Communication

Production of odd-carbon dicarboxylic acids in Escherichia coli using an engineered biotin-fatty acid biosynthetic pathway

Robert W. Haushalter, Ryan M. Phelan, Kristina M. Hoh, Cindy Su, George Wang, Edward E.K. Baidoo, and Jay D Keasling

J. Am. Chem. Soc., **Just Accepted Manuscript** • Publication Date (Web): 14 Mar 2017

Production of odd-carbon dicarboxylic acids in *Escherichia coli* using an engineered biotin-fatty acid biosynthetic pathway.

Robert W. Haushalter, Ryan M. Phelan, Kristina M. Hoh, Cindy Su, George Wang, Edward E.K. Baidoo, and Jay D. Keasling.

Joint BioEnergy Institute, U.S. Department of Energy, 5885 Hollis Street, Emeryville, California 94608, United States

Physical Bioscience Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

Corresponding author: jdkeasling@lbl.gov

ABSTRACT: Dicarboxylic acids are commodity chemicals used in the production of plastics, polyesters, nylons, fragrances, and medications. Bio-based routes to dicarboxylic acids are gaining attention due to environmental concerns about petroleum-based production of these compounds. Some industrial applications require dicarboxylic acids with specific carbon chain lengths, including odd-carbon species. Biosynthetic pathways involving cytochrome P450-catalyzed oxidation of fatty acids in yeast and bacteria have been reported, but these systems produce almost exclusively even-carbon species. Here we report a novel pathway to odd-carbon dicarboxylic acids directly from glucose in *Escherichia coli* by employing an engineered pathway combining enzymes from biotin and fatty acid synthesis. Optimization of the pathway will lead to industrial strains for the production of valuable odd-carbon diacids.

Currently, production routes for dicarboxylic acids (DCAs) (e.g, adipic,[1] pimelic[2] and suberic acid[3]) rely on petrochemical feedstocks. Production of DCAs from renewable feedstocks using engineered microbes presents an attractive and sustainable alternative to current synthetic methods. Several bio-based routes to adipic acid have been reported.[4-7] More recently, expression of a cytochrome P450 fatty acid hydroxylase with alcohol and aldehyde dehydrogenases in an engineered fatty acid-producing *E. coli* strain led to production of longer 12- and 14-carbon chain DCAs.[8] Using a combinatorial strategy Bowen *et. al.* were able to generate up to 84 mg/L of C_{12} DCA in shake flask experiments. These biosynthetic pathways involve oxidation of free fatty acids, which are predominantly even-carbon in industrial hosts like yeast and *E. coli*. As a consequence, the DCAs produced by fatty acid oxidation in these systems likewise contain even numbers of carbons. However, numerous industrial applications require odd-carbon DCAs. For example, azeleic acid (9-carbon DCA) is used as an acne medication.[9] Other examples include tridecanedioic acid (a.k.a. brassylic acid) and pentadecanedioic acid, both of which can be converted into high-value fragrance molecules.[10,11] The industrial significance of these compounds inspired us to seek a means of biologically producing exclusively odd-carbon DCAs.

We decided to investigate the use of biotin biosynthetic genes from *E. coli* because of the 7-carbon pimeloyl-ACP intermediate in the proposed pathway.[12-14] Biotin biosynthesis in *E. coli* begins with methylation of malonyl-ACP by BioC (Fig.1).[15] The vast majority of malonyl-ACP is consumed for fatty acid production under normal physiological conditions, but overexpression of *bioC* is expected to divert carbon flux toward the biotin pathway. Methylation of the malonyl group greatly decreases the polarity of the molecule, and thus allows extension of the resulting malonyl methyl ester by fatty acid synthase (FAS).[15] During biotin biosynthesis, FAS extends the malonyl methyl ester group twice to yield a pimeloyl-ACP methyl ester intermediate. At this step in biotin biosynthesis, BioH hydrolyzes the terminal methyl ester.[16] This restores the polar carboxyl group, prevents further extension by FAS, and allows processing of the pimeloyl group into biotin.[14] We hypothesized that, if BioH activity was removed, and an acyl-ACP thioesterase (i.e., *'tesA*) was overexpressed with *bioC*, FAS would extend the ACP-bound methyl ester further until it reached a length in the substrate range of the thioesterase, and would then be released as the monomethyl ester. Here we report the biological production of odd-chain DCAs directly from glucose in engineered *E. coli* by using biotin biosynthetic intermediates as novel starter units in DCA synthesis.

