One-pot synthesis of adipic acid from guaiacol in escherichia coli

One-Pot Synthesis of Adipic Acid from Guaiacol in *Escherichia coli*

Jack T. Suitor, Simon Varzandeh and Stephen Wallace*

Institute for Quantitative Biology, Biochemistry and Biotechnology, School of Biological Sciences, University of Edinburgh, King's Buildings, Alexander Crum Brown Road, Edinburgh, EH9 3FF, UK

Keywords: Biotransformation, biocatalysis, lignin

ABSTRACT: Adipic acid is one of the most important small molecules in the modern chemical industry. However, the damaging environmental impact of the current industrial synthesis of adipic acid has necessitated the development of greener, bio-based approaches to its manufacture. Herein we report the first one-pot synthesis of adipic acid from guaiacol, a lignin-derived feedstock, using genetically engineered whole-cells of *Escherichia coli*. The reaction is mild, efficient, requires no additional additives or reagents and produces no by-products. This study demonstrates how modern synthetic biology can be used to valorize abundant feedstocks into industrially-relevant small molecules in living cells.

Microorganisms can be programmed to perform chemical synthesis using synthetic biology. This enables the bioproduction of small molecules of industrial value directly from renewable feedstocks via fermentation. With a market value of $6.3 billion and an annual production of 2.6 million tons, adipic acid (hexane-1,6-dioic acid) is a highly important industrial chemical.[1-4] It is used, amongst other applications, as the precursor to the polymer nylon-6,6. Key to its industrial synthesis is the metal-catalyzed oxidative ring-opening of petrochemically-derived cyclohexanone/cyclohexanol ("KA-oil") using concentrated nitric acid (Figure 1A). Not only is this process highly energy-intensive, it also releases nitrous oxide (a greenhouse gas 300-times more potent than carbon dioxide) into the atmosphere. Alarmingly, in 2015 the adipic acid process was found to be responsible for 8-10% of anthropogenic N_2O emissions worldwide. As a result, research into the renewable synthesis of adipic acid via synthetic biology has received considerable attention in recent years.[1-4]

Until recently, semi-synthetic approaches to the synthesis of adipic acid have relied on the extraction and Pd-catalyzed hydrogenation of the metabolite *cis*,*cis*-muconic acid (ccMA, Figure 1B).[5-6] In addition to chemical hydrogenation, Balskus et al. recently demonstrated the reduction of ccMA using H_2 generated via microbial metabolism.[7] However, in 2017 Mahadevan *et al.* reported the discovery of a novel NADH-/flavin-dependent oxidoreductase from *Bacillus coagulans* (BcER) capable of reducing ccMA to adipic acid.[8-9] The enzyme contains an Fe-S cluster that facilitates C=C reduction via electron transfer from NADH. Expression of BcER in a ccMA-producing strain of *Saccharomyces cerevisiae* has since been used to access adipic acid directly from D-glucose (<1% yield, 2.6 mg/L).[5] The low yield in this strain is thought to stem from inactivation of the enzyme's Fe-S cluster by O_2 during cofactor maturation. Other metabolic engineering approaches have included reverse adipate degradation in *E. coli* and ω-/β-fatty acid biosynthesis in *Candida sp.* using D-glucose and palm oil, respectively.[10-11] All these approaches use central amino acid biosynthesis or 2C-metabolic siphons (e.g. acetyl-CoA) to build

