Biosynthetic Pathway

DOI: 10.1002/anie.201302406

LovG: The Thioesterase Required for Dihydromonacolin L Release and Lovastatin Nonaketide Synthase Turnover in Lovastatin Biosynthesis**

*Wei Xu, Yit-Heng Chooi, Jin W. Choi, Shuang Li, John C. Vederas, Nancy A. Da Silva, and Yi Tang**

Lovastatin, a polyketide produced by the fungus *Aspergillus terreus*,[1] is one of the most important natural products discovered to date. Both lovastatin and the semisynthetic derivative simvastatin are widely prescribed hypercholesterolemia drugs because of their inhibitory activities towards *3S*-hydroxy-3-methyl-glutaryl-CoA reductase (HMGR).[2] The fermentative production of lovastatin has therefore been one of highest grossing processes involving a natural product. We recently developed a biocatalytic process for converting an intermediate of the lovastatin biosynthetic pathway, monacolin J acid (**3**) into simvastatin acid (**5**) in a single enzymatic step.[3] Therefore, having a fast-growing heterologous host, such as *Saccharomyces cerevisiae*, that can produce **3** directly without the need to purify lovastatin followed by side chain hydrolysis, would be an attractive method. To do so, identification of the complete biosynthetic pathway leading to **3** is required.

The biosynthetic gene cluster for lovastatin acid (**1**) in *A. terreus* was first identified in a pioneering work by Kennedy et al.[4] Through genetic[4–5] and biochemical[6] characterizations, it is now known that two highly reducing iterative type I polyketide synthases (HR-PKSs) play central roles in the biosynthesis of **1**. The lovastatin nonaketide synthase (LNKS or LovB), together with the dissociated enoylreductase LovC, are responsible for the programmed assembly of dihydromonacolin L acid (**2**).[4] **2** is then modified by the cytochrome P450 monooxygenase LovA through hydroxylation and dehydration to afford monacolin L acid (**4**). The same enzyme also catalyzes hydroxylation at C-8 of **4** to produce **3**.[6c] Lastly, the α-methylbutyryl side chain is synthesized by the lovastatin diketide synthase (LDKS or LovF) and is transferred to the C-8 hydroxy group of **3** by the acyl transferase LovD to yield **1** (Figure 1).[6a] It is the function of LovD that was exploited for the enzymatic synthesis of simvastatin.[7]

Notwithstanding these insights into the *lov* pathway, one unresolved biochemical step is the release of **2** from LovB that allows for multiple turnovers by this enzyme. In our previous in vitro reconstitution study, we showed that purified LovB and LovC were sufficient to assemble **2** tethered to LovB from malonyl-CoA, but were not able to release the product.[6b] To test whether this is also the case in *S. cerevisiae,* LovB and LovC were coexpressed in *S. cerevisiae* BJ5464-NpgA, which is a vacuolar protease-deficient yeast strain with the *A. nidulans* phosphopantethienyl transferase *npgA*.[8] Following extraction of the three-day yeast culture, no trace of **2** was found by selective ion monitoring (Figure 3 A; Supporting Information, Figure S1 B), confirming the lack of a releasing function with LovB and LovC alone. Serendipitously, in our in vitro study, we identified non-native thioesterase (TE) domains from unrelated fungal PKSs that can release **2**.[6b] Co-expression of fungal TE domains, such as the PKS13 TE,[9] Hpm3 TE,[10] and others, with LovB and LovC resulted in detectable levels of **2** from BJ5464-NpgA, albeit with very low titers (less than 400 $\mu g L^{-1}$, Figure S2). Given the known expression levels of LovB (approximately 4 $mg L^{-1}$),[6b] this low titer corresponds to an average turnover of less than 90 for LovB. These studies strongly indicate that 1) the low titer of **2** from *S. cerevisiae* must be overcome by increasing the turnover rate of LovB; and 2) there must exist a natural TE that partners with LovB for the release of **2.**

To identify the TE that is involved in the biosynthesis of **2**, we reexamined the *A. terreus lov* gene cluster and searched for a likely candidate. The only known TE-like enzyme in the gene cluster, LovD, is highly specific towards LovF and does not function with LovB. A thorough bioinformatic analysis of genes of unassigned function suggested that a gene (ATEG_09962) located between *lovB* and *lovC* (Figure 1) warranted further scrutiny. This gene was initially assigned as *orf5*[4] and later as *lovG*,[11] with an annotation of the protein product as either a hypothetical protein or an oxidoreductase.

