lv M, Hu M, Li Z, Gao C, Xu P, Ma C.** 2014. Enzymatic production of 5-aminovalerate from L-lysine using L-lysine monooxygenase and 5-aminovaleramide amidohydrolase. Sci Rep **4:**5657.
50. **Park SJ, Kim EY, Noh W, Park HM, Oh YH, Lee SH, Song BK, Jegal J, Lee SY.** 2013. Metabolic engineering of Escherichia coli for the production of 5-aminovalerate and glutarate as C5 platform chemicals. Metab Eng **16:**42-47.
51. **Park SJ, Oh YH, Noh W, Kim HY, Shin JH, Lee EG, Lee S, David Y, Baylon MG, Song BK, Jegal J, Lee SY, Lee SH.** 2014. High-level conversion of L-lysine into 5-aminovalerate that can be used for nylon 6,5 synthesis. Biotechnol J **9:**1322-1328.
52. **Josef Ritz; Hugo Fuchs; Heinz Kieczka; William C. Moran, , , doi:10.1002/14356007.a05_031.pub2.** 2005. Caprolactam. Ullmann's Encyclopedia of Industrial Chemistry.
53. **Kallifidas D, Thomas D, Doughty P, Paget MS.** 2010. The sigmaR regulon of Streptomyces coelicolor A32 reveals a key role in protein quality control during disulphide stress. Microbiology **156:**1661-1672.
54. **Stefan C.H.J. Turk WPK, Dennis K Ninaber, Karin P.A.M Kolen, Julia Knutova, Erwin Suir, Martin Schürmann, Petronella C. Raemakers-Franken, Monica Muller, Stefaan M. A. De Wildeman, Leonie M Raamsdonk, Ruud van der Pol, Liang Wu, Margarida F Temudo, Rob van der Hoeven, Michiel Akeroyd, Roland E van der Stoel, Henk J. Noorman, Roel A.L. bovenberg, and Axel C. Trefzer.** 2015. Metabolic engineering towards sustainable production of Nylon-6. ACS Synthetic Biology (Just Accepted) doi:10.1021/acssynbio.5b00129.
55. **Stavila E, Loos K.** 2013. Synthesis of lactams using enzyme-catalyzed aminolysis. Tetrahedron Letters **54:**370-372.
56. **Hui Hong TF, Peter F. Leadlay.** 2013. A Common Origin for Guanidinobutanoate Starter Units in Antifungal Natural Products. Angewandte Chemie **124:**13334–13337.
57. **Zhang J, Kao, E., Wang G., Baidoo E. E.K., Chen, M., Keasling, J.** 2016. Metabolic Engineering of E. coli for the biosynthesis of 2-pyrrolidone. Met Eng Com accepted.
58. **Yuzawa S, Chiba N, Katz L, Keasling JD.** 2012. Construction of a part of a 3-hydroxypropionate cycle for heterologous polyketide biosynthesis in Escherichia coli. Biochemistry **51:**9779-9781.
59. **Zhang W, Heemstra JR, Jr., Walsh CT, Imker HJ.** 2010. Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins. Biochemistry **49:**9946-9947.
60. **Zhang W, Tang Y.** 2009. Chapter 16 In Vitro Analysis of Type II Polyketide Synthase. **459:**367-393.
61. **Gregory Bokinsky EEKB, Swetha Akella, Helcio Burd, Daniel Weaver, Jorge Alonso-Gutierrez, Héctor García-Martín, Taek Soon Lee, Jay D. Keasling.** 2013. HipA-Triggered Growth Arrest and β-Lactam Tolerance in Escherichia coli Are Mediated by RelA-Dependent ppGpp Synthesis. Journal of Bacteriology **195:**3173-3182.
62. **Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, Khandurina J, Trawick JD, Osterhout RE, Stephen R, Estadilla J, Teisan S, Schreyer HB, Andrae S, Yang TH, Lee SY, Burk MJ,**

Van Dien S. 2011. Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. Nat Chem Biol **7:**445-452.
63. **Adkins J, Pugh S, McKenna R, Nielsen DR.** 2012. Engineering microbial chemical factories to produce renewable "biomonomers". Front Microbiol **3:**313.
64. **Beardslee T, Picataggio S.** 2012. Bio-based adipic acid from renewable oils. Lipid Technology **24:**223-225.
65. **Keasling JD.** 2010. Manufacturing molecules through metabolic engineering. Science **330:**1355-1358.
66. **McKenna R, Nielsen DR.** 2011. Styrene biosynthesis from glucose by engineered E. coli. Metab Eng **13:**544-554.
67. **Nikolau BJ, Perera MA, Brachova L, Shanks B.** 2008. Platform biochemicals for a biorenewable chemical industry. Plant J **54:**536-545.
68. **Qian ZG, Xia XX, Lee SY.** 2009. Metabolic engineering of Escherichia coli for the production of putrescine: a four carbon diamine. Biotechnol Bioeng **104:**651-662.
69. **(ed).** 2005. Ullmann's Encyclopedia of Industrial Chemistry. Weinheim: Wiley-VCH, Accessed
70. **Amin.** 2013. Prediction and experimental validation of enzyme substrate specificity in protein structures. Proc Natl Acad Sci U S A **110:**E4195-4202.
71. **Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K.** 2012. Engineering the third wave of biocatalysis. Nature **485:**185-194.
72. **Dietrich.** 2010. High-throughput metabolic engineering: advances in small-molecule screening and selection. Annu Rev Biochem **79:**563-590.
73. **Esvelt.** 2011. A system for the continuous directed evolution of biomolecules. Nature **472:**499-503.
74. **Evans BS, Chen Y, Metcalf WW, Zhao H, Kelleher NL.** 2011. Directed evolution of the nonribosomal peptide synthetase AdmK generates new andrimid derivatives in vivo. Chem Biol **18:**601-607.
75. **Shiue.** 2014. Improving D-glucaric acid production from myo-inositol in E. coli by increasing MIOX stability and myo-inositol transport. Metab Eng **22:**22-31.
76. **Tyo KE, Nevoigt E, Stephanopoulos G.** 2011. Directed evolution of promoters and tandem gene arrays for customizing RNA synthesis rates and regulation. Methods Enzymol **497:**135-155.
77. **Wang HH, Church GM.** 2011. Multiplexed genome engineering and genotyping methods applications for synthetic biology and metabolic engineering. Methods Enzymol **498:**409-426.
78. **Ackerley EMJGJNCD.** 2014. Directed Evolution Library Creation. doi:10.1007/978-1-4939-1053-3.
79. **Tsvetanova B, Peng L, Liang X, Li K, Yang JP, Ho T, Shirley J, Xu L, Potter J, Kudlicki W, Peterson T, Katzen F.** 2011. Genetic assembly tools for synthetic biology. Methods Enzymol **498:**327-348.
80. **Peng Xu MAGK.** 2013. Assembly of Multi-gene Pathways and Combinatorial Pathway Libraries Through ePathBrick Vectors. Methods in Molecular Biology **1073** 107-129.
81. **Jiang W BD, Cox D, Zhang F, Marraffini LA.** 2013 RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol **31:**233-239.
82. **Auer TO, Duroure K, De Cian A, Concordet JP, Del Bene F.** 2014. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. Genome Res **24:**142-153.
83. **Baena-Lopez LA, Alexandre C, Mitchell A, Pasakarnis L, Vincent JP.** 2013. Accelerated homologous recombination and subsequent genome modification in Drosophila. Development **140:**4818-4825.
84. **Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA.** 2013. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res **41:**7429-7437.