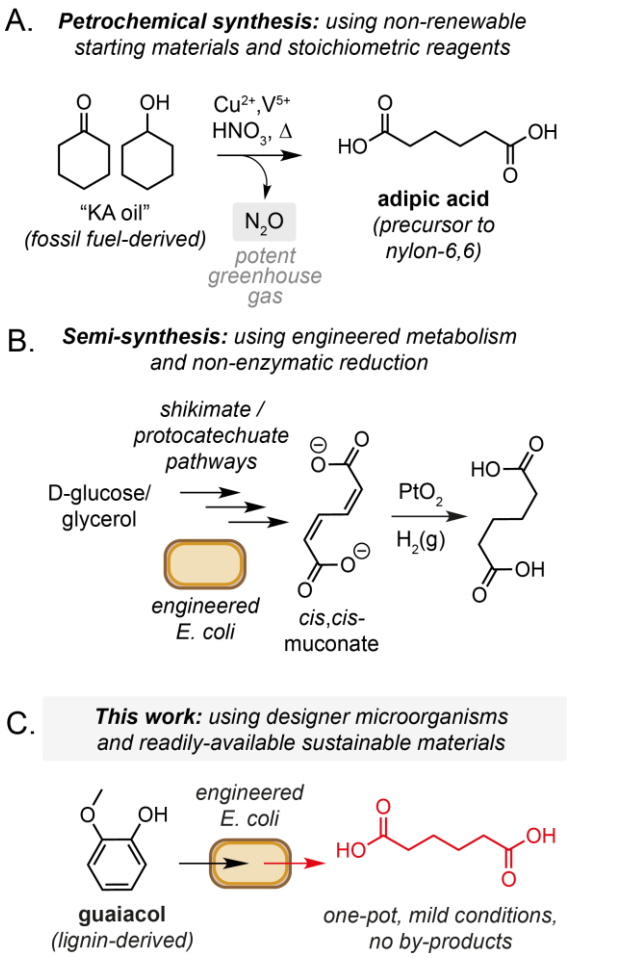

Figure 1. *Chemical and biological approaches to adipic acid. A) Industrial oxidation of KA oil. B) Semi-synthesis via metabolic engineering and chemical hydrogenation. C) Biotransformation of lignin-derived material using an engineered microorganism.*

the 6C-backbone of adipic acid. Increased production therefore comes at a metabolic cost to the organism, which in-turn inhibits its growth and subsequent product formation. This paradox often limits the yield of product available via engineered metabolism. One solution is the use of whole-cell reactions, where sustainable feedstocks can be directly converted into value-added products using heterologous biosynthetic pathways that operate independent to central metabolism. This includes the use of lignin, a naturally-occurring polymer that is considered to be one of the greatest untapped carbon sources on Earth. Valorization of lignin monomers directly into adipic acid is therefore an important challenge in chemical and synthetic biology.[2,6] Herein we report the synthesis of adipic acid from catechol (benzene-1,2-diol) and guaiacol (2-methoxyphenol) using engineered whole-cells of *E. coli* (Figure 1C). The one-pot reaction proceeds under mild conditions (37 ˚C, pH 7.2, aqueous media), requires no additives and produces no by-products. Optimizing heterologous enzyme solubility was essential to achieving rapid product formation in this strain using molecular chaperones – particularly at low cell density. Together, this study demonstrates how designer cells can be used to perform challenging chemical transformations and emphasizes the benefit(s) of engineering the background cellular environment to optimize enzymatic chemistry in whole-cell catalysis.

Our studies began by investigating the conversion of catechol to adipic acid by *E. coli* BL21(DE3). In this pathway, catechol is oxidized to ccMA by the Fe(III)- and O_2-dependent dioxygenase, CatA, from *Pseudomonas putida* followed by reduction to adipic acid by the Fe-S cluster-containing oxidoreductase, *Bc*ER, from *B. coagulans* (Figure 2A). The *catA* and *bcER* genes were co-expressed in *E. coli* BL21(DE3) using a pETDuet vector (pAA) and protein expression was confirmed via SDS-PAGE. To our delight, incubation of these cells in the presence of catechol (5 mM, Na-Pi, 3% w/v D-glucose, pH 7.2, 37 ˚C) for 24 h yielded adipic acid, which could be readily detected in lyophilized cell supernatant by ^{1}H NMR (Figure 2B). However, low mass recovery and overlapping signals in the downfield NMR region precluded the quantification of catechol or ccMA via this method. Pleasingly, analysis of the reaction by HPLC showed that adipic acid was produced in 85% yield (Figure 2C). Product formation was dependent on the presence of cells, catechol, glucose and IPTG.