[*] W. Xu, Dr. Y.-H. Chooi, Prof. S. Li, Prof. Y. Tang
Department of Chemical and Biomolecular Engineering,
Department of Chemistry and Biochemistry,
University of California, Los Angeles
Los Angeles, CA 90095 (USA)
E-mail: yitang@ucla.edu
Prof. S. Li
School of Bioscience and Bioengineering,
South China University of Technology
Guangzhou 510006 (PR China)
J. W. Choi, Prof. N. A. Da Silva
Department of Chemical Engineering and Material Science,
University of California, Irvine
Irvine, CA 92697 (USA)
Prof. J. C. Vederas
Department of Chemistry, University of Alberta,
Edmonton, Alberta T6G 2G2 (Canada)

[**] This work was supported by the US National Institutes of Health (1R01GM085128 and 1R01GM092217) to Y.T. Natural Sciences & Engineering Research Council of Canada (NSERC) and Canada Research Chair in Bioorganic & Medicinal Chemistry to J.V. We thank Prof. D. K. Ro for providing the expression plasmid of LovA.
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201302406.

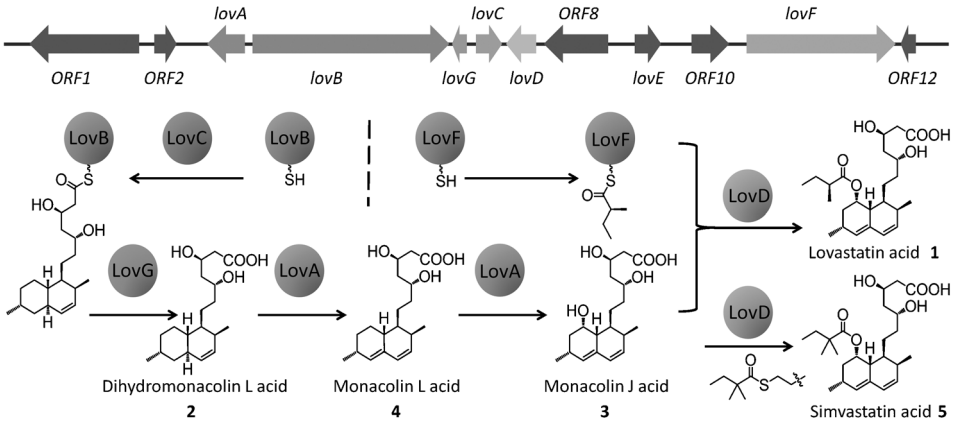

Figure 1. Lovastatin biosynthetic gene cluster and pathway.

Conserved domain analysis of LovG indicates that LovG in fact belongs to the esterase–lipase family of serine hydrolases (Figure S3). Notably, close homologues of LovG are also found in related biosynthetic pathways, including MlcF in the compactin pathway in *Penicillium citrinum*[12] (58% identity, 71% similarity) and MokD in the monacolin K (lovastatin) pathway in *Monascus pilosus*[13] (67% identity, 76% similarity). As is the case for *lovG*, both *mlcF* and *mokD* are located between the genes encoding a nonaketide PKS and the trans-acting enoylreductase. RT-PCR analysis showed that *mlcF* and *mokD* are moderately co-transcribed with the other biosynthetic genes in their corresponding clusters, which suggests that these enzymes should be involved in the biosynthetic pathways.[12–13] Therefore, we hypothesized that LovG is the missing link that is required for LovB turnover and release of **2**.