85. **Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadan AH, Moineau S.** 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature **468:**67-71.
86. **Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS.** 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell **154:**442-451.
87. **Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA.** 2010. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. Science **329:**1355-1358.
88. **Heigwer F, Kerr G, Boutros M.** 2014. E-CRISP: fast CRISPR target site identification. Nat Methods **11:**122-123.
89. **LaFlamme B.** 2014. A CRISPR method for genome-wide screening. Nature Genetics **46:**99-99.
90. **Marraffini LA, Sontheimer EJ.** 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science **322:**1843-1845.
91. **Menke DB.** 2013. Engineering subtle targeted mutations into the mouse genome. Genesis **51:**605-618.
92. **Pyzocha NK, Ran FA, Hsu PD, Zhang F.** 2014. RNA-guided genome editing of mammalian cells. Methods Mol Biol **1114:**269-277.
93. **Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F.** 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell **154:**1380-1389.
94. **Terns RM, Terns MP.** 2014. CRISPR-based technologies: prokaryotic defense weapons repurposed. Trends Genet **30:**111-118.
95. **Alper H SG.** 2007 Global transcription machinery engineering: a new approach for improving cellular phenotype. Metab Eng **9:**258-267.
96. **Lynch MD WT, Gill RT.** 2007 SCALEs: multiscale analysis of library enrichment. Nat Methods **4:**87-93.
97. **Nicholls AM, G. B.; Sheridan, R. P.; Good, A. C.; Warren, G.; Mathieu, M.; Muchmore, S. W.; Brown, S. P.; Grant, J. A.; Haigh, J. A.; Nevins, N.; Jain, A. N.; Kelley, B. J. .** 2010. Molecular Shape and Medical Chemistry: A Perspective. Med Chem **53:**3862– 3886.
98. **Anonymous.** Pubchem Struture Search.
99. **Anonymous.** 2013. Bacterial Gene Regulation and Transcriptional Networks. Caister Academic Press.
100. **Susanne B. von Bodman1 JMW, and Stephen P. Diggle.** 2008 Cell-Cell Communication in Bacteria: United We Stand. J Bacteriol **190:**4377-4391.
101. **Zhang F, Keasling J.** 2011. Biosensors and their applications in microbial metabolic engineering. Trends Microbiol **19:**323-329.
102. **Mary J Dunlop ZYD, Heather L Szmidt, Hou Cheng Chu, Taek Soon Lee, Jay D Keasling, Masood Z Hadi, Aindrila Mukhopadhyay.** 2011. Engineering microbial biofuel tolerance and export using efflux pumps. Molecular Systems Biology **7**.
103. **Anonymous.** 2014. Bacterial Metabolism. Academic Press.
104. **Chang.** 2015. BRENDA in 2015: exciting developments in its 25th year of existence. Nucleic Acids Res **43:**D439-446.
105. **Herai S, Hashimoto Y, Higashibata H, Maseda H, Ikeda H, Omura S, Kobayashi M.** 2004. Hyper-inducible expression system for streptomycetes. Proc Natl Acad Sci U S A **101:**14031-14035.
106. **Hidenobu komeda Yh, Michihiko kobayashi, Sakayu shimizu.** 1996. Transcriptional regulation of the Rhodococcus rhodochrous J1 nitA gene encoding a nitrilase. Proc Natl Acad Sci **93:**10572-10577.

107. **Pandey AK, Raman S, Proff R, Joshi S, Kang CM, Rubin EJ, Husson RN, Sassetti CM.** 2009. Nitrile-inducible gene expression in mycobacteria. Tuberculosis (Edinb) **89:**12-16.
108. **Toru Nagasawa TN, Hideaki Yamada.** 1990. Caprolactam a new powerful inducer for the formation of Rhodococcus rhodochrous J1 nitrilase. Arch Microbiol **155:**13-17.
109. **Brautaset T, Lale R, Valla S.** 2009. Positively regulated bacterial expression systems. Microb Biotechnol **2:**15-30.
110. **Hiroaki Iwaki YH, Masahiro Teraoka, Tai Tokuyama, Hélène Bergeron, Peter C. K. Lau.** 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.
111. **Iwaki H, Hasegawa Y, Teraoka M, Tokuyama T, Bernard L, Lau PCK.** 2002. Cyclohexanol Biodegradation Genes: A Pathway of Opportunities. **840:**80-92.
112. **Qiong Cheng SMT, Kristy Kostichka, James R. Valentine and Vasantha Nagarajan.** 2000
Genetic Analysis of a Gene Cluster for Cyclohexanol Oxidation in Acinetobacter sp. Strain SE19 by In Vitro Transposition. J Bacteriol **182:**4744.
113. **Steigedal M, 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.
114. **Thomas L. Ruegg E-MK, Blake A. Simmons, Jay D. Keasling, Steven W. Singer, Taek Soon Lee, Michael P. Thelen.** 2014 An auto-inducible mechanism for ionic liquid resistance in microbial biofuel production. Nature Communications **5**.
115. **Gomez-Escribano JP, Bibb MJ.** 2011. Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters. Microb Biotechnol **4:**207-215.
116. **Tiaden A, T. Spirig, S. S. Weber, H. Brüggemann, R. Bosshard, C. Buchrieser, and H. Hilbi.** 2007. The Legionella pneumophila response regulator LqsR promotes host cell interactions as an element of the virulence regulatory network controlled by RpoS and LetA. Cell Microbiol **9:**2902-2920.
117. **Higgins DA, M. E. Pomianek, C. M. Kraml, R. K. Taylor, M. F. Semmelback, and B. L. Bassler, .** 2007. The major Vibrio cholerae autoinducer and its role in virulence factor production. Nature **450:**883-886.
118. **Fuqua WC, S. C. Winans, and E. P. Greenberg.** 1994. Quorum sensing in bacteria: the LuxR/LuxI family of cell density-responsive transcriptional regulators. J Bacteriol **176:**269-275.
119. **Fouhy Y, K. Scanlon, K. Schouest, C. Spillane, L. Crossman, M. B. Avison, R. P. Ryan, and J. M. Dow.** 2007. DSF-dependent cell-cell signaling and virulence in the nosocomial pathogen Stenotrophomonas maltophilia. J Bacteriol **189:**4964-4968.
120. **Chen H, and G. R. Fink.** 2006. Feedback control of morphogenesis in fungi by aromatic alcohols. Genes Dev **20:**1150-1161.
121. **Cugini C, M. Calfee, J. M. Farrow III, D. K. Morales, E. C. Pesci, and D. A. Hogan.** 2007. Farnesol, a common sesquiterpene, inhibits PQS production in Pseudomonas aeruginosa. Mol Microbiol **65:**896-906.
122. **Déziel E, S. S. Gopalan, A. P. Tampakaki, F. Lépine, K. E. Padfield, M. Saucier, G. Xiao, and L. G. Rahme.** 2005. The contribution of MvfR toPseudomonas aeruginosa pathogenesis and quorum sensing circuitry regulation: multiple quorum sensing-regulated genes are modulated without affecting lasRI, rhlRI or the production of N-acyl-L-homoserine lactones. . Mol Microbiol **55:**998-1014.
123. **J. Lee AJ, T. K. Wood.** 2007. Indole is an inter-species biofilm signal mediated by SdiA. BMC Microbiology **7**.