To further optimize this, we first focused on the potential inactivation of BcER by O_2 during the reaction. Under microaerophilic conditions, we hypothesized that increasing the levels of BcER would mitigate any partial deactivation due to residual O_2 and increase the yield of adipic acid. To this end, we knocked-out the transcriptional repressor of Fe-S cluster biogenesis in *E. coli*, *iscR*, using a λ-Red recombinase. However, BcER expression and the yield of adipic acid remained unchanged in this strain (Tables S3 and S5). Moreover, visualization of the cells via microscopy showed a distorted cell morphology and the presence of inclusion bodies (Figure 3A and S2). Further analysis via SDS-PAGE revealed significant levels of insoluble BcER in both strains (68% and 71%, respectively, Figure 3B and S6) and therefore improper protein folding and/or non-functional aggregation in the cell interior was likely the reason for the reduced yield. To improve the solubility of the recombinant enzymes we focused on reducing the rate of protein synthesis and aiding protein folding using molecular chaperones. To begin, we tested *E. coli* Lemo21(DE3) cells. This BL21-derived strain contains the Lemo System™, a

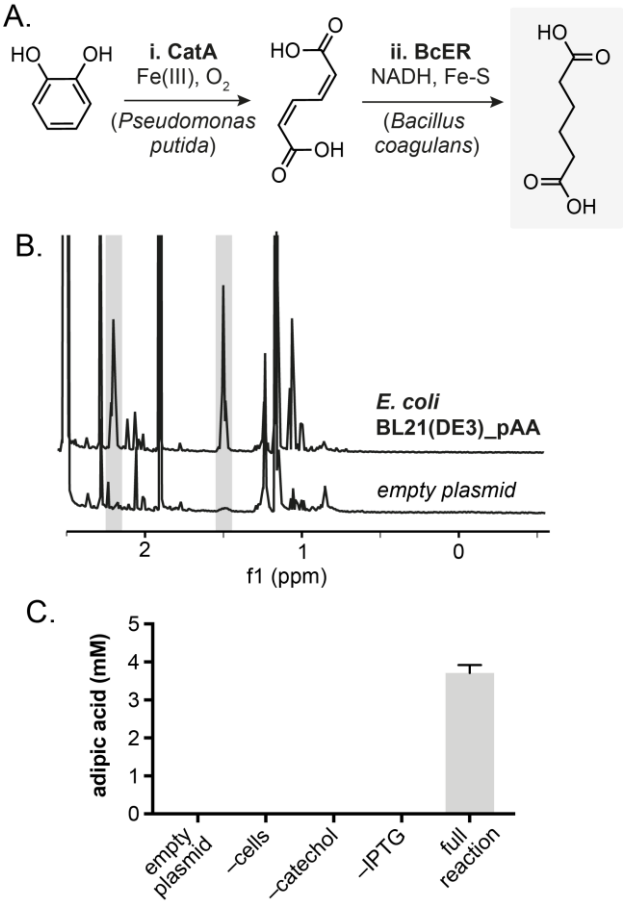

Figure 2. *Initial screen in E. coli BL21(DE3). A) The adipic acid pathway. B) Analysis of lyophilized culture supernatant by ^{1}H NMR. C) Control experiments examining the effect of the reaction components on product yield. Reactions were performed at 37 ˚C in sealed tubes under an atmosphere of air for 24 h. After inducing protein expression (0.4 mM IPTG in Terrific Broth at 18 ˚C), cells were suspended in Na-Pi buffer (pH 7.2) containing 5 mM catechol to OD_{600}=122. Product concentrations were determined by reverse-phase HPLC relative to an internal standard of caffeine. All data shown is an average of three independent experiments to one standard deviation. pAA refers to pETDuet-catA/bcER.*