We first determined the role of *lovG* in the biosynthesis of **1** in *A. terreus*. Genetic disruption of *lovG* was performed using double-crossover recombination with the zeocin-resistant marker, followed by identification of desired *ΔlovG* mutants using diagnostic PCR (Figure S4). Metabolite analysis of the *ΔlovG* mutants showed dramatically reduced levels of **1** (less than 5%) compared to the wild-type strain (Figure 2). No intermediates, such as **2**, **4**, or **3**, accumulated in the culture, indicating that LovG is involved in the early part of the pathway. The significant attenuation of the production of **1** indicates that *lovG* is intimately involved in the *lov* pathway, but not absolutely essential for the production of **1**. This is not unexpected given our previous findings: an endogenous TE in *A. terreus* could be recruited for the hydrolysis of **2**, albeit at a much lower efficiency. In a recent study, the region encoding *lovG* (between *lovB* and *lovC*) was replaced by sequences encoding the nonribosomal peptide synthetase (NRPS) module from the CheA PKS-NRPS to yield LovB fused with an NRPS module.[14] The authors of this study observed a similar decrease in the level of **1**, and attributed the drop in yield to the reduced efficiency of the chimeric LovB.[14] From the results shown in Figure 2, we believe the decreased titer of **1** in that study could also be from the unintended deletion of *lovG*.

To further confirm the hydrolytic role of LovG, we probed the direct involvement of this enzyme in turnover of **2** by in vitro and in vivo reconstitution assays. First, LovG was expressed and purified as an *N*-terminal His-tagged protein from *E. coli* BL21(DE3) cells. Chaperone proteins GroES/EL, DnaK/J, and GrpE were co-expressed[15] to assist with the folding of LovG and to improve the solubility (Figure S5). By incubating equimolar amounts (10 μM) of purified LovB, LovC, and LovG with malonyl-CoA and the necessary cofactors (NADPH, SAM) for 12 hours

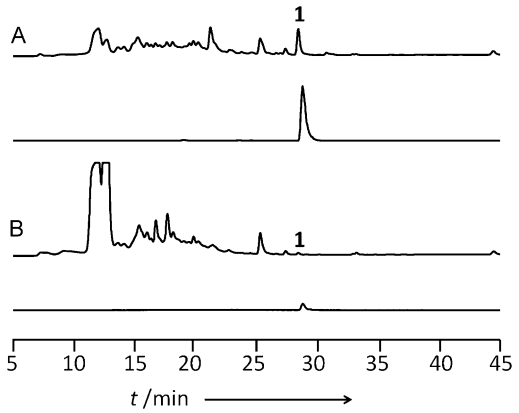

Figure 2. Metabolic profiles of fermentation extracts from cultures of different *A. terreus* strains. A) HPLC chromatograms at 238 nm and the extracted ion chromatograms of observed masses *m/z* $[M+Na]^+$ 445 corresponding to **1** in wild-type *A. terreus*. B) HPLC chromatograms at 238 nm and the extracted ion chromatograms of observed masses *m/z* $[M+Na]^+$ 445 corresponding to **1** in a *ΔlovG A. terreus* mutant. The traces are drawn to the same scale.

at 25 °C, we observed production of **2** without the need to add either base (1M KOH) or a heterologous TE (Figure 3B). With LovG included, the turnover rate of **2** was estimated to be approximately 0.12 h^{-1} (Figure S6). The slow, albeit observable, turnover may be due to a large fraction of the recombinant LovB being inactive in the in vitro assay, as previously noted.[6b] Furthermore, *lovG* was cloned under the *S. cerevisiae ADH2* promoter and coexpressed with *lovB* and *lovC*. This resulted in the production of **2** from *S. cerevisiae* BJ5464-NpgA at an elevated titer of approximately 35 mg L^{-1} after 50 hours of growth (Figure 3A; Figure S7). Collectively, these results confirmed that LovG is the natural partner for release of LovB during lovastatin biosynthesis in *A. terreus*.

