Published in final edited form as:

Chembiochem. 2014 November 3; 15(16): 2393–2401. doi:10.1002/cbic.201402275.

Construction of a Chimeric Biosynthetic Pathway for the *de novo* Biosynthesis of Rosmarinic Acid in *Escherichia coli*

Sarah E. Bloch^[a] and Claudia Schmidt-Dannert^{[a],*}

^[a]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, 1479 Gortner Avenue, St. Paul, MN 55108, USA

Abstract

Hydroxycinnamic acid esters (HCEs) are widely-distributed phenylpropanoid-derived plant natural products. Rosmarinic acid (RA **7**), the most well-known HCE, shows promise as a treatment for cancer and neurological disorders. In contrast to extraction from plant material or plant cell culture, a microbial production platform for HCEs could provide a sustainable, controlled means of production. Through the overexpression of a six-enzyme chimeric bacterial and plant pathway, we show the *de novo* biosynthesis of RA **7** and the related HCE isorinic acid (IA **8**) in *E. coli*. Probing the pathway through precursor supplementation showed several potential pathway bottlenecks. We show HCE biosynthesis using three plant RAS orthologs exhibiting different levels of HCE biosynthesis, but the same ratio of IA **8** to RA **7** produced. This work serves as a proof of concept for a microbial production platform for HCEs using a modular biosynthetic approach to access diverse natural and non-natural HCEs.

Keywords

rosmarinic acid; hydroxycinnamic acid esters; metabolic engineering; biosynthesis; phenylpropanoid

Introduction

Hydroxycinnamic acid esters (HCEs) are a diverse class of phenolic plant natural products consisting of a hydroxycinnamoyl moiety, typically caffeic acid **6**, conjugated with another aromatic molecule, such as phenyllactate, shikimate, or quinate.^[1–2] HCEs have received attention recently as potential nutraceutical supplements due to their anti-oxidant and anti-inflammatory properties. The most well-known HCE, rosmarinic acid (RA **7**), is the ester of caffeic acid **6** and 3,4-dihydroxyphenyllactic acid (3,4-DHPL **3**), and shows promise for a broad range of nutraceutical and pharmaceutical applications (Scheme 1).^[1–2] Of particular interest is its demonstrated anti-carcinogenic and anti-tumorigenic activities and utility in cancer prevention and treatment in cell culture^[3] and animal models.^[4–7] Additionally, RA

*Address correspondence to: Prof. Claudia Schmidt-Dannert, Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, 1479 Gortner Avenue, St. Paul, MN 55108, USA Tel. 612 625-5782; Fax. 612 625-5780; schmi232@umn.edu.

Supporting information for this article is given *via* a link at the end of the document

7 exhibits neuroprotective [8–11] and cognition-enhancing effects [12], and has shown promise for use in the prevention and treatment of diverse neurological disorders including Alzheimer's, [13] Parkinson's disease, [14] depression [15] and post-traumatic stress disorder. [16]

Currently, **RA 7** and other HCEs are produced primarily by extraction from plant tissue, which is a costly endeavor limited by the availability of plant material and the low concentration of HCEs in plant tissues. There are scant reports on the total [17–18] and chemoenzymatic [19] synthesis of **RA 7** and some derivatives; [20–21] however, these methods are inefficient for economical commercial production, requiring the use of harsh chemicals and giving low yields due to a chiral center in **RA 7**. Cell culturing methods could provide a lower-cost means of mass-production, and many groups have reported high production titers for **RA 7** from plant cell cultures. [22–27] However, none of these cell culturing methods have yet been applied on an industrial scale, perhaps due to the lack of understanding of regulatory networks that cause a decline in **RA 7** biosynthesis in long-term plant cell culture conditions. [2] By contrast, the development of a microbial production platform for **RA 7** and other HCEs could allow for a more controlled and stable means of producing these compounds. Additionally, the use of a microbial host could allow for a modular production system and, through combinatorial biochemistry, could lead to the biosynthesis of novel HCEs with valuable biological activities.

In plants, HCE biosynthesis relies on the convergence of two pathways – the synthesis of the hydroxycinnamoyl-CoA ester donor substrate and the conjugate acceptor substrate, which are subsequently ester-linked by a hydroxycinnamoyl transferase (Scheme 1). In the **RA 7** biosynthetic pathway, which has been completely elucidated in plants, the hydroxycinnamoyl transferase rosmarinic acid synthase (RAS) conjugates 4-hydroxyphenyllactic acid (4-HPL **2**) and *p*-coumaroyl-CoA, and the resulting ester is then hydroxylated at the 3 and 3' positions by cytochrome P450s to give **RA 7**. [1, 28] The native plant biosynthetic pathway for **RA 7** is impractical for use in a microbial host, requiring the co-overexpression of several plant enzymes including microsomal cytochrome P450s. Because they require localization to the endoplasmic reticulum for optimal activity and an associated NADPH-cytochrome P450 reductase for electron transfer, plant microsomal P450 expression is notoriously problematic in microorganisms commonly used in biotechnology, such as *E. coli* and *S. cerevisiae*. [29] Fortunately, however, many hydroxycinnamoyl transferases, including the three RASs previously characterized, show broad substrate flexibility and readily accept 3,4-hydroxylated acids and CoA esters as substrates. [30–39] This could allow for the design of a biosynthetic pathway in which the 3-hydroxylation of the donor and acceptor substrate precursors could be carried out by a non-cytochrome P450 bacterial monooxygenase prior to ester formation.