pACYC184 plasmid encoding lysozyme, *lysY*, under the control of a rhamnose-inducible PrhaBAD promoter. LysY is the natural inhibitor of T7-RNA polymerase and, as such, pLemo allows tunable control of T7-mediated gene expression using L-rhamnose.[12] Therefore, we co-transformed this strain with our pAA plasmid and examined the effect of increasing rhamnose concentration on the solubility of CatA/BcER and the reaction yield. Pleasingly, the solubility of BcER was increased using 10 μM L-rhamnose. However, this had no effect on the yield of adipic acid (Table S3). Increasing L-Rha to 40 and 100 μM increased the proportion of soluble enzyme but reduced overall protein expression levels, which in-turn decreased the yield to 56% and 32%, respectively. We therefore required a strategy to promote the solubility of CatA and BcER whilst also retaining high levels of protein expression *in vivo*. To achieve this, we moved on to test the use of molecular chaperones. Three sets of chaperones were chosen and co-expressed in *E. coli* BL21(DE3)_pAA using an arabinose-inducible pACYC plasmid. These encoded the native proteins DnaK-DnaJ-GrpE (pKJE7), GroEL-GroES (pGro7) and Trigger Factor (TF, pTf16). DnaK/J-GrpE and Tf are known to prevent aggregation of the unfolded polypeptide in the cytosol, whereas GroEL-

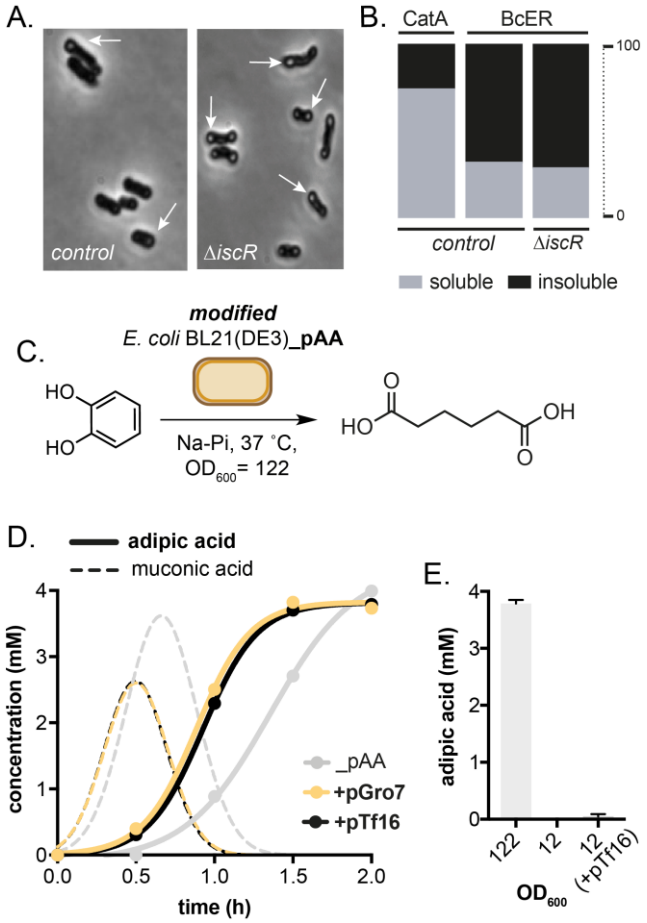

Figure 3. *Effects of an ΔiscR knock-out and examining product formation over time at high and low cell density. A) Microscopy images showing the presence of inclusion bodies (indicated by a white arrow). B) Relative protein levels in in soluble and insoluble cell fractions. C) Whole-cell transformation of catechol to adipic acid. D) Time-course analysis showing the effect of chaperone co-expression. E) Product formation at high and low cell density. Data shown is an average of three independent experiments to one standard deviation.*

GroES aids during protein folding.[13-14] In our parental strain, co-expression of GroEL-GroES and TF increased levels of soluble CatA by 30-40% and TF also increased soluble BcER levels by 75% (Table S5 and S6). In addition, GroEL-GroES or TF had no effect on the yield, whereas expression of DnaK/J-GrpE reduced the yield to 70% (Table S3).