124. **Walters M, V. Sperandio.** 2006. Autoinducer 3 and epinephrine signaling in the kinetics of locus of enterocyte effacement gene expression in enterohemorrhagic Escherichia coli. Infect Immun **74:**5445-5555.

Appendix

Sequence of synthetized DNA and plasmids

gBlock *davB* sequence

tataggggaattgtgagcggataacaatttcagaattcaaaagatcttttaagaaggagatatacatatgaacaagaagaaccgccaccccg
ccgacggcaagaagccgatcaccatttttcggcccggacttccctttgctttcgacgactggctggaacacccggcaggcctgggcagcat
tccggctgagcgccatggggaagaggtggccattgtcggtgccggtatcgccggcctggtagcggcctacgagctgatgaagctgggcc
tcaagccggtggtgtacgaggcttccaagctgggcggccggctgcgctcgcaagccttcaatggcactgacgggatcgttgccgagctgg
gtggcatgcgcttcccggtgtcgtccaccgccttctaccactacgtcgacaagctgggcctggagaccaagcccttccccaacccgctgac
cccggcttcgggcagcacggtgatcgacctggaaggccagacctactacgccgagaagcccaccgacctgccgcaactgtttcatgagg
tagccgacgcttgggccgatgctctggagagcggtgcgcagttcgccgatatccagcaggccatccgcgaccgtgatgtaccgcgcctga
aggaactctggaacaagctggtgccactgtgggacgaccgcaccttctacgacttcgtcgccacctcgcgctcttttgccaagctgagcttc
cagcaccgcgaagtgttcggccaggtcggtttcggcaccggcggttgggactcggacttccccaactcgatgctggaaatcttccgcgtgg
tgatgaccaactgcgacgaccaccagcacctggtggtcgggggcgtggaacaagtgccacaaggcatctggcgcgacgtaccggaacg
ctgcgtgcattggccagagggcaccagcctgagcacgctgcatggcggcgcaccgcgtaccggggtcaagcgcattgcccgcgccgcc
gatggccgcctggcggtcaccgacaactggggcgatacccgccactacagcgcagtactcgccacctgccagacctggttgctgaccac
ccagatcgactgcgaggaatcgctgttctcgcaaaagatgtggatggccctggaccgtacccgctacatgcagtcgtcgaaaaccttcgtc
atggtcgaccgcccgttctggaaggacaaggacccggaaaccggccgtgacctgctgagcatgaccctcaccgaccgcctcacccgcg
gcacttacctgttcgacaacggcaacgacaagcccggggtgatctgcctgtcgtactcgtggatgagcgacgcgctgaagatgctgccgc
acccggtggaaaagcgcgtacaactggccctggatgcgctgaagaagatctacccgaagaccgatatcgccgggcacatcatcggcgac
ccgatcacggtttcctgggaggccgacccgtacttcctcggcgccttcaaaggcgcgcttccgggccattaccgctacaaccagcgcatgt
acgcgcacttcatgcagcaggacatgccggcggagcagcgcggtatcttcattgccggtgacgacgtgtcatggacccccgcctgggttg
aaggcgcggtgcagacgtcgctgaatgcggtgtggggtatcatgaaccactttggtggccacacccaccccgacaaccccggcccgggc
gatgtgttcaacgaaatcggcccgatcgccctggcggattgaggatcttttaagaaggagatatacatatgcgcatcgctct

gBlock *davA* sequence

cccgatcgccctggcggattgaggatcttttaagaaggagatatacatatgcgcatcgctctgtaccagggcgcacccaagccactggatgt
gcccggcaacctgcaacggctgcgccaccaggcgcagttggcagccgaccgcggcgcacagttgctggtgtgcccggagatgttcctgt
ccggctacaacatcggcctggcccaggtcgagcgcctggccgaggccgccgatggcccggcagccatgacggtggtggagattgccc
aggcgcaccgtatcgccattgtctatggctacccggagcgcggcgatgacggggcgatctacaacagcgtgcagctgatcgatgcgcatg
gccgcagcctgagcaattaccgcaagacccacctgttcggtgaactggaccgctcgatgttcagccctggtgcggaccacttcccggtggt
ggaactggaaggctggaaggttggcctgctgatctgctacgacatcgagttcccggagaacgcccgacgcctggcgctggacggcgcc
gagctgatcctggtgccgacggcgaacatgacgccgtacgactttacctgccaggtgaccgtgagggcacgggcgcaggaaaaccagt
gctacctggtatatgccaactactgcggcgcggaagacgagatcgagtattgcgggcagagcagcatcatcggcccggatggcagcttgc
tggccatggccgggcgggatgagtgccagttgttggcagagctcgagcatgagcgggtggtgcaggggcgcagggcgtttccctacctg
accgatttgcgccaggagctgcacctgcgtaaaggctgaggatccaaactcgagtaaggatctccaggcatcaaataaaacgaaaggctc
agtc

pBbA7a-DavB-DavA

gacgtcctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgt
attgggcgccagggtggtttttcttttcaccagtgagacgggcaacagctgattgcccttcaccgcctggccctgagagagttgcagcaagc
ggtccacgctggtttgccccagcaggcgaaaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcc
cactaccgagatgtccgcaccaacgcgcagcccggactcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaaccagcat
cgcagtgggaacgatgccctcattcagcatttgcatggtttgttgaaaaccggacatggcactccagtcgccttcccgttccgctatcggctg
aatttgattgcgagtgagatatttatgccagccagccagacgcagacgcgccgagacagaacttaatgggcccgctaacagcgcgatttgc

tggtgacccaatgcgaccagatgctccacgcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtgtctggtcagagac
atcaagaaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctggtcatccagcggatagttaatgatcagcccactg
acgcgttgcgcgagaagattgtgcaccgccgctttacaggcttcgacgccgcttcgttctaccatcgacaccaccacgctggcacccagttg
atcggcgcgagatttaatcgccgcgacaatttgcgacggcgcgtgcagggccagactggaggtggcaacgccaatcagcaacgactgttt
gcccgccagttgttgtgccacgcggttgggaatgtaattcagctccgccatcgccgcttccacttttcccgcgttttcgcagaaacgtggctg
gcctggttcaccacgcgggaaacggtctgataagagacaccggcatactctgcgacatcgtataacgttactggtttcacattcaccaccctg
aattgactctcttccgggcgctatcatgccataccgcgaaaggttttgcgccattcgatggtgtccgggatctcgacgctctcccttatgcgact
cctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgcaaggaatggtgcatgcaaggagatggcgcccaa
cagtcccccggccacggggcctgccaccatacccacgccgaaacaagcgctcatgagcccgaagtggcgagcccgatcttccccatcg
gtgatgtcggcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacgatgcgtccggcgtagaggatcgagatcg
atctcgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaatttcagaattcaaaagatcttttaagaaggagatat
acatatgaacaagaagaaccgccacccccgccgacggcaagaagccgatcaccattttcggcccggacttccctttgctttcgacgactgg
ctggaacacccggcaggcctgggcagcattccggctgagcgccatggggaagaggtggccattgtcggtgccggtatcgccggcctggt
agcggcctacgagctgatgaagctgggcctcaagccggtggtgtacgaggcttccaagctgggcggccggctgcgctcgcaagccttca
atggcactgacgggatcgttgccgagctgggtggcatgcgcttcccggtgtcgtccaccgccttctaccactacgtcgacaagctgggcct
ggagaccaagcccttccccaacccgctgaccccggcttcgggcagcacggtgatcgacctggaaggccagacctactacgccgagaag
cccaccgacctgccgcaactgtttcatgaggtagccgacgcttgggccgatgctctggagagcggtgcgcagttcgccgatatccagcag
gccatccgcgaccgtgatgtaccgcgcctgaaggaactctggaacaagctggtgccactgtgggacgaccgcaccttctacgacttcgtc
gccacctcgcgctcttttgccaagctgagcttccagcaccgcgaagtgttcggccaggtcggtttcggcaccggcggttgggactcggactt
ccccaactcgatgctggaaatcttccgcgtggtgatgaccaactgcgacgaccaccagcacctggtggtcgggggcgtggaacaagtgc
cacaaggcatctggcgcgacgtaccggaacgctgcgtgcattggccagagggcaccagcctgagcacgctgcatggcggcgcaccgc
gtaccggggtcaagcgcattgcccgcgccgccgatggccgcctggcggtcaccgacaactggggcgatacccgccactacagcgcagt
actcgccacctgccagacctggttgctgaccacccagatcgactgcgaggaatcgctgttctcgcaaaagatgtggatggccctggaccgt
acccgctacatgcagtcgtcgaaaaccttcgtcatggtcgaccgcccgttctggaaggacaaggacccggaaaccggccgtgacctgctg
agcatgaccctcaccgaccgcctcacccgcggcacttacctgttcgacaacggcaacgacaagcccggggtgatctgcctgtcgtactcgt
ggatgagcgacgcgctgaagatgctgccgcacccggtggaaaagcgcgtacaactggccctggatgcgctgaagaagatctacccgaa
gaccgatatcgccgggcacatcatcggcgacccgatcacggtttcctgggaggccgacccgtacttcctcggcgccttcaaaggcgcgct
tccgggccattaccgctacaaccagcgcatgtacgcgcacttcatgcagcaggacatgccggcggagcagcgcggtatcttcattgccgg
tgacgacgtgtcatggacccccgcctgggttgaaggcgcggtgcagacgtcgctgaatgcggtgtggggtatcatgaaccactttggtggc
cacacccaccccgacaaccccggcccgggcgatgtgttcaacgaaatcggcccgatcgccctggcggattgaggatcttttaagaagga
gatatacatatgcgcatcgctctgtaccagggcgcacccaagccactggatgtgcccggcaacctgcaacggctgcgccaccaggcgca
gttggcagccgaccgcggcgcacagttgctggtgtgcccggagatgttcctgtccggctacaacatcggcctggcccaggtcgagcgcct
ggccgaggccgccgatggcccggcagccatgacggtggtggagattgcccaggcgcaccgtatcgccattgtctatggctacccggagc
gcggcgatgacggggcgatctacaacagcgtgcagctgatcgatgcgcatggccgcagcctgagcaattaccgcaagacccacctgttc
ggtgaactggaccgctcgatgttcagccctggtgcggaccacttcccggtggtggaactggaaggctggaaggttggcctgctgatctgct
acgacatcgagttcccggagaacgcccgacgcctggcgctggacggcgccgagctgatcctggtgccgacggcgaacatgacgccgta
cgactttacctgccaggtgaccgtgagggcacgggcgcaggaaaaccagtgctacctggtatatgccaactactgcggcgcggaagacg
agatcgagtattgcgggcagagcagcatcatcggcccggatggcagcttgctggccatggccgggcgggatgagtgccagttgttggca
gagctcgagcatgagcgggtggtgcaggggcgcagggcgtttccctacctgaccgatttgcgccaggagctgcacctgcgtaaaggctg
aggatccaaactcgagtaaggatctccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcgg
tgaacgctctctactagagtcacactggctcaccttcgggtgggcctttctgcgtttatacctagggatatattccgcttcctcgctcactgactc
gctacgctcggtcgttcgactgcggcgagcggaaatggcttacgaacggggcggagatttcctggaagatgccaggaagatacttaacag
ggaagtgagagggccgcggcaaagccgtttttccataggctccgccccccctgacaagcatcacgaaatctgacgctcaaatcagtggtgg
cgaaacccgacaggactataaagataccaggcgtttcccccctggcggctccctcgtgcgctctcctgttcctgccttttcggtttaccggtgtca
ttccgctgttatggccgcgtttgtctcattccacgcctgacactcagttccgggtaggcagttcgctccaagctggactgtatgcacgaacccc
ccgttcagtccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggaaagacatgcaaaagcaccactggcagcagccac

tggtaattgatttagaggagttagtcttgaagtcatgcgccggttaaggctaaactgaaaggacaagttttggtgactgcgctcctccaagcca
gttacctcggttcaaagagttggtagctcagagaaccttcgaaaaaccgccctgcaaggcggttttttcgttttcagagcaagagattacgcg
cagaccaaaacgatctcaagaagatcatcttattaatcagataaaatatttctagatttcagtgcaatttatctcttcaaatgtagcacctgaagtc
agccccatacgatataagttgttactagtgcttggattctcaccaataaaaaacgcccggcggcaaccgagcgttctgaacaaatccagatg
gagttctgaggtcattactggatctatcaacaggagtccaagcgagctcgtaaacttggtctgacagttaccaatgcttaatcagtgaggcac
ctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggcccc
agtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaa
gtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttg
ttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcc
cccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcag
cactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggc
gaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcgg
ggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcacc
agcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttc
ctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggttccgcg
cacatttccccgaaaagtgccacct

pET28a-MBP-ORF27

tggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccc
tagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctcccttttagggttcc
gatttagtgctttacggcacctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgccc
tttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggat
tttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaatttcaggtggca
cttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgaattaattcttagaaaaactcatcg
agcatcaaatgaaactgcaatttattcatatcaggattatcaataccatatttttgaaaaagccgtttctgtaatgaaggagaaaactcaccgag
gcagttccataggatggcaagatcctggtatcggtctgcgattccgactcgtccaacatcaatacaacctattaatttcccctcgtcaaaaataa
ggttatcaagtgagaaatcaccatgagtgacgactgaatccggtgagaatggcaaaagtttatgcatttctttccagacttgttcaacaggcca
gccattacgctcgtcatcaaaatcactcgcatcaaccaaaccgttattcattcgtgattgcgcctgagcgagacgaaatacgcgatcgctgtta
aaaggacaattacaaacaggaatcgaatgcaaccggcgcaggaacactgccagcgcatcaacaatattttcacctgaatcaggatattcttc
taatacctggaatgctgttttcccggggatcgcagtggtgagtaaccatgcatcatcaggagtacggataaaatgcttgatggtcggaagag
gcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttgccatgtttcagaaacaactctggcgcat
cgggcttcccatacaatcgatagattgtcgcacctgattgcccgacattatcgcgagcccatttatacccatataaatcagcatccatgttggaa
tttaatcgcggcctagagcaagacgtttcccgttgaatatggctcataacacccccttgtattactgtttatgtaagcagacagttttattgttcatga
ccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaa
tctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggctt
cagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgc
tctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgca
gcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctat
gagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgaggga
gcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggg
cggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccct
gattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgag
gaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctct
gatgccgcatagttaagccagtatacactccgctatcgctacgtgactgggtcatggctgcgccccgacacccgccaacacccgctgacgc
gccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatc
accgaaacgcgcgaggcagctgcggtaaagctcatcagcgtggtcgtgaagcgattcacagatgtctgcctgttcatccgcgtccagctcg