Recent studies have reported the bioconversion of phenolic precursors to hydroxycinnamoyl amides, [40–41] shikimates and quinates [42] and the biosynthesis of hydroxycinnamoyl anthranilates from glucose [43] in *E. coli*. In this study, we designed and constructed a novel, modular biosynthetic pathway for the biosynthesis of **RA 7** from endogenous tyrosine and 4-hydroxyphenylpyruvate (4-HPP **1**) (Scheme 1). Through the overexpression of this engineered pathway, we observe *de novo* biosynthesis of **RA 7** and isorinic acid (**IA 8**),

another hydroxycinnamoyl-hydroxyphenyllactate conjugate, in *E. coli* using three orthologous RAS enzymes. This work serves as a proof of concept for a microbial production platform for RA **7** and other HCEs. Advances in metabolic engineering of the shikimate pathway in *E. coli* have shown the utility of this approach for the mass production of aromatic compounds from simple feedstocks.^[44–49] This pathway, used in a tyrosine-overproducing strain and with future optimization of pathway enzyme activity and expression, could lead to an inexpensive production platform for HCEs.

Results and Discussion

Design of a chimeric RA biosynthetic pathway for *E. coli*

In order to design an effective system for microbial production of RA **7** and minimize the number of plant enzymes required, we designed a novel chimeric bacterial-plant pathway in which the meta-hydroxylation of the donor and acceptor substrates takes place before condensation by RAS (Scheme 1). In this pathway, caffeic acid **6** and 3,4-DHPL **3** serve as the donor and acceptor substrates, respectively, rather than *p*-coumaric acid **5** and 4-HPL **2** as in the plant pathway. All pathway genes were initially cloned into pUCBB, one of our group's custom set of BioBrick™-compatible vectors, for subsequent assembly of multi-gene plasmids to create the biosynthetic pathway modules described later.^[50] Monocistronic “bricks” containing gene expression cassettes allow for “stacking” using the standard BioBrick™ restriction sites^[51] that can be reused repeatedly.

The biosynthesis of caffeic acid **6** has been demonstrated previously in *E. coli*;^[43, 47, 52–54] however, at the time of this work, biosynthesis of 3,4-DHPL **3** had not been shown in *E. coli*. To achieve 3,4-DHPL **3** biosynthesis, we first chose a 2-hydroxyacid dehydrogenase from *Lactobacillus delbrueckii* subsp. *bulgaricus*, HdhA, that was originally characterized as a D-isomer specific 2-hydroxyisocaproate dehydrogenase and was shown to have some activity against phenylpyruvic acid.^[55] We hypothesized that the overexpression of the gene *hdhA* would lead to the reduction of endogenous 4-HPP **1** to 4-HPL **2** as a first step in 3,4-DHPL **3** biosynthesis (Scheme 1). Therefore, we amplified the gene encoding HdhA from *L. delbrueckii* genomic DNA and created the plasmid pUCBB-hdhA, which contains a constitutive expression cassette for *hdhA* (Figure 1). *E. coli* BW27784 transformed with pUCBB-hdhA were cultured in minimal media and samples were taken over time, cleared of cells, and extracted with ethyl acetate for HPLC analysis. We observed production of 4-HPL **2** to 39 ± 6 mg/L without the addition of any precursors to the culture media (Figure 2), suggesting that HdhA reduces endogenous 4-HPP **1** to 4-HPL **2**. No 4-HPL **2** biosynthesis was observed in transformants of the empty pUCBB plasmid.

The next challenge in 3,4-DHPL **3** biosynthesis was the meta-hydroxylation of 4-HPL **2** to 3,4-DHPL **3** (Scheme 1). Two flavin-dependent monooxygenases, Sam5 from *Saccharothrix espanaensis* and the *E. coli* 4-hydroxyphenylacetate 3-hydroxylase complex, HpaBC, have been shown to carry out the meta-hydroxylation of *p*-coumaric acid **5** and facilitate biosynthesis of caffeic acid **6** in *E. coli* when co-overexpressed with a tyrosine ammonia lyase (TAL).^[47, 52–54] Due to the similarity in structure between *p*-coumaric acid **5** and 4-HPL **2**, we hypothesized that one of these two monooxygenases may be able to accept 4-

HPL **2** as a substrate and produce 3,4-DHPL **3**. We amplified the gene encoding Sam5 from *S. espanaensis* genomic DNA, amplified the genes encoding HpaB and HpaC individually from *E. coli* BL21 genomic DNA, and created the constitutive overexpression plasmids pUCBB-sam5 and pUCBB-hpaBC (Figure 1). *E. coli* BW27784 transformed with pUCBB-hpaBC were cultured in minimal media in the presence of 1 mM *p*-coumaric acid **5** or 1 mM 4-HPL **2** (Scheme 1); culture samples were taken over time, cleared of cells, extracted and analyzed *via* HPLC. The overexpression of *sam5* led to the complete conversion of *p*-coumaric acid **5** to caffeic acid **6** within 24 h, as had been shown previously; however, no conversion of 4-HPL **2** to 3,4-DHPL **3** could be observed. By contrast, the overexpression of *hpaBC* led to the complete conversion of *p*-coumaric acid **5** to caffeic acid **6** within 24 h and the detection of 3,4-DHPL **3** after 24 h (Figure 3). Additionally, because we chose to employ the same enzyme to carry out two simultaneous conversions in the RA biosynthetic pathway, we incubated transformants of pUCBB-hpaBC with both 1 mM 4-HPL **2** and 1 mM *p*-coumaric acid **5** to determine if the presence of one substrate would inhibit the activity of the enzyme against the other. We did not observe any significant effect of feeding both substrates simultaneously on enzyme activity with either substrate (Figure S1). We therefore decided to use HpaBC for the meta-hydroxylation of 4-HPL **2** as well as for caffeic acid **6** biosynthesis. After confirming the activities of HdhA and HpaBC with 4-HPP **1** and 4-HPL **2**, respectively, as key RA **7** precursors, we cloned genes encoding enzymes with known activities for the remaining steps of the chimeric pathway as shown in Scheme 1. In our chimeric pathway design, acceptor substrate biosynthesis combines the activities of HdhA and HpaBC for the production of 3,4-DHPL **3**. For donor substrate biosynthesis, we chose a highly active TAL from *Rhodobacter sphaeroides* (RsTAL) that we have previously shown to convert tyrosine into *p*-coumaric acid **5**.^[56] RsTAL and HpaBC together catalyze the conversion of endogenous L-tyrosine **4** to caffeic acid **6**. For CoA activation of the donor substrate caffeic acid **6**, we selected 4-coumarate-CoA ligase 2 from *Arabidopsis thaliana* (At4CL2) due to the reported preference of this isoenzyme for caffeic acid **6** as a substrate^[57] and cloned the gene from *A. thaliana* cDNA. Finally, to carry out the condensation of 3,4-DHPL **3** and caffeic acid **6** to form RA **7**, we chose the first characterized RAS from *Coleus blumei* (CbRAS)^[37, 58] and cloned the gene from cDNA synthesized from RNA extracted from *C. blumei* plant tissue.