Having increased the solubility of CatA and BcER we next moved on to examine product formation over time and at low cell density. We hypothesized that improved enzyme solubility in these modified strains would increase the rate of product formation relative to a cell containing predominantly insoluble aggregates. To test this, we compared the formation of adipic acid in the parent strain *E. coli* BL21(DE3)_pAA, *E. coli* BL21(DE3)_pAA_pGro7 and *E. coli* BL21(DE3)_pAA_pTf16 at various cell densities over time. Pleasingly, adipic acid was formed quantitatively in <2 h in all three strains (Figure 3D). This is encouraging as the production of adipic acid via engineered metabolism in *E. coli* requires 144 h via fed-batch fermentation.[5] Product formation reached 24% after 1 h in the parent strain, whereas co-expression of GroEL-GroES or Tf increased the conversion >2.5-fold after this time (66% and 62% yield, respectively, Figure 3D and S5). Increasing the solubility of recombinant CatA and BcER therefore accelerates product formation via this pathway. Reducing cell density ten-fold reduced the yield to 0% in all strains (Figure 3E and Table S3). Formation of adipic acid could not be restored by increasing the concertation of D-glucose, lowering the reaction temperature or by increasing the reaction time. However, to our surprise, re-suspending cells in M9 media at OD_{600}=12 increased the yield to 59% and co-expression of Tf or GroEL-GroES increased this further to 77% and 89% yield, respectively (Table 1 and Table S4). Chaperone co-expression therefore enhances whole-cell reactivity at low cell density in this system. In addition, cells isolated from spent reactions run at OD_{600}=12 could re-grow when inoculated into LB media whereas cells from reactions run at OD_{600}=122 could not, indicating increased metabolic activity under these conditions (Figure S9). Furthermore, this low cell density can be achieved in stirred-tank bioreactors and therefore suggests this system could be amenable to large-scale bioproduction.

Table 1. *Reaction optimization at low cell density[a,b]*

Entry	*E. coli* Bl21(DE3)	Reaction Media	OD_{600} (a.u)	Yield of **1** (%±S.D)
1	_pAA	Na-Pi	122	85 ± 5
2	_pAA	Na-Pi	12	0
3	_pAA	Na-Pi + 10% glu	12	0
4	_pAA + pTf16	Na-Pi	12	0
5	_pAA + pGro7	Na-Pi	12	0
6	_pAA	M9	12	59 ± 21[c]
7	_pAA + pTf16	M9	12	77 ± 1
8	_pAA + pGro7	M9	12	89 ± 1

[a] Reactions were performed as outlined in Figure 1. All data shown is an average of three independent experiments to one standard deviation. [b] reactions were run for 24 h. [c] n=9.

Having optimized the production of adipic acid from catechol to near-quantitative levels, we moved on to extend this pathway to encompass the use of a renewable substrate. Guaiacol is the major component of softwood lignin, a renewable natural resource of considerable interest to the field of industrial biotechnology. Recent studies have demonstrated the conversion of guaiacol to ccMA in engineered *Amycolatopsis sp.* ATCC 39116.[15-16] However, the direct production of adipic acid from guaiacol via engineered metabolism has yet to be achieved. Towards this goal, we constructed an expression vector harboring the *catA* and *bcER* genes in addition to *gcoAB* from *Amycolatopsis sp.* ATCC 39116. GcoA is a hemin-dependent cytochrome P450 and GcoB is its associated 2Fe-2S cluster-containing, NADH- and FAD-dependent reductase. We codon-optimized and assembled the four genes in a pQLinkN backbone using iterative rounds of ligation-independent cloning (Figure 4A). Pleasingly, incubation of the cells with guaiacol (5 mM, PBS, pH 7.2, 37 ˚C) for 24 h resulted in a 14% yield of adipic acid (Figure 4C and Table S7). This was accompanied by a significant amount of unreacted substrate, suggesting that the activity of CatA and BcER had been retained in this strain and that the demethylation of guaiacol was rate-limiting. This was confirmed by adding equimolar volumes of catechol and formaldehyde to the cells and observing a 98% yield of adipic acid (Table S8). In addition, negligible formaldehyde was detected in spent reactions using a Fluoral-P colorimetric assay, confirming