ttgagtttctccagaagcgttaatgtctggcttctgataaagcgggccatgttaagggcggttttttcctgtttggtcactgatgcctccgtgtaag
ggggatttctgttcatgggggtaatgataccgatgaaacgagagaggatgctcacgatacgggttactgatgatgaacatgcccggttactg
gaacgttgtgagggtaaacaactggcggtatggatgcggcgggaccagagaaaaatcactcagggtcaatgccagcgcttcgttaataca
gatgtaggtgttccacagggtagccagcagcatcctgcgatgcagatccggaacataatggtgcagggcgctgacttccgcgtttccagac
tttacgaaacacggaaaccgaagaccattcatgttgttgctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatcg
gtgattcattctgctaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggggcc
gccatgccggcgataatggcctgcttctcgccgaaacgtttggtggcgggaccagtgacgaaggcttgagcgagggcgtgcaagattccg
aataccgcaagcgacaggccgatcatcgtcgcgctccagcgaaagcggtcctcgccgaaaatgacccagagcgctgccggcacctgtc
ctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcgcccaccggaaggagctgactgggttgaag
gctctcaagggcatcggtcgagatcccggtgcctaatgagtgagctaacttacattaattgcgttgcgctcactgcccgctttccagtcggga
aacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgccagggtggtttttctttcacca
gtgagacgggcaacagctgattgcccttcaccgcctggccctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcga
aaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcaccaacgcgc
agcccggactcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaaccagcatcgcagtgggaacgatgccctcattcagc
atttgcatggtttgttgaaaaccggacatggcactccagtcgccttcccgttccgctatcggctgaatttgattgcgagtgagatatttatgccag
ccagccagacgcagacgcgccgagacagaacttaatgggcccgctaacagcgcgatttgctggtgacccaatgcgaccagatgctccac
gcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtgtctggtcagagacatcaagaaataacgccggaacattagtgc
aggcagcttccacagcaatggcatcctggtcatccagcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgc
cgctttacaggcttcgacgccgcttcgttctaccatcgacaccaccacgctggcacccagttgatcggcgcgagatttaatcgccgcgacaa
tttgcgacggcgcgtgcagggccagactggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgttgtgccacgcggttgg
gaatgtaattcagctccgccatcgccgcttccactttttcccgcgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacggtct
gataagagacaccggcatactctgcgacatcgtataacgttactggtttcacattcaccaccctgaattgactctcttccgggcgctatcatgc
cataccgcgaaaggttttgcgccattcgatggtgtccgggatctcgacgctctcccttatgcgactcctgcattaggaagcagcccagtagta
ggttgaggccgttgagcaccgccgccgcaaggaatggtgcatgcaaggagatggcgcccaacagtcccccggccacggggcctgcca
ccatacccacgccgaaacaagcgctcatgagcccgaagtggcgagcccgatcttccccatcggtgatgtcggcgatataggcgccagca
accgcacctgtggcgccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactcacta
tagggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggagatataccatgggcagcagccatcatcatca
tcatcacagcagcggcaaaatcgaagaaggtaaactggtaatctggattaacggcgataaaggctataacggtctcgctgaagtcggtaag
aaattcgagaaagataccggaattaaagtcaccgttgagcatccggataaactggaagagaaattcccacaggttgcggcaactggcgatg
gccctgacattatcttctgggcacacgaccgctttggtggctacgctcaatctggcctgttggctgaaatcaccccggacaaagcgttccagg
acaagctgtatccgtttacctgggatgccgtacgttacaacggcaagctgattgcttacccgatcgctgttgaagcgttatcgctgatttataac
aaagatctgctgccgaacccgccaaaaacctgggaagagatcccggcgctggataaagaactgaaagcgaaaggtaagagcgcgctga
tgttcaacctgcaagaaccgtacttcacctggccgctgattgctgctgacgggggttatgcgttcaagtatgaaaacggcaagtacgacatta
aagacgtgggcgtggataacgctggcgcgaaagcgggtctgaccttcctggttgacctgattaaaaacaaacacatgaatgcagacaccg
attactccatcgcagaagctgcctttaataaaggcgaaacagcgatgaccatcaacggcccgtgggcatggtccaacatcgacaccagca
aagtgaattatggtgtaacggtactgccgaccttcaagggtcaaccatccaaaccgttcgttggcgtgctgagcgcaggtattaacgccgcc
agtccgaacaaagagctggcaaaagagttcctcgaaaactatctgctgactgatgaaggtctggaagcggttaataaagacaaaccgctgg
gtgccgtagcgctgaagtcttacgaggaagagttggcgaaagatccacgtattgccgccactatggaaaacgcccagaaaggtgaaatca
tgccgaacatcccgcagatgtccgctttctggtatgccgtgcgtactgcggtgatcaacgccgccagcggtcgtcagactgtcgatgaagc
cctgaaagacgcgcagactagcagcggcctggtgccgcgcggcagcCATATGCGCCCAATGACCGCTAAAAT
CTTCGCCGTCGACTCCGTCCGTCCGATCGACGAGTTTGAGCAGGACGCACTGCGCGT
TGCGGATGTGATTCGCGAACGTGGCGTGTGTCTGGGTGACCGTGTGATGTTGAAGGC
GGGCAACAGCGCGTCGTACGTTTGCGTTTTGTATGCGCTGATGCACATCGGTGCGAG
CATCGTTTTGGTCGATCAGCAAGAGCATAAAGAGGAAACCCGTCGTATCGCGCTGC
GTACCGGCGTAAAAGTCACGTTTGTGGATGATGAAACCCCGATTGATCAAGATGCG
GACCCGATTCACCTGTACGAGCTGATGGTGGCTACCCAGAACCGTCCTCCGATGGAC

AGCGCACTGAGCTTCGACGCGTGGGGTGAACTGTCTGACGGTCTGATTATGTGGACG
AGCGGCAGCACCGGTAGCCCGAAGGGTGTCGTGAAGAGCGGTGGTAAATTCCTGGC
GAATCTGCGCCGTAACGCGCATCAAGTGGGTCATCGTCCGGATGACGTGCTGATGCC
GCTGCTGCCGTTCGCGCACCAGTACGGTCTGTCTATGGTGCTGATTGCATGGCTGAC
GCGCTGCTCCCTGGTTATTGCGCCATACCGCCGTCTGGATCGTGCTTTGCGTATGGCC
CGTGACAGCGGCACGACCGTTATCGATGCCACGCCGAGCAGCTATCGCAGCATCCT
GGGCCTGGTCACGCGTAAACCGGCCCTGCGTGCACACCTGGCCGGCACCCGCATGTT
CTGTGTGGGCGCAGCGCCGTTGGATGCGCCGCTGGTCGAAAGCTACGTTCAAGAGTT
TGGTCTGCCGCTGTTGGACAGCTATGGTTCTACCGAGCTGAACAATATCGCTTTCGC
GACCCTGGATAATCCGGTTTCCTGTGGTCGCGCAATGGAAGGTATCGGTCTGCGTAT
TGTTGACGAAGATGGTCGTGAAGTTGCGGCAGGCCAACCGGGCGAAATCGAGGTTG
ACACTCCGGATGCCCTGGAGGGTCAAATCGCCGAGGATGGTAGCATTATTCCGGCA
CCGACCGGCTGGCAGCGTACGGGCGATCTGGGTCACTTGGACGCCGACGGCAACCT
GTATGTCCTGGGTCGTAAGTTTGCGGTCCACCGCATGGGTTATACTTTGTACCCAGA
GCTGATTGAGCGCAAAGTGGCCGCTGAGGGCTGCCCGACCCGCATTGTTCCGCTGCC
GGACGAGCTGCGTGGTAGCCAACTGGTCTTTTTCGTGGAAGATGATGAACAGCGTG
ACGCAGGTTACTGGCGTGAACGTCTGTGCGGTTTGCTGCCGGCGTTCGAGCAGCCGA
ACAAGGTGGTCGTTCTGGAGCAGTTTCCTCTGAATCGCAATGGCAAGCCGGACAAG
AAAGAGCTGACCCGTATGGCGGCAGAATGATAAGGATCCgaattcgagctccgtcgacaagcttgcg
gccgcactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctga
gcaataactagcataaccccttggggcctctaaacgggtcttgaggggtttttgctgaaaggaggaactatatccggat

pBbA8a-ChnR

gacgtcttatgacaacttgacggctacatcattcacttttttcttcacaaccggcacggaactcgctcgggctggccccggtgcatttttttaaata
cccgcgagaaatagagttgatcgtcaaaaccaacattgcgaccgacggtggcgataggcatccgggtggtgctcaaaagcagcttcgcct
ggctgatacgttggtcctcgcgccagcttaagacgctaatccctaactgctggcggaaaagatgtgacagacgcgacggcgacaagcaaa
catgctgtgcgacgctggcgatatcaaaattgctgtctgccaggtgatcgctgatgtactgacaagcctcgcgtacccgattatccatcggtg
gatggagcgactcgttaatcgcttccatgcgccgcagtaacaattgctcaagcagatttatcgccagcagctccgaatagcgcccttcccctt
gcccggcgttaatgatttgcccaaacaggtcgctgaaatgcggctggtgcgcttcatccgggcgaaagaaccccgtattggcaaatattgac
ggccagttaagccattcatgccagtaggcgcgcggacgaaagtaaacccactggtgataccattcgcgagcctccggatgacgaccgtag
tgatgaatctctcctggcgggaacagcaaaatatcacccggtcggcaaacaaattctcgtccctgatttttcaccacccccctgaccgcgaatg
gtgagattgagaatataacctttcattcccagcggtcggtcgataaaaaaatcgagataaccgttggcctcaatcggcgttaaacccgccacc
agatgggcattaaacgagtatcccggcagcaggggatcattttgcgcttcagccatacttttcatactcccgccattcagagaagaaaccaat
tgtccatattgcatcagacattgccgtcactgcgtcttttactggctcttctcgctaaccaaaccggtaaccccgcttattaaaagcattctgtaac
aaagcgggaccaaagccatgacaaaaacgcgtaacaaaagtgtctataatcacggcagaaaagtccacattgattatttgcacggcgtcac
actttgctatgccatagcatttttatccataagattagcggatttctacctgacgcttttttatcgcaactctctactgtttctccatacccgtttttttggg
aattcaaaagatcttttaagaaggagatatacatatgagcacagacaaagcaaatacgctgatcaaacccgaagatgtcgtgttatggattcc
gggtaatgtcacaattgacagcatgaatgccggttgggaaaacattgcaatcagagggtacgaatataccaacctcgatgtgcatattcctgc
catgcgtgactacatgatcgtcaactataaaaaaagtgcggcggaaatgcgtagaaaaggcgatgcctcttgggatacccaagtggttaag
ccgggttatgtctccttgttgacctgtggtgaagattcccgctgggcgtggaatgaccatattgccgtcacccatgtctacatttcgcatgactc
catcacctcaatggcgaataaggtgtttgattatgatatcgcttcgatccgaatcagagacgaagtcggtgtggaagatcatgttttacctgctc
tgacttcacttttagaactagaattaaagcaaggtggtttaggtggaaacctgtatttagagagcattaaaaaccagatcgccctgcatttactc
cgtcagtatgccaaattagattttaaggaaggacagtgccgttctggttttactcccctacaacgcagactgttattagaatttatcaatgaaaac
atgagcattaaaattaccctcgaagatttagcgggattagtcaagatgagcgtgcctcatttaatgagaaaatttaaagtcgattttggtaattcc
cctgctgcctacatcatgaatctcagggtgcaatttgctaaacgtttgctcacttcaaaaaaagaaattccactgaaagtgattgccagtgaag

ccggtttttgcgatcagagccatatgacccgagtatttcaaaaatttttgggaaaacacccatcgaaatcagacaggaacacaccaatctcgt
gtctgaaaattcagtctcctctattgttttttgaggatccaaactcgagtaaggatctccaggcatcaaataaaacgaaaggctcagtcgaaag
actgggcctttcgttttatctgttgtttgtcggtgaacgctctctactagagtcacactggctcaccttcgggtgggcctttctgcgtttatacctag
ggatatattccgcttcctcgctcactgactcgctacgctcggtcgttcgactgcggcgagcggaaatggcttacgaacggggcggagatttc
ctggaagatgccaggaagatacttaacagggaagtgagagggccgcggcaaagccgtttttccataggctccgccccctgacaagcatc
acgaaatctgacgctcaaatcagtggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggcggctccctcgtgcgctc
tcctgttcctgccttttcggtttaccggtgtcattccgctgttatggccgcgtttgtctcattccacgcctgacactcagttccgggtaggcagttcg
ctccaagctggactgtatgcacgaacccccccgttcagtccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggaaagac
atgcaaaagcaccactggcagcagccactggtaattgatttagaggagttagtcttgaagtcatgcgccggttaaggctaaactgaaaggac
aagttttggtgactgcgctcctccaagccagttacctcggttcaaagagttggtagctcagagaaccttcgaaaaaccgccctgcaaggcgg
ttttttcgttttcagagcaagagattacgcgcagaccaaaacgatctcaagaagatcatcttattaatcagataaaatatttctagatttcagtgca
atttatctcttcaaatgtagcacctgaagtcagccccatacgatataagttgttactagtgcttggattctcaccaataaaaaacgcccggcggc
aaccgagcgttctgaacaaatccagatggagttctgaggtcattactggatctatcaacaggagtccaagcgagctcgtaaacttggtctgac
agttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacg
atacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagc
cagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagt
tcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcc
caacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggcc
gcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaacca
agtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaa
gtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcaccca
actgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgaca
cggaaatgttgaatactcatactcttccttttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttaga
aaaataaacaaatagggttccgcgcacatttccccgaaaagtgccacct