De novo biosynthesis of RA and IA using a modular biosynthetic pathway

One of the biggest advantages of using a heterologous system for the assembly and expression of biosynthetic pathways is the relative ease with which pathway design and precursor input can be manipulated once a working pathway has been established. The biosynthesis of HCEs lends itself well to combinatorial biochemistry, because the hydroxycinnamoyl transferases that carry out condensation of the donor and acceptor substrates have been shown to have relatively broad substrate specificity and catalyze the condensation of a range of esters and amides^[36–38]. The construction of a modular HCE biosynthetic pathway would therefore allow for the interchange of modules for the biosynthesis of different acceptor and donor substrates and the condensation of these to synthesize a range of known or novel HCEs. To this end, we assembled and tested three

modular plasmid systems for the production of RA **7** (Figure 1) using our previously designed, compatible BioBrick™ plasmids for pathway assembly in *E. coli*.^[50]

Because the conversion of 4-HPL **2** to 3,4-DHPL **3** by HpaBC appeared to be relatively slow (Figure 3), we assembled the genes for 3,4-DHPL **3** biosynthesis onto the high copy number vector pUCBB. We first attempted a two-plasmid pathway setup that consisted of 3,4-DHPL **3** (acceptor substrate) biosynthetic genes and *cbRAS* being expressed on the plasmid pUCBB-ABCR and the donor substrate biosynthetic genes on the lower-copy plasmids pCDFBB-TBC4 or pACBB-TBC4 (Figure 1). Because we have found co-expression of 4CL enzymes with TALs to be problematic, perhaps due to the depletion of CoA pools within the cell, *at4CL2* was expressed under the control of an arabinose inducible promoter (*P_{araBAD}*) while all other genes were expressed constitutively from a modified *P_{lac'}* promoter.^[50] When *E. coli* BW27784 was co-transformed with the acceptor substrate and *cbRAS* plasmid pUCBB-ABCR and the donor substrate plasmids pCDFBB-TBC4 or pACBB-TBC4 and expression of *at4CL2* was induced 24 h after inoculation, no RA **7** could be detected in the media.

Next, we assembled a three-plasmid modular system, consisting of the acceptor substrate module pUCBB-ABC, containing *hdhA* and *hpaBC*; the donor substrate module pCDFBB-TBC, containing *rsTAL* and *hpaBC*; and the hydroxycinnamoyl transfer module, pACBB-4cbR, containing *at4CL2* and *cbRAS* (Figure 1). Because 3,4-DHPL **3** production by HpaBC was first seen 24 h after the addition of 4-HPL **2** (Figure 3), in this three-module system we expressed both *at4CL2* and *cbRAS* under the control of the arabinose inducible promoter *P_{araBAD}*, while all other genes were expressed constitutively from *P_{lac'}*. Transformants harboring these three plasmids were grown in minimal media, and *at4CL2* and *cbRAS* expression was induced 24 h after inoculation. This led to the detection of RA **7** in the culture media 48 h after induction and up to $0.33 \pm 0.05 \mu\text{M}$ RA **7** 72 h after the induction of *at4CL2* and *cbRAS* expression (Figure 4). *E. coli* cells harboring pUCBB-ABC, pCDFBB-TBC and empty pACBB or all three empty plasmid backbones were tested as negative controls; no RA **7** biosynthesis was observed in these cultures. We were concerned that replicating three plasmids may put an unnecessary metabolic burden on the *E. coli* cells and that expressing *hpaBC* at high levels from both pUCBB and pCDFBB may overwhelm the host's protein folding machinery and lead to aggregation of insoluble recombinant proteins in the cell. Therefore, we attempted one non-modular plasmid, pUCBB-TABC4R, with *cbRAS* and *at4CL2* under the control of the arabinose promoter and *rsTAL*, *hdhA* and *hpaBC* expressed constitutively (Figure 1). When transformants of this plasmid were grown in minimal media and *at4CL2* and *cbRAS* expression was induced 24 h after inoculation, no production of RA **7** was detected.

Extracts of culture media of cells expressing the three-plasmid pathway also contained an HPLC peak with a different retention time from RA **7** but a nearly identical UV spectrum (Figure S2). Analysis *via* LCMS revealed this peak to be IA **8**, the ester of 4-HPL **2** and caffeic acid **6**, based on mass shift of the fragment corresponding to 3,4-DHPL **3** from 197 to 181 m/z, and the mass fragment corresponding to caffeic acid **6** (161 m/z) remaining the same (Figure 4). The biosynthesis of IA **8** alongside RA **7** can be explained by the

previously reported slight preference of CbRAS to bind 4-HPL **2** over 3,4-DHPL **3** [37] and the apparently slow turnover of 4-HPL **2** to 3,4-DHPL **3** by HpaBC (Figure 3).