that *E. coli* also consumes the by-product of the demethylation reaction via an endogenous process (Figure 4D and S8). Interestingly, replacing the *Aa*GcoAB enzymes with GcoA and GcoB from *Rhodococcus rhodochrous* (pAA'-*Rr*) reduced the yield of adipic acid to 1%, despite having activity in *P. putida* EM42.[17] However, to our delight, when we expressed all four enzymes in *E. coli* BL21(DE3)_pAA' in M9 or LB media and reduced the cell density of the subsequent reaction to OD_{600}=20 the yield of adipic acid increased to 61% (Figure 4C and Table

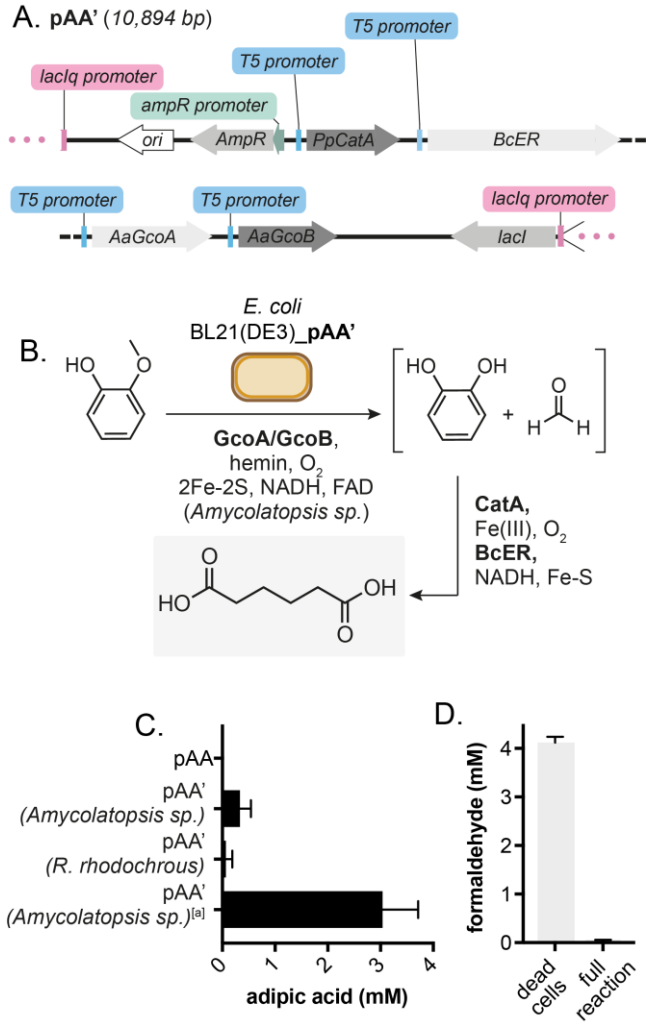

Figure 4. *One-pot synthesis of adipic acid from guaiacol. A) Map of the pAA' plasmid B) The biosynthetic pathway. C)* Product quantification by HPLC. Caffeine was used as an internal standard. *D) Formaldehyde quantification using a Fluoral-P colorimetric assay. After induced protein expression (0.4 mM IPTG, 0.76 mM iron(III) ammonium citrate, 0.76 mM 4-aminolevulinic acid in LB or M9 at 18 ˚C), cells were suspended in M9-glucose media (pH 7.2) to OD_{600}=122 before the addition of 5 mM guaiacol. pAA' refers to pQLinkN-AagcoAB/catA/bcER. [a] OD_{600}=20. All data shown is an average of three independent experiments to one standard deviation.*

S7). To best of our knowledge, this is the first synthesis of adipic acid from guaiacol in an engineered microorganism.[5-6,18] Although the precise reason(s) for this latter increase in yield are currently unclear, the accumulation of ccMA and guaiacol in low- and high-OD reactions, respectively, indicates that a fine balance exists between GcoAB and BcER expression and activity in this pathway (Table S7).