The successful production of **2** following LovG coexpression prompted us to examine the ability of the yeast host to produce more advanced intermediates, such as **3**. The 2μ vector encoding both *lovA* and the endogenous *A. terreus* cytochrome P450 oxidoreductase (CPR) from *A. terreus*[6c] was introduced into BJ5464-NpgA and coexpressed with LovB, LovC, and LovG. In the absence of expression of LovA and CPR, which are controlled by the divergent *GAL1-*

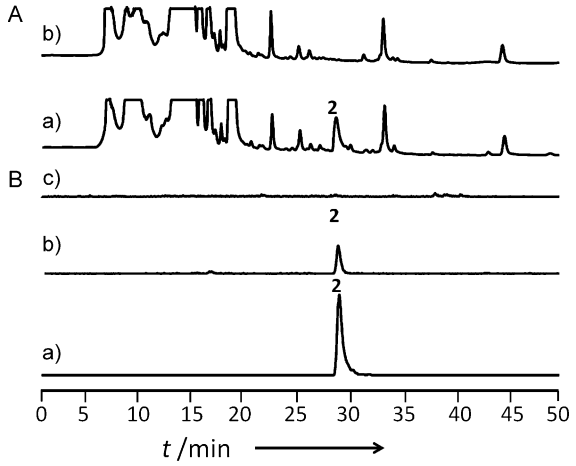

Figure 3. The release of **2** from LovB using LovG. A) HPLC chromatogram at 200 nm of fermentation extracts from three-day-old cultures of BJ5464-NpgA expressing a) LovB, LovC, and LovG; b) LovB and LovC. B) Extracted ion chromatograms of m/z $[M-H]^-$ 323 corresponding to **2** from in vitro experiments containing a) LovB, LovC, and LovG; b) LovB and LovC treated with 1 M KOH; c) LovB and LovC with no 1 M KOH treatment. All enzymes were added at a final concentration of 10 μM. The reactions were performed at 25 °C for 12 h.

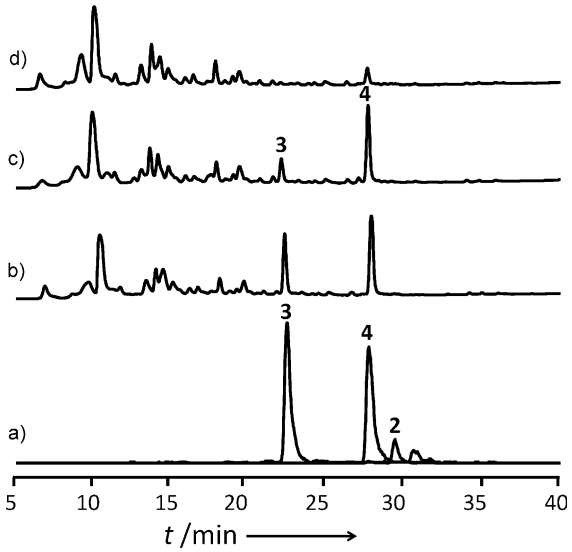

Figure 4. Production of **3** from *S. cerevisiae* BJ5464-NpgA co-expressing LovB, LovC, LovG, LovA, and *A. terreus* cytochrome P450 oxidoreductase (CPR). a) Extracted ion chromatograms showing detection of **2**, **3**, and **4** in culture 48 h after induction of LovA and CPR expression. m/z $[M-H]^-$ for **2**: 323, **3:** 337 and **4:** 321. b–d) HPLC chromatograms at 238 nm of fermentation extracts. b) 48 h after induction of LovA and CPR expression; c) 24 h after induction of LovA and CPR expression; d) before induction of LovA and CPR expression.

GAL10 promoter, we were only able to observe the accumulation of **2**. When galactose was added to the culture medium at day two, nearly all of **2** was oxidized to **4** and **3.** After 48 hours, both **3** and **4** can be detected from the yeast culture at approximately 20 mg L^{-1} each (Figure 4). It is important to note that this initial demonstration of the biosynthesis of **3** has not been optimized. All five of the *lov* enzymes were expressed episomally from three separate 2μ vectors. With additional metabolic engineering approaches, we expect the conversion from **4** to **3**, as well as the titer of **3**, could be significantly improved.