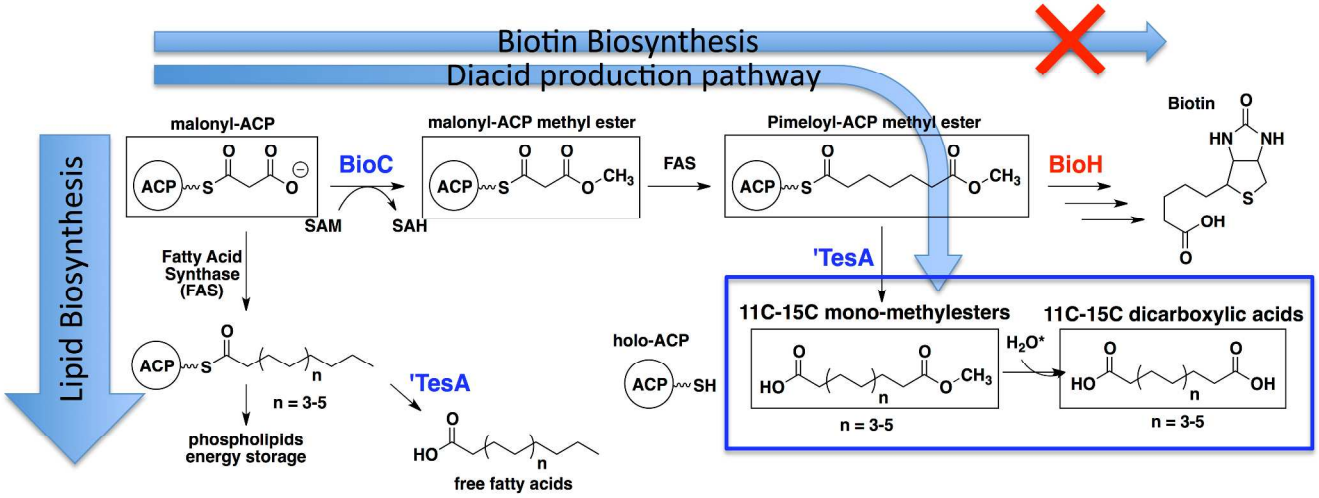


Figure 1. Engineered of the odd- carbon dicarboxylic acid production pathway. Both fatty acid and biotin biosynthesis begin with malonyl-ACP. This intermediate is either directly used by FAS, or methylated by BioC and extended by the same FAS complex to generate pimeloyl-ACP methyl ester. The methyl ester of this intermediate is normally hydrolyzed by BioH to yield a pimeloyl-ACP intermediate for downstream biotin biosynthetic steps. Overexpression of BioC increases metabolic flux down this pathway, while knockout of BioH allows extension beyond the C_7 pimeloyl intermediate. The extended ACP-bound methyl ester is eventually cleaved off of ACP by 'TesA-catalyzed thioester hydrolysis. *In vitro* data suggests that 'TesA plays a role in hydrolyzing the resulting mono-methyl esters to dicarboxylic acids.

The BioC enzyme from *B. cereus* was the first homolog we investigated. Previous biochemical studies of BioC enzymes from a variety of bacterial sources revealed that this homolog expresses well in *E. coli* and the enzyme is stable enough for purification and activity assays.[15] Overexpression of *bioC* resulted in growth inhibition in this initial study, presumably due to methylation of more malonyl-ACP than is tolerable for proper fatty acid synthesis. We hypothesized that this toxic effect could be relieved by an esterase that removes these aberrant intermediates from the ACP. The native *E. coli* enzyme TesA is normally localized to the periplasm, where it would not be able to catalyze the hydrolysis of methylated ACP-bound intermediates, but a recombinant version of the enzyme, lacking a leader peptide that directs it to the periplasm, is known to release free fatty acid intermediates from ACP in the cytosol.[17] We anticipated that co-expression of this truncated gene, *'tesA,* would help relieve this toxic effect by removing ACP-bound methyl esters and restoring ACP to its catalytically viable *holo-* state.