JZ-BCS-2

accagagggcgccccagctggcaattccgaattcatgagatcttaagtaagtaagagtatacgtatatcggctaataacgtattaaggcgctt
cggcgccttttttatgggggtattttcatcccaatccacacgtccaacgcacagcaaacaccacgtcgaccctatcagctgcgtgctttctatg
gaagaggggaagaggagtggccgcgggatatcactagtgcggccggtggtggtggtccaaactatccgaatatgtaagtctaatcacttca
aggaagaggtttcccatttcccttctttctagcagatgaagaagcttgcaactaaaagagattgtttggatcagttacccaaaatcgttgaaaag
attttaactcttcgatttttatttttttaggtaatcctagccctctcgggggctaggattaaaaattttaagttattccaacacgaatgacaaattgttca
atgcaaaataaaaacatacaatatataaatatatttttttaattaaaacataagattacaataaaataagaatttttatttggagtttgttttttttctacaa
tgatcattatgtacaatttttaggttcaccccatccaagccttgtgattgcattcctgcgattctttattcaatgaataagcaatgctattaatcagca
atgaataaccagcactgcagattttgaataaattcacatgtcgtaatggagggtaccatgcgtaaaggagaagaagataacatggctatcatt
aaagagttcatgcgcttcaaagttcacatggagggttctgttaacggtcacgagttcgagatcgaaggcgaaggcgagggccgtccgtatg
aaggcacccagaccgccaaactgaaagtgactaaaggcggcccgctgccttttgcgtgggacatcctgagcccgcaatttatgtacggttc
taaagcgtatgttaaacacccagcggatatcccggactatctgaagctgtcttttccggaaggtttcaagtgggaacgcgtaatgaattttgaa
gatggtggtgtcgtgaccgtcactcaggactcctccctgcaggatggcgagttcatctataaagttaaactgcgtggtactaattttccatctga
tggcccggtgatgcagaaaaagacgatgggttgggaggcgtctagcgaacgcatgtatccggaagatggtgcgctgaaaggcgaaatta
aacagcgcctgaaactgaaagatggcggccattatgacgctgaagtgaaaaccacgtacaaagccaagaaacctgtgcagctgcctggc
gcgtacaatgtgaatattaaactggacatcacctctcataatgaagattatacgatcgtagagcaatatgagcgcgcggagggtcgtcattcta
ccggtggcatggatgaactatacaaataataaggatctagagaatataaaaagccagattattaatccggctttttattatttggatcctaactc
gagtaaggatctccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacgctctctac
tagagtcacactggctcaccttcgggtgggcctttctgcgtttatacctagggcgttcggctgcggcgagcggtatcagctcactcaaaggc
ggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaag
gccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca

ggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttc
tcccttcgggaagcgtggcgctttctatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacga
acccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcag
ccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggac
agtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcgg
tggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtgga
acgaaaactcacgttaagggattttggtcatgactagtgcttggattctcaccaataaaaaacgcccggcggcaaccgagcgttctgaacaa
atccagatggagttctgaggtcattactggatctatcaacaggagtccaagcgagctctcgaaccccagagtcccgctcagaagaactcgtc
aagaaggcgatagaaggcgatgcgctgcgaatcgggagcggcgataccgtaaagcacgaggaagcggtcagcccattcgccgccaag
ctcttcagcaatatcacgggtagccaacgctatgtcctgatagcggtccgccacacccagccggccacagtcgatgaatccagaaaagcg
gccattttccaccatgatattcggcaagcaggcatcgccatgggtcacgacgagatcctcgccgtcgggcatgcgcgccttgagcctggcg
aacagttcggctggcgcgagcccctgatgctcttcgtccagatcatcctgatcgacaagaccggcttccatccgagtacgtgctcgctcgat
gcgatgtttcgcttggtggtcgaatgggcaggtagccggatcaagcgtatgcagccgccgcattgcatcagccatgatggatactttctcgg
caggagcaaggtgagatgacaggagatcctgccccggcacttcgcccaatagcagccagtcccttcccgcttcagtgacaacgtcgagca
cagctgcgcaaggaacgcccgtcgtggccagccacgatagccgcgctgcctcgtcctgcagttcattcagggcaccggacaggtcggtc
ttgacaaaaagaaccgggcgcccctgcgctgacagccggaacacggcggcatcagagcagccgattgtctgttgtgcccagtcatagcc
gaatagcctctccacccaagcggccggagaacctgcgtgcaatccatcttgttcaatcatgcgaaacgatcctcatcctgtctcttgatcagat
catgatcccctgcgccatcagatccttggcggcaagaaagccatccagtttactttgcagggcttcccaacctt

pBbSLactamC-mCherry

gatatctggcgaaaatgagacgttgatcggcacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgtatttttgagtt
atcgagattttcaggagctaaggaagctaaaatggagaaaaaaatcactggatataccaccgttgatatatcccaatggcatcgtaaagaac
attttgaggcatttcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggccttttaaagaccgtaaagaaaaataagc
acaagttttatccggcctttattcacattcttgcccgcctgatgaatgctcatccggaatttcgtatggcaatgaaagacggtgagctggtgatat
gggatagtgttcacccttgttacaccgttttccatgagcaaactgaaacgttttcatcgctctggagtgaataccacgacgatttccggcagtttc
tacacatatattcgcaagatgtggcgtgttacggtgaaaacctggcctatttccctaaagggtttattgagaatatgtttttcgtctcagccaatcc
ctgggtgagtttcaccagttttgatttaaacgtggccaatatggacaacttcttcgcccccgttttcaccatgggcaaatattatacgcaaggcg
acaaggtgctgatgccgctggcgattcaggttcatcatgccgtttgtgatggcttccatgtcggcagaatgcttaatgaattacaacagtactg
cgatgagtggcagggcggggcgtaatttgatatcgagctcgcttggactcctgttgatagatccagtaatgacctcagaactccatctggattt
gttcagaacgctcggttgccgccgggcgtttttattggtgagaatccaagcactagggacagtaagacgggtaagcctgttgatgataccg
ctgccttactgggtgcattagccagtctgaatgacctgtcacgggataatccgaagtggtcagactggaaaatcagagggcaggaactgct
gaacagcaaaaagtcagatagcaccacatagcagacccgccataaaacgccctgagaagcccgtgacgggcttttcttgtattatgggtag
tttccttgcatgaatccataaaaggcgcctgtagtgccatttacccccattcactgccagagccgtgagcgcagcgaactgaatgtcacgaaa
aagacagcgactcaggtgcctgatggtcggagacaaaaggaatattcagcgatttgcccgagcttgcgagggtgctacttaagcctttagg
gttttaaggtctgttttgtagaggagcaaacagcgtttgcgacatccttttgtaatactgcggaactgactaaagtagtgagttatacacagggc
tgggatctattctttttatctttttttattctttctttattctataaattataaccacttgaatataaacaaaaaaaacacacaaaggtctagcggaattta
cagagggtctagcagaatttacaagttttccagcaaaggtctagcagaatttacagatacccagatcacccgggaaaaggactagtaattatc
attgactagcccatctcaattggtatagtgattaaaatcacctagaccaattgagatgtatgtctgaattagttgttttcaaagcaaatgaactagc
gattagtcgctatgacttaacggagcatgaaaccaagctaattttatgctgtgtggcactactcaaccccacgattgaaaaccctacaaggaa
agaacggacggtatcgttcacttataaccaatacgctcagatgatgaacatcagtagggaaaatgcttatggtgtattagctaaagcaaccag
agagctgatgacgagaactgtggaaatcaggaatcctttggttaaaggctttgagattttccagtggacaaactatgccaagttctcaagcga
aaaattagaattagtttttagtgaagagatattgccttatcttttccagttaaaaaaattcataaaatataatctggaacatgttaagtcttttgaaaac
aaatactctatgaggatttatgagtggttattaaaagaactaacacaaaagaaaactcacaaggcaaatatagagattagccttgatgaatttaa
gttcatgttaatgcttgaaaataactaccatgagtttaaaaggcttaaccaatgggttttgaaaccaataagtaaagatttaaacacttacagcaa
tatgaaattggtggttgataagcgaggccgcccgactgatacgttgattttccaagttgaactagatagacaaatggatctcgtaaccgaactt
gagaacaaccagataaaaatgaatggtgacaaaataccaacaaccattacatcagattcctacctacataacggactaagaaaaacactaca

cgatgctttaactgcaaaaattcagctcaccagttttgaggcaaaatttttgagtgacatgcaaagtaagtatgatctcaatggttcgttctcatg
gctcacgcaaaaacaacgaaccacactagagaacatactggctaaatacggaaggatctgaggttcttatggctcttgtatctatcagtgaag
catcaagactaacaaacaaaagtagaacaactgttcaccgttacatatcaaagggaaaactgtccatatgcacagatgaaaacggtgtaaaa
aagatagatacatcagagcttttacgagtttttggtgcattcaaagctgttcaccatgaacagatcgacaatgtaacagatgaacagcatgtaa
cacctaatagaacaggtgaaaccagtaaaacaaagcaactagaacatgaaattgaacacctgagacaacttgttacagctcaacagtcaca
catagacagcctgaaacaggcgatgctgcttatcgaatcaaagctgccgacaacacgggagccagtgacgcctcccgtggggaaaaaat
catggcaattctggaagaaatagcgctttcagccggcaaaccggctgaagccggatctgcgattctgataacaaactagcaacaccagaac
agcccgtttgcgggcagcaaaacccgtaccctaggtataaacgcagaaaggcccacccgaaggtgagccagtgtgactctagtagagag
cgttcaccgacaaacaacagataaaacgaaaggcccagtctttcgactgagcctttcgttttatttgatgcctggagatccttactcgagtttgg
atccttatttgtatagttcatccatgccaccggtagaatgacgaccctccgcgcgctcatattgctctacgatcgtataatcttcattatgagaggt
gatgtccagtttaatattcacattgtacgcgccaggcagctgcacaggtttcttggctttgtacgtggttttcacttcagcgtcataatggccgcc
atctttcagtttcaggcgctgtttaatttcgcctttcagcgcaccatcttccggatacatgcgttcgctagacgcctcccaacccatcgtctttttct
gcatcaccgggccatcagatggaaaattagtaccacgcagtttaactttatagatgaactcgccatcctgcagggaggagtcctgagtgacg
gtcacgacaccaccatcttcaaaattcattacgcgttcccacttgaaaccttccggaaaagacagcttcagatagtccgggatatccgctggg
tgtttaacatacgctttagaaccgtacataaattgcgggctcaggatgtcccacgcaaaaggcagcgggccgcctttagtcactttcagtttgg
cggtctgggtgccttcatacggacggccctcgccttcgccttcgatctcgaactcgtgaccgttaacagaaccctccatgtgaactttgaagc
gcatgaactctttaatgatagccatgttatcttcttctcctttacgcatggtaccctccattacgacatgtgaatttattcaaaatctgcagtgctggt
tattcattgctgattaatagcattgcttattcattgaataaagaatcgcaggaatgcaatcacaaggcttggatggggtgaacctaaaaattgtac
ataatgatcattgtagaaaaaaaacaaactccaaataaaaattcttattttattgtaatcttatgttttaattaaaaaatatatttatatattgtatgttttt
attttgcattgaacaatttgtcattcgtgttggaataacttaaaatttttaatcctagcccccgagagggctaggattacctaaaaaataaaaatcg
aagagttaaaatcttttcaacgattttgggtaactgatccaaacaatctcttttagttgcaagcttcttcatctgctagaaagaagggaaatggga
aacctcttccttgaagtgattagacttacatattcggatagtttggaccaccaccaccggccgcactagtgatatcccgcggccactcgatcct
ctacgccggacgcatcgtggccggcatcaccggcgccacaggtgcggttgctggcgcctatatcgccgacatcaccgatggggaagatc
gggctcgccacttcgggctcatgagcgcttgtttcggcgtgggtatggtggcaggccccgtggccgggggactgttgggcgccatctcctt
gcatgcaccattccttgcggcggcggtgctcaacggcctcaacctactactgggctgcttcctaatgcaggagtcgcataagggagagcgt
cgagatcccggacaccatcgaatggcgcaaaaccttcgcggtatggcatgatagcgcccggaagagagtcaattcagggtggtgaatat
gagcacagacaaagcaaatacgctgatcaaacccgaagatgtcgtgttatggattccgggtaatgtcacaattgacagcatgaatgccggtt
gggaaaacattgcaatcagagggtacgaatataccaacctcgatgtgcatattcctgccatgcgtgactacatgatcgtcaactataaaaaaa
gtgcggcggaaatgcgtagaaaaggcgatgcctcttgggatacccaagtggttaagccgggttatgtctccttgttgacctgtggtgaagatt
cccgctgggcgtggaatgaccatattgccgtcacccatgtctacatttcgcatgactccatcacctcaatggcgaataaggtgtttgattatgat
atcgcttcgatccgaatcagagacgaagtcggtgtggaagatcatgttttacctgctctgacttcacttttagaactagaattaaagcaaggtgg
tttaggtggaaacctgtatttagagagcattaaaaaccagatcgccctgcatttactccgtcagtatgccaaattagattttaaggaaggacagt
gccgttctggttttactcccctacaacgcagactgttattagaatttatcaatgaaaacatgagcattaaaattaccctcgaagatttagcgggat
tagtcaagatgagcgtgcctcatttaatgagaaaatttaaagtcgattttggtaattcccctgctgcctacatcatgaatctcagggtgcaatttg
ctaaacgtttgctcacttcaaaaaaagaaattccactgaaagtgattgccagtgaagccggtttttgcgatcagagccatatgacccgagtattt
caaaaatttttgggaaaacacccatcgaaatcagacaggaacacaccaatctcgtgtctgaaaattcagtctcctctattgtttttgagcgcaa
cgcaattaatgtaagttagctcactcattaggcaccgacgtc