LovG joins a growing list of fungal TEs that are involved in product release through either hydrolysis or acyl transfer to an on-pathway intermediate (such as LovD[6a]). All of the TEs associated with HR-PKSs characterized to date are stand-alone enzymes, in contrast to the fused TE domains in mammalian fatty acid synthases (FAS)[16] or some of the nonreducing PKSs (NR-PKS), in which the fused TEs can act as Claisen-like cyclases (TE/CLC).[17] Recently, the TE/CLC from an NR-PKS was also shown to possess editing functions during the PKS function and hydrolyze stalled products.[18] Given that LovB is known to be highly accurate in the assembly of **2** as its only product, both in vitro and here, in vivo, we hypothesized that LovG may have a proofreading function during the iterative LovB functions to remove aberrant tailored products. This function could ensure the timely discharge of incorrect, stalled intermediates and free LovB for more efficient turnover of **2**.

To assay the possible editing function of LovG, we performed in vitro reconstitution experiments in the absence of LovC, of which the earliest function in the pathway is to reduce the α-β enoyl intermediate at the tetraketide stage. We previously observed small amounts of discharged products from LovB, when LovC was excluded, in the forms of methylated, conjugated α-pyrones such as **6** and **7** (Figure 5A) resulting from continued chain elongation and spontaneous esterification; and ketones (not visible at the scale drawn in Figure 5A) through hydrolysis of β-keto thioesters and decarboxylation.[4,6b] Upon addition of LovG at

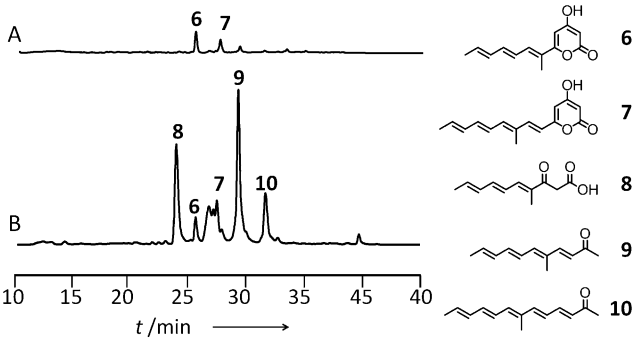

Figure 5. The release of aberrant products from LovB by LovG. HPLC chromatograms were extracted at 330 nm. A) In vitro assay with LovB; B) in vitro assay with LovB and LovG.

an equimolar concentration as LovB, the in vitro assay mixture turned significantly more yellow and the extract contained high amounts of discharged products **8**–**10**, in addition to **6** and **7** (Figure 5B). The compounds **8**–**10** were not stable, which precluded precise structural determination by NMR spectroscopy. However, by comparison of the UV absorption spectra and mass increase upon using [2-^{13}C] malonate (Figures S8–S12), as well as considering previously deciphered programming rules of LovB,[6b] we were able to

deduce the structures of these compounds as shown in Figure 5. **8** is a hydrolyzed β-keto acid, while **9** and **10** are ketones resulting from the decarboxylation of longer polyene β-keto acids. Interestingly, while the levels of **8**–**10** are high in the LovG-containing assay, the relative amounts of pyrones **6** and **7** remained nearly the same as in the assay with LovB alone, indicating that LovG can readily hydrolyze the incorrectly tailored, β-keto thioesters from LovB to give **8**–**10**. In contrast, the spontaneous hydrolysis of β-keto intermediates is much slower without LovG, resulting in LovB having to discharge the products through an additional round of chain elongation, which can then be cyclically released as **6** and **7** (Figure S13). A recent study showed that NR-PKS TEs can also promote the formation of pyrones.[19] Therefore, LovG may also assist in the formation of **6** and **7**, albeit at a much lower level as is evident in Figure 5.

In conclusion, we have identified LovG as a multifunctional esterase from the lovastatin gene cluster. LovG is not only involved in the release of the correct product **2** from LovB, but is also shown to play a role in the clearance of aberrant intermediates from LovB. Construction of the LovG-containing pathway capable of de novo synthesis of **3** also opens up new metabolic engineering opportunities for statin production from yeast.