The gene *'tesA,* and *bioC* from *B. cereus,* were co-expressed in *E. coli* MG1655 and the DCA content of the cells were analyzed. We used LCMS analysis for our initial experiments because it is capable of discerning between mono-methyl ester and diacid products, and because it is much more sensitive than GCMS analysis. We co-expressed both genes using P_{lacUV5} in strains JBEI-3111 (*E. coli* MG1655 *ΔfadE*) and JBEI-7954 (*E. coli* MG1655 *ΔfadE ΔbioH*) to investigate the effect of BioH activity on DCA titer. The *rfp* gene was co-expressed with *'tesA* in the negative control strain. We did not observe any growth defect in strains expressing *bioC* compared to the RFP control, suggesting that *'tesA* does indeed alleviate the toxic effects of BioC reported by Lin and Cronan.[15] When *bioC* and *'tesA* were co-expressed a mixture of odd-carbon DCAs ranging from C_9 to C_{15} were produced (Figure 2). Negligible amounts of these products were observed in the negative control expressing *'tesA* and *rfp*. Brassylic acid (C_{13}) dominated the product distribution, comprising 80-90% of DCAs produced. Additionally, while produced at significantly lower levels (<1 mg/L), compounds containing 9, 11 and 15 carbons were generated (see figure S3 for raw LCMS spectral counts). We also detected a trace amount of C_{13} mono-methyl ester. We found that deleting *bioH* increased DCA titers modestly from 6.2±0.5 mg/L to 6.9±0.5 mg/L. Interestingly, even in the absence of *bioH*, most of the observed products were DCAs rather than methyl esters, indicating that hydrolysis of the terminal methyl ester is taking place spontaneously or enzymatically. To test whether 'TesA is responsible for accelerating methyl ester hydrolysis, we purified recombinant 'TesA from *E. coli* and tested its ability to

hydrolyze mono-methyl ester standards. Indeed, methyl esters incubated with 1uM 'TesA showed complete hydrolysis after 10 hours incubating at 22C (see Figure S2). To test the effect of expression level of *bioC*, we compared DCA titers when *bioC* expression was driven with P_{lacUV5} or P_{T7} . The promoter used is indicated in the plasmid name in the legend, ie. E5C= P_{lacUV5} and E7C= P_{T7} (Figure 2). Surprisingly, increasing the expression strength of the *bioC* gene with P_{T7} resulted in a decrease in DCA titers. The highest titers were observed when *'tesA* and *bioC* from *B. cereus* were both expressed using P_{lacUV5} in a *ΔbioH* strain.

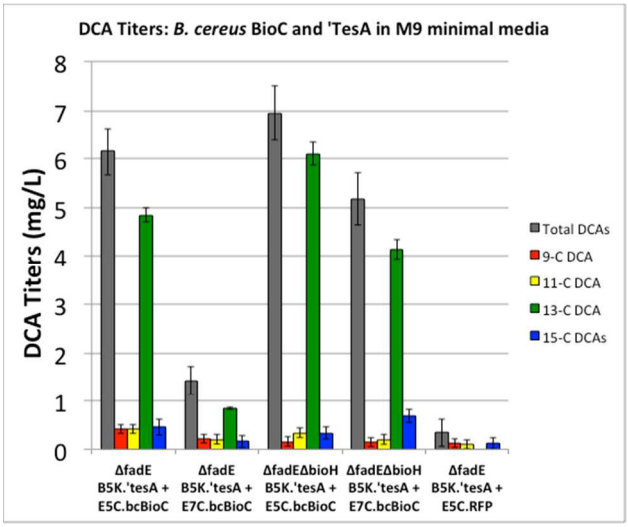


Figure 2. Dicarboxylic acid production in M9 minimal media as measured by LCMS-TOF. Expression of *bioC* from *B. cereus* was driven by P_{lacUV5} or P_{T7} in the presence or absence of *bioH*. Error bars indicate standard deviations measured from triplicate samples.