To conclude, we have demonstrated the first synthesis of adipic acid from guaiacol in the bacterium *Escherichia coli*. This pathway uses the remarkable chemistry of biogenic Fe2+/Fe3+ in heme, non-heme *and* Fe-S cluster-containing metalloenzymes to achieve the overall linearization of an aromatic ring. Co-expression of molecular chaperones was used to accelerate product formation and the addition of M9 media was found to be essential to maintain high yields at low cell density. Overall, this study demonstrates how synthetic biology can be used to program living cells as biological reagents for sustainable chemical synthesis.

ASSOCIATED CONTENT

Supporting Information

(Word Style "TE_Supporting_Information"). (Word Style "TE_Supporting_Information"). A listing of the contents of each file supplied as Supporting Information should be included. For instructions on what should be included in the Supporting Information as well as how to prepare this material for publication, refer to the journal's Instructions for Authors.

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

AUTHOR INFORMATION

Corresponding Author

* Institute for Quantitative Biology, Biochemistry and Biotechnology, Roger Land Building, Alexander Crum Brown Road, The King's Buildings, Edinburgh, EH9 3FF. stephen.wallace@ed.ac.uk

Author Contributions

The manuscript was written through contributions of all authors. / All authors have given approval to the final version of the manuscript.

Funding Sources

J.T.S acknowledges a PhD scholarship from the Carnegie Trust for the Universities of Scotland (PHD007724) and S.W acknowledges a Future Leaders Fellowship from UKRI (MR/S033882/1).

ACKNOWLEDGMENT

The authors would like to thank R. C. Brewster, E. Darmon and K. Nicholson for technical assistance.

REFERENCES

[1] Bart, J. C. J.; Cavallaro, S. Transiting from Adipic Acid to Bioadipic Acid. Part II. Biosynthetic Pathways. *Ind. Eng. Chem. Res.* **2015**, *54*, 567-576. DOI: 10.1021/ie502074d

[2] Kruyer N.S.; Peralta-Yahya, P. Metabolic engineering strategies to bio-adipic acid production. *Curr. Opin. Biotechnol.* **2017**, 45, 136-143. DOI: 10.1016/j.copbio.2017.03.006

[3] Deng, Y.; Ma, L.; Mao, Y. Biological production of adipic acid from renewable substrates: Current and future methods. *Biochem. Eng. J.* **2016**, 105, 16-26. DOI: 10.1016/j.bej.2015.08.015

[4] Skoog E.; Shin J. H.; Saez-Jimenez V.; Mapeli, V.; Olsson, L. Biobased adipic acid - The challenge of developing the production host. *Biotechnol. Adv.* **2018**, 36, 2248-2263. DOI: 10.1016/j.biotechadv.2018.10.012

[5] Vardon, D. R.; Franden, M. A.; Johnson, C. W.; Karp, E. M.; Guarnieri, M. T.; Linger, J. G.; Salm, M. J.; Strathmann, T. J.; Beckham, G. T. Adipic acid production from lignin. *Energy Environ. Sci.* **2015**, *8*, 617-628. 10.1039/C4EE03230F

[6] van Duuren, J. B. J. H.; de Wild, P. J.; Starck, S.; Bradtmöller, C.; Selzer, M.; Mehlmann, K.; Schneider, R.; Kohlstedt, M.; Poblete-Castro, I.; Stolzenberger, J.; Barton, N.; Fritz, M.; Scholl, S.; Venus, J.; Wittman, C. Limited life cycle and cost assessment for the bioconversion of lignin-derived aromatics into adipic acid. *Biotechnol. Bioeng.* **2020**; 1-13. DOI: 10.1002/bit.27299

[7] Sirasani, G.; Tong, L.; Balskus, E. P. A Biocompatible Alkene Hydrogenation Merges Organic Synthesis with Microbial Metabolism. *Angew. Chem. Int. Ed.* **2014**, *53*, 7785-7788. DOI: 10.1002/anie.201403148