Received: March 22, 2013
Published online: May 7, 2013

Keywords: biosynthesis · lovastatin · polyketides · product release · thioesterases

[1] A. W. Alberts, J. Chen, G. Kuron, V. Hunt, J. Huff, C. Hoffman, J. Rothrock, M. Lopez, H. Joshua, E. Harris, A. Patchett, R. Monaghan, S. Currie, E. Stapley, G. Albers-Schonberg, O. Hensens, J. Hirshfield, K. Hoogsteen, J. Liesch, J. Springer, *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 3957 – 3961.
[2] E. S. Istvan, J. Deisenhofer, *Science* **2001**, *292*, 1160 – 1164.
[3] X. K. Xie, Y. Tang, *Appl. Environ. Microbiol.* **2007**, *73*, 2054 – 2060.
[4] J. Kennedy, K. Auclair, S. G. Kendrew, C. Park, J. C. Vederas, C. R. Hutchinson, *Science* **1999**, *284*, 1368 – 1372.
[5] L. Hendrickson, C. R. Davis, C. Roach, D. K. Nguyen, T. Aldrich, P. C. McAda, C. D. Reeves, *Chem. Biol.* **1999**, *6*, 429 – 439.
[6] a) X. Xie, M. J. Meehan, W. Xu, P. C. Dorrestein, Y. Tang, *J. Am. Chem. Soc.* **2009**, *131*, 8388 – 8389; b) S. M. Ma, J. W. Li, J. W. Choi, H. Zhou, K. K. Lee, V. A. Moorthie, X. Xie, J. T. Kealey, N. A. Da Silva, J. C. Vederas, Y. Tang, *Science* **2009**, *326*, 589 – 592; c) J. Barriuso, D. T. Nguyen, J. W. Li, J. N. Roberts, G. MacNevin, J. L. Chaytor, S. L. Marcus, J. C. Vederas, D. K. Ro, *J. Am. Chem. Soc.* **2011**, *133*, 8078 – 8081.
[7] X. Xie, K. Watanabe, W. A. Wojcicki, C. C. Wang, Y. Tang, *Chem. Biol.* **2006**, *13*, 1161 – 1169.
[8] K. K. Lee, N. A. Da Silva, J. T. Kealey, *Anal. Biochem.* **2009**, *394*, 75 – 80.
[9] M. Wang, H. Zhou, M. Wirz, Y. Tang, C. N. Boddy, *Biochemistry* **2009**, *48*, 6288 – 6290.
[10] H. Zhou, K. Qiao, Z. Gao, M. J. Meehan, J. W. Li, X. Zhao, P. C. Dorrestein, J. C. Vederas, Y. Tang, *J. Am. Chem. Soc.* **2010**, *132*, 4530 – 4531.
[11] K. Auclair, J. Kennedy, C. R. Hutchinson, J. C. Vederas, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1527 – 1531.
[12] Y. Abe, T. Suzuki, C. Ono, K. Iwamoto, M. Hosobuchi, H. Yoshikawa, *Mol. Genet. Genomics* **2002**, *267*, 636 – 646.
[13] Y. P. Chen, G. F. Yuan, S. Y. Hsieh, Y. S. Lin, W. Y. Wang, L. L. Liaw, C. P. Tseng, *J. Agric. Food Chem.* **2010**, *58*, 287 – 293.
[14] D. Boettger, H. Bergmann, B. Kuehn, E. Shelest, C. Hertweck, *ChemBioChem* **2012**, *13*, 2363 – 2373.
[15] K. Nishihara, M. Kanemori, M. Kitagawa, H. Yanagi, T. Yura, *Appl. Environ. Microbiol.* **1998**, *64*, 1694 – 1699.
[16] T. Maier, M. Leibundgut, N. Ban, *Science* **2008**, *321*, 1315 – 1322.
[17] I. Fujii, A. Watanabe, U. Sankawa, Y. Ebizuka, *Chem. Biol.* **2001**, *8*, 189 – 197.
[18] A. L. Vagstad, S. B. Bumpus, K. Belecki, N. L. Kelleher, C. A. Townsend, *J. Am. Chem. Soc.* **2012**, *134*, 6865 – 6877.
[19] A. G. Newman, A. L. Vagstad, K. Belecki, J. R. Scheerer, C. A. Townsend, *Chem. Commun.* **2012**, *48*, 11772 – 11774.