We then tested two alternative *bioC* genes from different bacterial sources. When Lin and Cronan first characterized the activity of *B. cereus* BioC, they noted that the *bioC* homologs from *P. putida* and *Kurthia* were particularly toxic to cell growth.[15] We hypothesized that this notable toxicity is an indication of the high activity of these enzymes *in vivo*. In the absence of an ACP-thioesterase, this increased activity would result in an accumulation of acyl-ACP methyl esters and have a deleterious effect on cell growth. However, in the presence of 'TesA, we anticipated that increased BioC activity would lead to a desirable increase in DCA titers. It is possible that the optimal level of BioC activity falls between the levels resulting from bcBioC production with P_{lacUV5} and $P_{T7,}$ and that expressing a more active isoform with P_{lacUV5} could help achieve this activity level. We also hypothesized that growing in defined rich media could help increase DCA titers.

The *bioC* genes from *B cereus*, *P. putida*, and *Kurthia* were each co-expressed with *'tesA* in strain JBEI-7954 (*E. coli* MG1655 *ΔfadE ΔbioH*), cultured in EZ-rich media, and the DCA titers were measured. We used GCMS analysis for these experiments because it allows simultaneous measurement of DCAs and free fatty acids. It is necessary to fully derivitize both fatty acids and DCAs to methyl esters for GCMS analysis, and thus DCA titers measured are a combination of DCAs and any mono-methyl esters in the culture (which we showed were only present in trace amounts using LCMS). We found that the *Kurthia* enzyme outperformed the *B. cereus* enzyme slightly, with DCA titers of 24.7±1.4 mg/L and 20.7±1.2 mg/L respectively. The *P. putida* enzyme was the least active of the three, producing only 2.3±0.7 mg/L DCAs.

We repeated the experiment by comparing both the fatty acid and DCA titers in *Kurthia* strain (inoculated from freezer stocks) with the RFP negative control. We found that the RFP strain produced 2.86 ± 0.79 g/L free fatty acids with no detectable DCAs. In contrast, the strain expressing *Kurthia bioC* produced only 0.6 ±0.11 g/L free fatty acids while producing 15.7 ± 3.0 mg/L DCAs. These results suggest that *bioC* expression diverts carbon flux away from fatty acid biosynthesis.

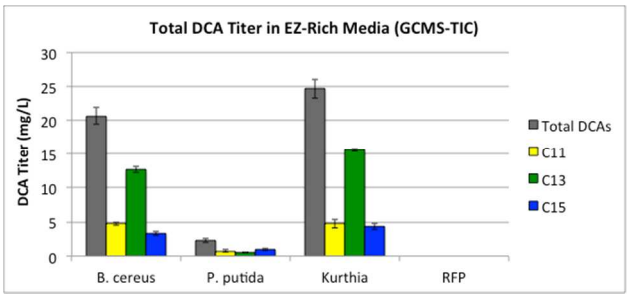


Figure 3. Dicarboxylic acid titers in EZ-rich media with *bioC* homologs from different species as measured by GCMS in TIC mode. Error bars indicate standard deviations from triplicate measurements.

We have demonstrated a novel biosynthetic pathway to odd-carbon DCAs by allowing biotin intermediates to be fully extended by the endogenous FAS biosynthetic machinery. All three BioC enzymes lead to odd-carbon DCA production, with the *Kurthia* homolog performing the best of the three. It is not surprising that C_{13} DCA is the most prevalent product when considering the substrate specificity of 'TesA. This ACP-thioesterase is more commonly used to release free fatty acids from ACP, which deregulates fatty acid synthase and leads to accumulation of free fatty acids. 'TesA similarly releases a distribution of chain lengths rather than a single species, with C_{14} and C_{16} produced at the highest levels.[18] When a C_{13} methyl ester is appended to ACP, the acyl chain has a total length of 15 atoms and is thus sterically similar to the C_{14} and C_{16}

fatty acids on which 'TesA is most active. It is therefore likely that the lengths of the DCA species produced can be tailored by choosing alternative ACP-thioesterases with alternate chain length specificities.

This work provides a novel, sustainable bio-based route to odd-carbon DCAs without even-carbon by-products, and adds to the increasing number of petrochemicals that can now be accessed from renewable feedstocks using metabolic engineering. The next step is to increase productivity in a scalable manner. A fundamental hurdle to overcome is the competition between the engineered DCA pathway and native fatty acid and biotin synthesis. The cell requires very little biotin, and so supplementing the media with small amounts of this cofactor relieves the toxicity imposed by blocking *de novo* synthesis. However, diverting enough malonyl-ACP away from fatty acid synthesis to increase DCA titers without impacting normal cell growth and function is a challenge. With this unique hybrid route to odd-chain DCAs established, we are currently investigating multiple regulation strategies to improve titers. Once optimized, it will provide the commodity chemical industry with a sustainable alternative for odd-carbon DCA production.

The Supporting Information is available free of charge on the ACS Publications website.

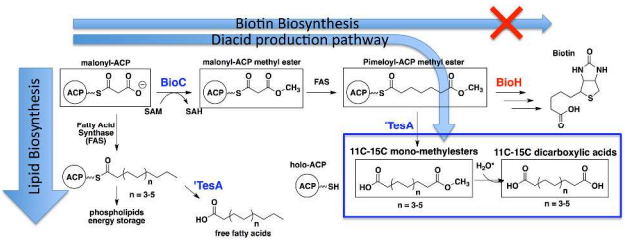
Experimental details, plasmids and strains used in this stud, DNA sequences, raw GCMS/LCMS data.

Acknowledgements

This work was part of the DOE Joint BioEnergy Institute (http:// www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy. We thank Anisa Azad, Brian Laus, and Maren Wehrs for technical assistance.

References

1. Castellan, A., Bart, J. C. J., and Cavallaro, S. *Catal Today* (1991) *9*, 301-322.
2. Warner, D. T., and Moe, O. A. *J. Am. Chem. Soc.* (1952) *74*, 371-373.
3. Cason, J., Wallcave, L., and Whiteside, C. N. *J Org Chem* (1949) *14*, 37-44.
4. Niu, W., Draths, K. M., and Frost, J. W. (2002) *Biotechnol. Prog.* (2002) *18*, 201-211.
5. Hagen, A., Poust, S., Rond, T., Fortman, J. L., Katz, L., Petzold, C. J., and Keasling, J. D. *ACS Synth. Biol.* (2016) *5*, 21-27.
6. Yu, J-L., Xia X-X., Zhong, J-J., and Qian, Z-G. *Biotechnol. Bioeng.* (2014) 111(12): 2580-2586.
7. Clomberg, J.M., Blankschien, M.D., Vick, J.E., Chou, A., Kim, S., and Gonzalez, R. *Metab Eng* (2015) 28: 202-212
8. Bowen, C. H., Bonin, J., Kogler, A., Barba-Ostria, C., and Zhang, F. *ACS Synth. Biol.* (2016) *5*, 200-206.
9. Bladon, P. T., Burke, B. M., Cunliffe, W. J., Forster, R. A., Holland, K. T., and King, K. *Br. J. Dermatol.* (1986) *114*, 493-499.
10. Ravi, S., Padmanabhan, D., and Mamdapur, V.R. *J. Indian Inst. Sci.* (2001) 81, 299-312.
11. De Cardayre, S.B., Schirmer, A.W., and Ko, M. PCT/US2015/025452 (2015)
12. Lin, S., Hanson, R. E., and Cronan, J. E. *Nat. Chem. Biol.* (2010) *6*, 682-688.
13. Lin, S., and Cronan, J. E. *Mol. Biosyst.* (2011) *7*, 1811-1821.
14. Cronan, J. E., and Lin, S. *Curr. Opin. Chem. Biol.* (2011) *15*, 407-413.
15. Lin, S., and Cronan, J. E. *J Biol Chem* (2012) *287*, 37010-37020.
16. Akatsuka, H., Kawai, E., Sakurai, N., and Omori, K. *Gene* (2003) *302*, 185-192.
17. Cho, H., and Cronan, J.E. *J. Biol. Chem.* (1995) 270(9):4216-9.
18. Haushalter, R. W., Kim, W., Chavkin, T. A., The, L., Garber, M. E., Nhan, M., Adams, P. D., Petzold, C. J., Katz, L., and Keasling, J. D. *Metab. Eng.* (2014) *26*, 111-118.


Biotin Biosynthesis
Diacid production pathway
Lipid Biosynthesis
malonyl-ACP
BioC
SAM
SAH
malonyl-ACP methyl ester
FAS
Pimeloyl-ACP methyl ester
BioH
Biotin
Fatty Acid Synthase (FAS)
n = 3-5
phospholipids energy storage
'TesA
free fatty acids
holo-ACP
'TesA
11C-15C mono-methylesters
n = 3-5
H_2O^*
11C-15C dicarboxylic acids
n = 3-5