[8] Joo, J. C.; Khusnutdinova, A. N.; Flick, R.; Kim, T.; Bornscheuer, U.T.; Yakunin, A.F.; Mahadevan, R. Alkene hydrogenation activity of enoate reductases for an environmentally benign biosynthesis of adipic acid. *Chem. Sci.* **2017**, *8,* 1406-1413. DOI: 10.1039/c6sc02842j

[9] Raj, K.; Partow, S.; Correia, K.; Khusnutdinova, A.N.; Yakunin, A.F.; Mahadevan, R. Biocatalytic production of adipic acid from glucose using engineered *Saccharomyces cerevisiae*. *Metab. Eng. Commun.* **2018**, *6*, 28-32. DOI: 10.1016/j.meteno.2018.02.001

[10] Zhao, M.; Huang, D.; Zhang, X.; Koffas, M. A. G.; Zhou, J.; Deng, Y. Metabolic engineering of Escherichia coli for producing adipic acid through the reverse adipate-degradation pathway. *Metab. Eng.* **2018**, *47*, 254-262. DOI: 10.1016/j.ymben.2018.04.002

[11] Beardslee, T.; Picataggio, S. Bio-based adipic acid from renewable oils. *Lipid Tech.* **2012**, *24*, 223-225. DOI: 10.1002/lite.201200230

[12] Wagner, S.; Klepsch, M. M.; Schlegel, S.; Appel, A.; Draheim, R.; Tarry,M.; Högbom, M.; van Wijk, K. J.; Slotboom, D.J.; Persson, J. O.; de Gier, J. Tuning Escherichia coli for membrane protein overexpression *PNAS* **2008**, *38*, 14371-14376. DOI: 10.1073/pnas.0804090105

[13] de Marco, A.; Vigh, L.; Diamant, S.; Goloubinoff, P. Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol–overexpressed molecular chaperones *Cell Stress Chaperones* **2005**, *10*, 329. DOI: 10.1379/CSC-139R.1

[14] Nishihara, L.; Kanemori, M.; Yanagi, H.; Yura, T. Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in Escherichia coli *Appl. Environ. Microbiol.* **2000** *66*, 884-889. DOI: 10.1128/aem.66.3.884-889.2000:

[15] Machovina, M. M.; Mallinson, S. J. B.; Knott, B. C.; Meyers, A. W.; Gacria-Borràs, M.; Bu, L.; Gado, J. E.; Oliver, A.; Schmidt, G. P.; Hinchen, D. J.;, M. F.; Johnson, C. W.; Neidle, E. L.; Payne, C. M.; Houk, K. N.; Beckham, G. T.; McGeehan, J. E.; Dubois, J. L. Enabling microbial syringol conversion through structure-guided protein engineering *PNAS* **2019**, *116*, 13970-13976. DOI: 10.1073/pnas.1820001116

[16] Mallinson, S. J. B.; Machovina, M. M.; Silveira, R. L.; Gacria-Borràs, M.; Gallup, N.; Johnson, C. W.; Allen, M. D.; Skaf, M. S.; Crowley, M. F.; Neidle, E. L.; Houk, K. N.; Beckham, G. T.; Dubois, J. L.; McGeehan, J. E. A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion *Nat. Commun.* **2018**, *9*, 2487. DOI: 10.1038/s41467-018-04878-2

[17] García-Hidalgo, J.; Ravi, K.; Kuré, L.; Lidén, G.; Gorwa-Grauslund, M. Identification of the two-component guaiacol demethylase system from *Rhodococcus rhodochrous* and expression in *Pseudomonas putida* EM42 for guaiacol assimilation. *AMB Express* **2019** *9*, 34. DOI: 10.1186/s13568-019-0759-8.

[18] Niu, W.; Willett, H.; Mueller, J.; He, X.; Kramer, L.; Ma, b.; Guo, J. Direct biosynthesis of adipic acid from lignin-derived aromatics using engineered Pseudomonas putida KT2440 *Metab. Eng.* **2020**, *59*, 151-161. DOI: 10.1016/j.ymben.2020.02.006

Table of Contents artwork: