2004

Rational Pathway Engineering of Type I Fatty Acid Synthase Allows the Biosynthesis of Triacetic Acid Lactone from D-Glucose in Vivo

Wenjuan Zha, *University of Illinois at Urbana-Champaign*
Zengyi Shao, *University of Illinois at Urbana-Champaign*
John W. Frost, *Michigan State University*
Huimin Zhao, *University of Illinois at Urbana-Champaign*

J|A|C|S
COMMUNICATIONS

Published on Web 03/23/2004

Rational Pathway Engineering of Type I Fatty Acid Synthase Allows the Biosynthesis of Triacetic Acid Lactone from D-Glucose in Vivo

Wenjuan Zha,[†] Zengyi Shao,[†] John W. Frost,[‡] and Huimin Zhao*,[†]

Department of Chemical and Biomolecular Engineering, University of Illinois, 600 South Mathews Avenue, Urbana, Illinois 61801, and Department of Chemistry, Michigan State University, 528 Chemistry Building, East Lansing, Michigan 48824

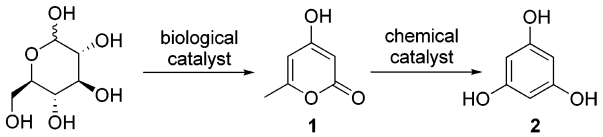
Received December 15, 2003; E-mail: zhao5@uiuc.edu

Metabolic pathway engineering has emerged as a powerful tool to synthesize structurally diverse and complex chemicals through genetic manipulation of multistep catalytic systems involved in cell metabolism. Because it allows the use of annually renewable feedstocks such as corn, metabolic pathway engineering is often considered to be an enabling technology for green chemistry and the biobased economy.[1] Here, we report the rational design of a fatty acid biosynthetic pathway, *Brevibacterium ammoniagenes* fatty acid synthase B (FAS-B), that allows the microbial synthesis of triacetic acid lactone (TAL) **1** from D-glucose. TAL can be chemically converted to phloroglucinol **2**,[2] which is a core structure for the synthesis of various high value bioactive compounds[3] and energetic compounds such as 1,3,5-triamino-2,4,6-trinitrobenzene (TATB).[4] The synthesis of phloroglucinol from D-glucose using this combined biological and chemical synthesis (Scheme 1) may offer significant advantages over the current phloroglucinol manufacture,[3] including environmental friendliness and reduction in the cost of phloroglucinol. More importantly, it represents a novel strategy for the benzene-free synthesis of aromatic chemicals.[2]

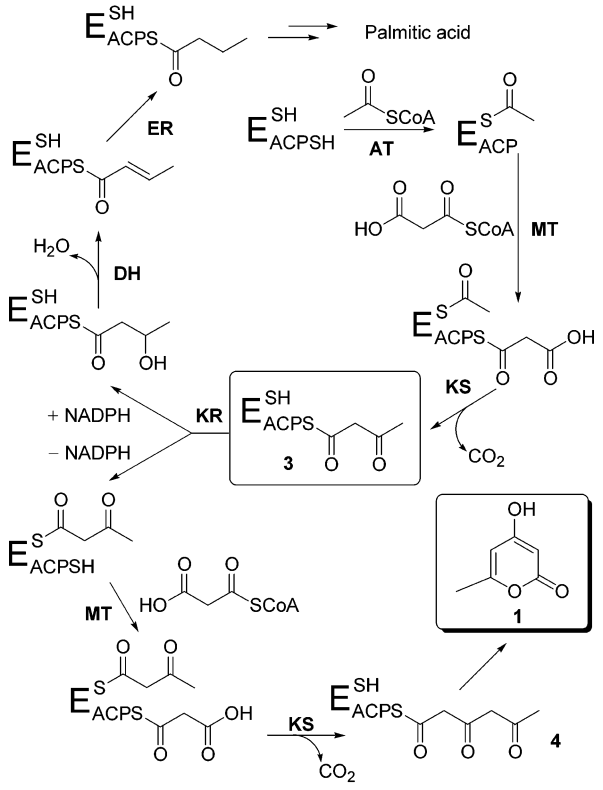
As a type I fatty acid synthase, FAS-B is a single, large multifunctional enzyme that contains multiple functional domains to catalyze all of the reactions necessary for fatty acid synthesis.[5] The exact domain structure of FAS-B is unclear, but it is believed to include a ketoacyl-synthase (KS) domain, an enoyl reductase (ER) domain, a ketoacyl-reductase (KR) domain, an acyl transferase (AT) domain, and an acyl carrier protein (ACP) domain.[5] Previously, it was shown that TAL is the exclusive product synthesized in vitro by FAS-B in the absence of NADPH.[6] Consequently, a reaction mechanism (Scheme 2) was proposed. Fatty acid synthesis proceeds from acetyl-CoA and malonyl-CoA to acetoacetyl-ACP **3**. Under conditions of NADPH limitation, acetoacetyl-ACP **3**, instead of being reduced by the KR domain (nonfunctional due to the absence of NADPH) as in the physiological condition, undergoes the KS domain catalyzed condensation with malonyl-CoA to give 3,5-diketohexanoyl-ACP **4**. Subsequent enolization and intramolecular cyclization of **4** produces TAL. Therefore, we hypothesized that inactivating the KR domain of FAS-B that is heterologously expressed in a microorganism would lead to TAL formation in vivo.

To test this hypothesis, we sought to eliminate the enzyme activity of the FAS-B KR domain by mutating its key residue(s) involved in catalysis. Unfortunately, neither the location of the FAS-B KR domain nor its catalytic residues were known. Thus, to address this problem, we used various publicly available bioinformatics tools on the Biology Workbench.[7] We first tried to identify the linker regions that define the exact FAS-B KR domain by sequence alignment. It was assumed that because the multifunctional proteins likely result from gene fusion or duplication, the linker regions connecting functional domains tend to be less conserved in sequence and structure than the functional domains. Thus, a large fragment of 800 amino acids bracketing the previously identified NADPH-binding motif was selected as a probe for BLASTP search on the San Diego Supercomputer Center nonredundant protein database. Multiple sequence alignment by CLUSTALW[7] of the resulting homologous sequences narrowed the FAS-B KR domain to between residue 2054 and residue 2319. We then tried to refine the primary structure of the KR domain by predicting the secondary structure of FAS-B using PELE.[7] It was assumed that the first and the last residues of the ketoreductase domain should not be in the

Scheme 1 An Environmentally Benign Synthetic Route to Phloroglucinol



Scheme 2. Proposed Reaction Mechanism of FAS-B (AT, Acyl Transferase; DH, Dehydratase; ER, Enoyl Reductase, KR, Ketoacyl Reductase; KS, Ketoacyl Synthase; MT, Malonyl Transferase)



[†] University of Illinois.
[‡] Michigan State University.

10.1021/ja0317271 CCC: $27.50

middle of a defined secondary structure such as an α-helix or β-sheet. The final location of the ketoreductase domain was predicted to be from residue 2051 to residue 2319. Furthermore, although this putative FAS-B KR domain has low sequence homology (21.6%, determined by ALIGN[7]) with the ketoreductase (FabG) enzyme from *E. coli*, a member of the short-chain dehydrogenase/reductase (SDR) superfamily,[8] it shares several common features with the SDR superfamily including a $(\alpha/\beta)_6$ structure with the signature Rossman fold motif and a highly conserved catalytic $Y^{2226}XXXK$ motif.[9] SDRs catalyze a broad range of NAD(P)/NAD(P)H-dependent reduction and dehydrogenation reactions in which the catalytic tyrosine and lysine residues are involved in catalysis through hydrogen bonding and proton transfer.[8−10] Thus, we hypothesized that Y2226 is an important catalytic residue and replacement of this residue with a catalytically inert phenylalanine (F) may significantly reduce or potentially eliminate the KR activity. A similar mutation was made to eliminate the activity of the KR domain in a modular polyketide synthase, 6-deoxyerythronolide B synthase (DEBS), which led to the formation of triketide lactone products.[11]

The Y2226F FAS-B mutant was created by overlap extension PCR based site-directed mutagenesis[12] in plasmid pCH4.267A.[6] The mutation was confirmed by DNA sequencing. The resultant Y2226F FAS-B mutant and the wild-type FAS-B were expressed in *E. coli* strain JM109 together with the *B. ammoniagenes* phosphopantetheine transferase (PPT1) gene situated immediately downstream of the FAS-B gene. PPT1 is required for conversion of the nonfunctional apo-FAS-B to the functional holo-form.[15] Upon verifying protein expression in *E. coli* by SDS-PAGE, the FAS-B enzymes were partially purified and assayed for ketoreductase activity.[16] No ketoreductase activity of the Y2226F mutant was detected.

In vivo TAL production was analyzed by HPLC on a C8 reverse phase column with culture supernatants. Inconsistent with the fact that the partially purified Y2226F FAS-B mutant synthesized TAL in vitro (data not shown), no TAL was detected in the supernatant from JM109 harboring this mutant. However, 6 mg/L TAL was detected by HPLC and GC-MS after concentrating the supernatant of the Y2226F mutant by 1000-fold using ethyl acetate extraction. This extremely low maximal titer of TAL (6 μg/L) might result from the poor expression of FAS-B in *E. coli*. Thus, we attempted to use yeast as the new host to synthesize TAL. The FAS-B and PPT1 genes were subcloned into plasmid pMR228 and pKOS12-128a to replace the corresponding 6-MSAS gene and Sfp gene, respectively, via yeast in vivo homologous recombination[13] for coexpression in *S. cerevisiae* strain INV*Sc*1.[14] Yeast cells harboring the Y2226F FAS-B mutant produced a compound that migrated on HPLC at the same retention time as authentic TAL (retention time = 12.1 min), while no such compound was detected with yeast cells harboring the wild-type FAS-B.[14] This compound was further purified from the culture supernatant of the Y2226F mutant[17] and identified spectroscopically[18] as TAL. The time-course of TAL production for the Y2226F mutant was determined (Figure 1). The amount of TAL produced reached a plateau after ~60 h of growth, with a maximum titer of 52 mg/L.

Previous designs for benzene-free synthesis of aromatics have recruited the shikimate pathway[19] and inositol biosynthesis.[20] TAL has previously been synthesized using 6-methylsalicylic acid synthase, which is a polyketide synthase.[21] The FAS-B variant described in this report now establishes fatty acid biosynthesis as an alternative route for benzene-free synthesis of aromatics. Of equal importance, the identification of the KR domains of FAS-B and other type I fatty acid synthases as members of the SDR protein superfamily suggests that the same strategy could be used to engineer other type I fatty acid synthases for the production of TAL.

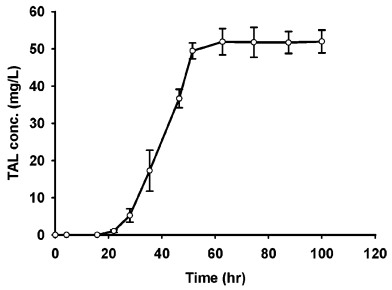


Figure 1. Time course of TAL production in yeast strain INV*Sc*1.

Acknowledgment. We thank Kosan Bioscience for providing plasmids pMR228 and pKOS12-128a, and Dr. Eckhart Schweizer for providing plasmid pGM44 containing the FAS-B and PPT1 genes. This work was supported by the Office of Naval Research.

Supporting Information Available: Sequence alignment, HPLC, NMR, and details of experimental procedures (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Schoemaker, H. E.; Mink, D.; Wubbolts, M. G. *Science* **2003**, *299*, 1694−1697. (b) de Boer, A. L.; Schmidt-Dannert, C. *Curr. Opin. Chem. Biol.* **2003**, *7*, 273−278. (c) Chotani, G.; Dodge, T.; Hsu, A.; Kumar, M.; LaDuca, R.; Trimbur, D.; Weyler, W.; Sanford, K. *Biochim. Biophys. Acta* **2000**, *1543*, 434−455.
(2) Hansen, C. A.; Frost, J. W. *J. Am. Chem. Soc.* **2002**, *124*, 5926−5927.
(3) Leston, G. In *Kirk-Othmer Encyclopedia of Chemical Technology*; Kroschwitz, J. I., Howe-Grant, M., Eds.; Wiley: New York, 1996; Vol. 19, pp 778−812.
(4) Bellamy, A. J.; Golding, P.; Ward, S. J. UK Patent Application 2 335 713, 2001.
(5) Meurer, G.; Biermann, G.; Schutz, A.; Harth, S.; Schweizer, E. *Mol. Gen. Genet.* **1992**, *232*, 106−116.
(6) Hansen, C. A. Ph.D. Thesis, Michigan State University, East Lansing, MI, 2002.
(7) Biology Workbench: http://workbench.sdsc.edu. Programs including BLASTP, CLUSTALW, PELE, and ALIGN were used in this study.
(8) Price, A. C.; Zhang, Y. M.; Rock, C. O.; White, S. W. *Biochemistry* **2001**, *40*, 12772−12781.
(9) Jornvall, H.; Persson, B.; Krook, M.; Atrian, S.; Gonzalez-Duarte, R.; Jeffery, J.; Ghosh, D. *Biochemistry* **1995**, *34*, 6003−6013.
(10) Oppermann, U. C.; Filling, C.; Berndt, K. D.; Persson, B.; Benach, J.; Ladenstein, R.; Jornvall, H. *Biochemistry* **1997**, *36*, 34−40.
(11) Reid, R.; Paigentini, M.; Rodriguez, E.; Ashley, G.; Viswanathan, N.; Carney, J.; Santi, D.; Hutchinson, C. R.; McDaniel, R. *Biochemistry* **2003**, *42*, 72−79.
(12) Horton, R. M.; Cai, Z. L.; Ho, S. N.; Pease, L. R. *Biotechniques* **1990**, *8*, 528−535.
(13) Prado, F.; Aguilera, A. *Curr. Genet.* **1994**, *25*, 180−183.
(14) See the Supporting Information for details.
(15) Stuible, H. P.; Meier, S.; Schweizer, E. *Eur. J. Biochem.* **1997**, *248*, 481−487.
(16) Stuible, H. P.; Meurer, G.; Schweizer, E. *Eur. J. Biochem.* **1997**, *247*, 268−273.
(17) See the Supporting Information for details.
(18) TAL: ^{1}H NMR (DMSO-d_6, 500 MHz) δ (ppm) 5.94 (m, 1H), 5.19 (d, J = 2.1 Hz, 1H), 2.14 (br s, 3H); ^{13}C NMR (DMSO-d_6, 500 MHz) δ (ppm) 170.9, 164.1, 163.3, 100.4, 88.1, 19.4; high-resolution EI-MS, calcd for TAL (M^+) 126.0317, found 126.0316.
(19) Ran, N.; Knop, D. R.; Draths, K. M.; Frost, J. W. *J. Am. Chem. Soc.* **2001**, *123*, 10927−10934.
(20) Hansen, C. A.; Dean, A. B.; Draths, K. M.; Frost, J. W. *J. Am. Chem. Soc.* **1999**, *121*, 3799−3800.
(21) Richardson, M. T.; Pohl, N. L.; Kealey, J. T.; Khosla, C. *Met. Eng.* **1999**, *1*, 180−187.

JA0317271