Feeding cultures with pathway precursors shows potential bottlenecks and a loss of acceptor substrates

Because of the relatively low concentration of HCEs produced and the detection of both RA **7** and IA **8** in the media, we fed the *E. coli* cells expressing the engineered RA **7** pathway modules selected precursors and intermediates and measured the impact on HCE production in order to gain insight into potential pathway bottlenecks. Cultures of *E. coli* BW27784 harboring all three RA **7** pathway modules (pUCBB-ABC, pCDFBB-TBC and pACBB-4cbR) were supplemented with either 1 mM L-tyrosine **4**, 1 mM *p*-coumaric acid **5**, 1 mM caffeic acid **6**, or 1 mM caffeic acid **6** and 1 mM 4-HPL **2** 24 h after inoculation and at the same time of induction of *at4CL2* and *cbRAS* expression, and the levels of HCEs and pathway intermediates were measured *via* HPLC 48 h after induction (Figure 5A).

The addition of 1 mM L-tyrosine **4** yielded a significant increase in both RA **7** and IA **8** production ($p < 0.001$), suggesting that precursor availability is a significant bottleneck to the productivity of this pathway. Additionally, supplementing 1 mM *p*-coumaric acid **5** or 1 mM caffeic acid **6** resulted in a further increase in RA **7** production ($p < 0.05$), but only 1 mM caffeic acid **6** led to a significantly greater increase in IA **8** biosynthesis than feeding of 1 mM tyrosine ($p < 0.01$). These results suggest that, while the availability of tyrosine is a limiting factor in the donor substrate biosynthetic pathway, the conversion of tyrosine to *p*-coumarate by RsTAL is also a bottleneck, as has been shown in several other engineered hydroxycinnamic acid biosynthetic pathways.^[43, 47, 53–54]

Supplementation of 1 mM caffeic acid **6** and 1 mM 4-HPL **2** led to significantly higher RA **7** and IA **8** biosynthesis than any of the other added pathway intermediates ($p < 0.05$), suggesting that acceptor substrate availability and therefore the conversion of 4-HPP **1** to 4-HPL **2** by HdhA may also be a limiting factor. Interestingly, none of the supplemented cultures showed a significant difference in the ratio of IA **8** to RA **7** compared to the unfed culture, except when 1 mM caffeic acid **6** and 1 mM 4-HPL **2** were added, which increased the ratio of IA **8** to RA **7** 10-fold (Table S1). This increase in the ratio of IA **8** to RA **7** due to the addition of 4-HPL **2** illustrates the potential of this or related HCE biosynthetic pathways to be pushed toward production of the desired product by the addition of pathway precursors.

In order to investigate in more detail the flux through the donor and acceptor substrate branches of the engineered RA **7** pathway, we measured the concentrations of pathway intermediates when supplementing cultures with donor or acceptor substrates and precursors over time. When 1 mM caffeic acid **6** and 1 mM 4-HPL **2** were added to cultures, the sum of the molar concentrations of *p*-coumaric acid **5**, caffeic acid **6** and HCE products remained stable over time (Figure 5B). By contrast, the sum of the molar concentrations of acceptor substrates (4-HPL **2** and 3,4-DHPL **3**) and HCE products decreased significantly within the first 24 h (Figure 5C). This suggests that a leak exists in the acceptor substrate biosynthetic pathway. One possible explanation is that acceptor substrates are being lost to polyphenol

formation. Melanin production has been well-documented in *E. coli* through the overexpression of tyrosinase enzymes which catalyze the 3-hydroxylation of L-tyrosine **4** to L-DOPA and dopaquinone.^[59] Indeed, we observe the accumulation of a dark brown pigment on agar plates or liquid cultures of cells overexpressing *hpaBC* (Figure S3), presumably due to the conversion of endogenous L-tyrosine **4** to L-DOPA.^[53, 60–61] Similarly, 3,4-DHPL **3** may be further oxidized to 3,4-dihydroxyphenyllactone, non-enzymatically polymerized to a polyphenol and thus removed from the precursor pool of the HCE biosynthetic pathway.

Three orthologous RASs give different levels of RA and IA production, but a consistent ratio of IA to RA

Because of the low concentration of HCEs detected in the media and the detection of a high ratio of IA **8** to RA **7** in the media, we sought to compare the *in vivo* activity and product profile of CbRAS with other RAS orthologs. The DNA encoding two other characterized RAS genes, from *Lavandula angustifolia* (*laRAS*)^[36] and *Melissa officinalis* (*moRAS*)^[38] was synthesized based on their respective cDNA sequences, and the genes were assembled into the hydroxycinnamoyl transfer module of the HCE biosynthetic pathway, giving pACBB-4laR and pACBB-4moR (Figure 1). *E. coli* BW27784 were transformed with pUCBB-ABC, pCDFB-TBC and either pACBB-4cbR, pACBB-4laR, pACBB-4moR or empty pACBB as a control,^[50] transformants were cultured in minimal media, and *at4CL2* and RAS-encoding gene expression was induced 24 h after inoculation. At 72 h after induction, there was a significant difference in the concentrations of RA **7** and IA **8** detected with the three different RAS enzymes (Figure 6): LaRAS led to significantly higher HCE production than CbRAS ($p < 0.05$), up to 0.9 ± 0.2 μM RA **7** and 3.2 ± 0.4 μM IA **8** in the media; MoRAS led to significantly higher production than both LaRAS and CbRAS ($p < 0.05$), up to 1.8 ± 0.3 μM RA **7** and 5.3 ± 0.7 μM IA **8** in the media. An expression study of these enzymes did not show a difference in protein levels that correlated with HCE production, suggesting that the difference in HCE production levels may be due to differences in catalytic activity rather than expression (Figure S4). Interestingly, the ratio of IA **8** to RA **7** detected did not significantly differ between orthologs ($p = 0.2$); the average ratio of IA **8** to RA **7** in all samples was 4.0 ± 0.8 (Table S2). The lack of characterized RAS genes that favor production of RA **7** over IA **8** suggests that an enzyme engineering approach would be necessary to increase RAS substrate specificity toward 3,4-DHPL **3** and thus direct the pathway flux toward the RA **7** product.

Conclusions

In this work, we have successfully designed a chimeric biosynthetic pathway tailored for a microbial host that allows heterologous production of the plant-specific HCEs, RA **7** and IA **8**. To accomplish this, we implemented bacterial enzymes with known biosynthetic utility for hydroxycinnamic acid production, identified new substrates for known bacterial enzymes to establish 4-HPL **2** and 3,4-DHPL **3** biosynthesis in *E. coli*, and incorporated the plant enzymes At4CL2 and RAS to carry out hydroxycinnamoyl transfer. Incorporation of three plant RAS orthologs showed a significant variation in overall HCE biosynthesis, but not in the product profile of the HCEs observed, which remained about 4:1 IA **8** to RA **7** for all

three RASs. *E. coli* cultures expressing the RA **7** pathway containing the RAS from *M. officinalis* produced the highest levels of HCEs without precursor supplementation, up to $1.8 \pm 0.3 \mu\text{M}$ RA **7** and $5.3 \pm 0.7 \mu\text{M}$ IA **8**. Because of the modular pathway design, the RA **7** pathway could easily be altered for the biosynthesis of different HCEs through swapping the donor substrate, acceptor substrate and/or the hydroxycinnamoyl transfer modules. Feeding of alternative precursor compounds likewise would allow biosynthesis of structurally diverse HCEs; their economic production, however, would depend on access to inexpensive precursor sources. For example, 4-HPL **2** and 3,4-DHPL **3** are infeasible feeding precursors for industrial applications given their cost (about \$370/g and \$1850/g, respectively).

Although we achieved RA **7** and IA **8** biosynthesis, the current production levels are far less than sufficient for industrial application. One potential means of optimizing the HCE biosynthetic pathway would be to increase flux through the shikimate and tyrosine biosynthetic pathways of the host strain. Indeed, the levels of HCEs we observe are comparable to previously reported levels of the hydroxycinnamic acid amide avenathramide D produced from glucose in *E. coli* prior to optimizing tyrosine and shikimate pathway flux.^[43] In that work, a more than 100 fold increase in avenathramide D production was achieved through co-overexpression of rate-limiting enzymes and feedback-resistant mutant enzymes to increase flux and decrease feedback inhibition of the shikimate and tyrosine pathways.^[43] Introduction of our engineered HCE pathway into a tyrosine- or shikimate-overproducing strain would likely have a similar effect on production levels.

During the completion of this work, Yao et al. showed the production of 3,4-DHPL **3** (also known as salvianic acid A or danshensu) in *E. coli* through the overexpression of a mutant D-lactate dehydrogenase from *Lactobacillus pentosus* and HpaBC.^[62] The key mutations they used had been shown previously to convert the enzyme from a D-lactate dehydrogenase to a D-hydroxyisocaproate dehydrogenase^[63] and increase activity against phenylpyruvic acid.^[64] While Yao et al. do not report on the level of *de novo* production of 4-HPL **2** or 3,4-DHPL **3** without precursor feeding or engineering of upstream pathways for tyrosine overproduction, they were able to show up to 7.1 g/L 3,4-DHPL **3** with their tyrosine-overproducing strain in a batch-fed process.^[62] The implementation of our 3,4-DHPL **3** biosynthetic pathway in a tyrosine overproducing strain could therefore lead to similarly high titers of 3,4-DHPL **3**.

Several downstream pathway steps are also prime targets for optimization, such as TAL activity, which has been shown before to limit *E. coli* hydroxycinnamic acid biosynthetic pathways.^[43, 47, 53–54] Another optimization strategy must address the problem of acceptor substrate diversion. This could be achieved by co-localization of the acceptor substrate biosynthetic pathway and the RAS, which has proved to be a successful strategy to increase the flux through heterologous biosynthetic pathways.^[65] Besides an increase in production levels, the ability to eliminate byproducts and tune the engineered pathway towards the biosynthesis of RA **7** alone, rather than RA **7** and IA **8**, would increase the value of this pathway. For example, the ratio of RA **7** to IA **8** could be increased either by increasing the catalytic efficiency of 4-HPL **2** hydroxylation to 3,4-DHPL **3** and thus decreasing the pool of 4-HPL **2** in the cell or by altering a RAS to more specifically bind 3,4-DHPL **3**. In summary, given the numerous options for pathway optimization, this work provides the

basis for the development of a valuable microbial production platform for RA **7** and other HCEs.

Experimental Section

Chemicals and Enzymes

Chemicals used in this study were procured as follows: L-tyrosine **4**, caffeic acid **6**, 4-hydroxyphenylpyruvic acid **1**, DL-4-hydroxyphenyllactic acid, rosmarinic acid **7** (Sigma Aldrich, St. Louis, MO, USA), *p*-coumaric acid **5** (MP Biomedical, LLC, Solon, OH, USA), sodium danshensu (3,4-dihydroxyphenyllactic acid, Select Chemicals, Houston, TX, USA), L-(+)-arabinose (ICN Biomedical, Inc., Aurora, OH, USA). The following enzymes and kits were used for DNA manipulation according to the manufacturers' protocols: Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA), RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), SuperScript® III Reverse Transcriptase, Zero Blunt® TOPO® PCR Cloning Kit, (Life Technologies, Carlsbad, CA, USA), restriction enzymes, T4 DNA ligase, Phusion® High Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA).

Strains, Media and Growth Conditions

E. coli C2566 (New England Biolabs, Ipswich, MA, USA) was used for plasmid manipulation and propagation and was cultivated in Luria-Bertani (LB) medium at 37 °C. *L. delbrueckii* subsp. *bulgaricus* (DSM-20081) was cultivated anaerobically in MRS broth (DSMZ Medium 11) at 37 °C. *S. espanaensis* (DSM-44229) was cultivated on GYM agar plates (DSMZ Medium 65) at 28 °C. All biosynthesis experiments were carried using *E. coli* BW27784 (*E. coli* Genetic Stock Center, New Haven, CT).^[66] Strains used in this study are listed in Table S3.

For production experiments, seed cultures of *E. coli* BW27784 containing pathway plasmids were cultivated from single colonies overnight and grown in LB media at 37 °C with the appropriate antibiotics: ampicillin (100 µg/mL), chloramphenicol (50 µg/mL), and/or streptomycin (50 µg/mL). After overnight growth these seed cultures were used to inoculate cultures in modified M9 media containing yeast extract (1.25 g/L) and glycerol (5 g/L) instead of glucose and supplemented with the appropriate antibiotics. For cultures to test the biosynthesis of 4-HPL **2**, the bioconversion of *p*-coumaric acid **5** and 4-HPL **2** to caffeic acid **6** and 3,4-DHPL **3**, and the relative expression levels of three RAS orthologs, flasks cultures (100 mL) were inoculated with the seed culture (1 mL); for all other experiments, test tube cultures (10 mL) were inoculated with the seed culture (100 µL). Where appropriate, the expression of *at4CL2* and RAS-encoding genes was induced 24 h after culture inoculation by the addition of arabinose to a final concentration of 0.2 %w/v. For precursor supplementation experiments, *p*-coumaric acid **5**, caffeic acid **6**, or caffeic acid **6** plus 4-HPL **2** (50 µL each) were added to the cultures from 200 mM stocks in methanol to a final concentration of 1 mM after 24 h at the same time as induction; tyrosine was added to a final concentration of 1 mM by the addition of dry powder (2 mg) at the time of induction. All biosynthesis experiments were carried out in three biological replicates.

Gene Cloning, Plasmid Construction and Enzyme Expression

All gene cloning and pathway assembly was carried out using our custom set of BioBrick™-compatible vectors^[50] which are available from Addgene (www.addgene.org). The open reading frame (ORF) encoding cbRAS (GenBank: CAK55166.1) was amplified from *C. blumei* cDNA, which was synthesized from RNA extracted from plant tissue donated by the University of Minnesota Horticulture Club. An ORF encoding At4CL2 (GenBank: AAD47192.1) was amplified from *A. thaliana* cDNA which was a generous gift from the Olszewski laboratory at the University of Minnesota. The gene encoding D-2-hydroxyisocaproate dehydrogenase HdhA (GenBank: CAI97812.1) was amplified from *L. delbrueckii* subsp. *bulgaricus* genomic DNA in two pieces using mutagenic primers designed to remove an EcoRI site in the ORF. The resulting PCR products were amplified in an overlap-extension PCR to give the full-length sequence with a silent mutation eliminating the EcoRI site. The gene encoding Sam5 (GenBank: CCH33123.1) was amplified from *S. espanaensis* genomic DNA. The genes encoding HpaB (GenBank: CAQ34705.1) and HpaC (GenBank: AAZ90964.1) were amplified individually from *E. coli* BL21 (DE3) genomic DNA. The gene encoding RsTAL from *R. sphaeroides* (GenBank: ABA81174.1) was amplified from pUC-TAL^[56] in two pieces with mutagenic primers to remove the EcoRI site, as described above for HdhA. ORFs encoding LaRAS (GenBank: ABI48360.1) and MoRAS (GenBank: CBW35684.1) were synthesized with their native codon usage (Life Technologies, Carlsbad, CA, USA). All genes were amplified with forward primers containing either a BglII or a NdeI site at the 5' end, and a reverse primer containing either an NsiI or NotI site and a stop codon at the 5' end. The genes for At4CL2 and CbRAS were first TOPO-cloned using the Zero Blunt® TOPO® PCR Cloning Kit according to the manufacturer's protocol. Plasmids TOPO-cbRAS and TOPO-4CL2 were used for future sub-cloning into pUCBB. All other amplified genes were digested with the appropriate restriction enzymes and ligated into pUCBB digested with the same enzymes. The resulting pUCBB-gene plasmids were used for subsequent subcloning and pathway assembly. Full primer sequences are listed in Table S4.

An initial screen of *at4CL2* and *cbRAS* expression did not show visible expression on an SDS-PAGE gel (Figure S5A). Therefore, an N-terminal His-tag was added to each of these enzymes by amplification of the genes with forward primers containing NdeI sites and the reverse primers containing NotI sites and subsequent digestion and cloning into pUCBB-ntH6^[50], which encodes a His-tag and thrombin cleavage site just upstream of the NdeI site. Subsequent expression studies showed that N-terminally His-tagged At4CL2 protein was visible *via* SDS-PAGE (Figure S5B). Additionally, an anti-His-tag Western blot showed the presence of both N-terminally-His-tagged At4CL2 and CbRAS (Figure S5C). Because of the apparent increase in At4CL2 protein levels and the confirmed presence of CbRAS protein, the N-terminally-tagged versions of *at4CL2* and *cbRAS* were used for all subsequent subcloning and cultivations. Additionally, *laRAS* and *moRAS* were cloned into pUCBB-ntH6 before further subcloning, such that all RAS enzymes tested carried the same N-terminal His-tag.

The assembly of plasmids containing multiple enzymes was carried out *via* the BioBrick™ prefix-stacking strategy as described in^[50]. Briefly, for example, after cloning *hpaB* and

hpaC individually into the pUCBB backbone, giving pUCBB-*hpaB* and pUCBB-*hpaC*; pUCBB-*hpaC* was then digested with *EcoRI* and *XbaI*, while pUCBB-*hpaB* was digested with *EcoRI* and *SpeI*. The resulting DNA fragment containing the modified constitutive lac promoter P_{lac} ,^[50] the ribosome binding site and the *hpaB* ORF was subsequently ligated into the digested pUCBB-*hpaC*, to give pUCBB-*hpaBC*, a plasmid containing a BioBrick™ comprised of a distinct expression cassette for each gene. All plasmids containing multiple genes were constructed in this way by prefix-stacking of the appropriate cassette into the appropriate plasmid with the pUCBB backbone. Once the multi-enzyme BioBrick™ was constructed in the pUCBB backbone, if needed, it was subsequently moved to the pACBB or pCDFBB backbones^[50] through digestion with *EcoRI* and *SpeI* and ligation to the desired backbone digested with the same enzymes. A list of plasmids created for this study and their descriptions is given in Table S5.

HPLC and LC-MS Analysis of Pathway Intermediates and Products

For biosynthesis and feeding experiments, cultures were sampled at various timepoints by removing a sample of the culture (1 mL), pelleting the cells, and storing the supernatant at -20°C . Samples were later thawed at room temperature, acidified with 1N hydrochloric acid (50 μL) and extracted twice with ethyl acetate (2 mL). The ethyl acetate fraction was dried to completion and resuspended in methanol (100 μL). For the detection and quantification of pathway intermediates, samples were separated on a Zorbax SB-C18 column (5 μM , 4.6 X 150 mm) on an Agilent 1100 Series system. A flow rate of 1 mL/min according was used according to the following timetable: 5 min at 10% methanol, a gradient to 50% methanol for 20 min, and 10 min post-run at 10% methanol. For the detection and quantification of HCEs, samples were separated on a Zorbax SB-C18 column (5 μM , 4.6 X 250 mm) at a flow rate of 1 mL/min according the following timetable: 5 min at 45% methanol, a gradient to 55% methanol for 10 min, and 5 min post-run at 45% methanol. All buffers contained trifluoroacetic acid (0.1 % v/v). RA **7**, IA **8** and caffeic acid **6** were detected at 330 nm; 4-HPL **2** and 3,4-DHPL **3** were detected at 223nm; *p*-coumaric acid **5** was detected at 310 nm. The identities and quantities of all intermediates and products except IA **8** were verified by the comparison of retention times and UV spectra to authentic standards. Because IA **8** is not commercially available, and due to the similarity of the UV spectra of IA **8** and RA **7** (Figure S2), IA **8** was quantified using the RA **7** standard curve. For LC-MS analysis of HCEs, the HPLC method was transferred to a Thermo Finnigan Spectra System HPLC in line with a Thermo Finnigan LCQ ion trap. The same HPLC method was used as the one listed for HCE detection above, except that for elution the trifluoroacetic acid was replaced with glacial acetic acid (0.1 % v/v). Negative-mode ESI was used for ionization, and ms/ms scanning events were set up for parent ion masses for RA **7** (359.2 ± 2 m/z) and the mass of dehydroxylated RA **7** (343.2 ± 2 m/z) using 30% ionization energy for fragmentation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Dr. Neil Olszewski and Dr. Kerry Sokol for the gift of *A. thaliana* cDNA and the University of Minnesota Horticulture Club for *C. blumei* tissue for RNA extraction. We also thank Tom Krick for assistance setting up the LC-MS protocol. This work was supported by funding from the Biocatalysis Initiative at the University of Minnesota administered by the Biotechnology Institute and a grant from the Academic Health Center of the University of Minnesota. S.E.B. was supported by a predoctoral National Institute of Health traineeship (GM008347).

References

- Petersen M, Simmonds MS. Phytochemistry. 2003; 62:121–125. [PubMed: 12482446]
- Bulgakov VP, Inyushkina YV, Fedoreyev SA. Crit Rev Biotechnol. 2012; 32:203–217. [PubMed: 21838541]
- Moon DO, Kim MO, Lee JD, Choi YH, Kim GY. Cancer Lett. 2010; 288:183–191. [PubMed: 19619938]
- Paluszczak J, Krajka-Kuzniak V, Baer-Dubowska W. Toxicol Lett. 2010; 192:119–125. [PubMed: 19840838]
- Xu Y, Jiang Z, Ji G, Liu J. Planta Med. 2010; 76:956–962. [PubMed: 20157877]
- Anusuya C, Manoharan S. J Environ Pathol Toxicol Oncol. 2011; 30:199–211. [PubMed: 22126613]
- Venkatachalam K, Gunasekaran S, Jesudoss VA, Namasivayam N. Exp Toxicol Pathol. 2013; 65:409–418. [PubMed: 22236574]
- Fallarini S, Miglio G, Paoletti T, Minassi A, Amoruso A, Bardelli C, Brunelleschi S, Lombardi G. Br J Pharmacol. 2009; 157:1072–1084. [PubMed: 19466982]
- Kelsey NA, Wilkins HM, Linseman DA. Molecules. 2010; 15:7792–7814. [PubMed: 21060289]
- Shimojo Y, Kosaka K, Noda Y, Shimizu T, Shirasawa T. J Neurosci Res. 2010; 88:896–904. [PubMed: 19798750]
- Braidy N, Matin A, Rossi F, Chinain M, Laurent D, Guillemin GJ. Neurotox Res. 2014; 25:226–234. [PubMed: 24097334]
- Park DH, Park SJ, Kim JM, Jung WY, Ryu JH. Fitoterapia. 2010; 81:644–648. [PubMed: 20230877]
- Hamaguchi T, Ono K, Murase A, Yamada M. Am J Pathol. 2009; 175:2557–2565. [PubMed: 19893028]
- Wang J, Xu H, Jiang H, Du X, Sun P, Xie J. J Mol Neurosci. 2012; 47:113–119. [PubMed: 22205146]
- Jin X, Liu P, Yang F, Zhang YH, Miao D. Neurochem Res. 2013; 38:1828–1837. [PubMed: 23756732]
- Nie H, Peng Z, Lao N, Wang H, Chen Y, Fang Z, Hou W, Gao F, Li X, Xiong L, Tan Q. Prog Neuropsychopharmacol Biol Psychiatry. 2014; 51:16–22. [PubMed: 24418162]
- Zinsmeister HD, Becker H, Eicher T. Angew Chem. 1991; 30:130–147.
- Eicher T, Ott M, Speicher A. Synthesis-Stuttgart. 1996:755.
- Pabsch K, Petersen M, Rao NN, Alfermann AW, Wandrey C. Recueil Des Travaux Chimiques Des Pays-Bas-Journal of the Royal Netherlands Chemical Society. 1991; 110:199–205.
- Reimann E, Maas HJ, Pflug T. Monatshefte Fur Chemie. 1997; 128:995–1008.
- Huang LJ, Li CH, Lu ZM, Ma ZB, Yu DQ. J Asian Nat Prod Res. 2006; 8:561–566. [PubMed: 16931433]
- Petersen MS. Phytochemistry. 1991; 30:2877–2881.
- Su WW, Lei F, Su LY. Biotechnol Bioeng. 1993; 42:884–890. [PubMed: 18613136]
- Georgiev M, Pavlov A, Ilieva M. Process Biochem. 2006; 41:2068–2071.
- Francoise B, Hossein S, Halimeh H, Zahra NF. Pak J Biol Sci. 2007; 10:3395–3399. [PubMed: 19090157]
- Lee SY, Xu H, Kim YK, Park SU. World J Microbiol Biotechnol. 2008; 24:969–972.

27. Georgiev M, Abrashev R, Krumova E, Demirevska K, Ilieva M, Angelova M. *Appl Biochem Biotechnol.* 2009; 159:415–425. [PubMed: 19050834]
28. Eberle D, Ullmann P, Werck-Reichhart D, Petersen M. *Plant Mol Biol.* 2009; 69:239–253. [PubMed: 18982412]
29. Renault H, Bassard JE, Hamberger B, Werck-Reichhart D. *Curr Opin Plant Biol.* 2014; 19C:27–34. [PubMed: 24709279]
30. Negrel J, Javelle F. *Eur J Biochem.* 1997; 247:1127–1135. [PubMed: 9288939]
31. Yang Q, Reinhard K, Schiltz E, Matern U. *Plant Mol Biol.* 1997; 35:777–789. [PubMed: 9426598]
32. Yu M, Facchini PJ. *Planta.* 1999; 209:33–44. [PubMed: 10467029]
33. Yonekura-Sakakibara K, Tanaka Y, Fukuchi-Mizutani M, Fujiwara H, Fukui Y, Ashikari T, Murakami Y, Yamaguchi M, Kusumi T. *Plant Cell Physiol.* 2000; 41:495–502. [PubMed: 10845463]
34. Back K, Jang SM, Lee BC, Schmidt A, Strack D, Kim KM. *Plant Cell Physiol.* 2001; 42:475–481. [PubMed: 11382813]
35. Comino C, Lanteri S, Portis E, Acquadro A, Romani A, Hehn A, Larbat R, Bourgaud F. *Bmc Plant Biology.* 2007; 7
36. Landmann C, Hucherig S, Fink B, Hoffmann T, Dittlein D, Coiner HA, Schwab W. *Planta.* 2011; 234:305–320. [PubMed: 21424826]
37. Sander M, Petersen M. *Planta.* 2011; 233:1157–1171. [PubMed: 21318289]
38. Weitzel C, Petersen M. *Phytochemistry.* 2011; 72:572–578. [PubMed: 21354582]
39. Kim IA, Kim BG, Kim M, Ahn JH. *Phytochemistry.* 2012; 76:25–31. [PubMed: 22285622]
40. Kang K, Park M, Park S, Kim YS, Lee S, Lee SG, Back K. *Biotechnol Lett.* 2009; 31:1469–1475. [PubMed: 19479321]
41. Kang K, Back K. *Metab Eng.* 2009; 11:64–68. [PubMed: 18805501]
42. Kim BG, Jung WD, Mok H, Ahn JH. *Microb Cell Fact.* 2013; 12:15. [PubMed: 23383718]
43. Eudes A, Juminaga D, Baidoo EE, Collins FW, Keasling JD, Loque D. *Microb Cell Fact.* 2013; 12:62. [PubMed: 23806124]
44. Gosset G. *Curr Opin Biotechnol.* 2009; 20:651–658. [PubMed: 19875279]
45. Escalante A, Calderon R, Valdivia A, de Anda R, Hernandez G, Ramirez OT, Gosset G, Bolivar F. *Microb Cell Fact.* 2010; 9:21. [PubMed: 20385022]
46. Munoz AJ, Hernandez-Chavez G, de Anda R, Martinez A, Bolivar F, Gosset G. *J Ind Microbiol Biotechnol.* 2011; 38:1845–1852. [PubMed: 21512819]
47. Kang SY, Choi O, Lee JK, Hwang BY, Uhm TB, Hong YS. *Microb Cell Fact.* 2012; 11:153. [PubMed: 23206756]
48. Santos CN, Xiao W, Stephanopoulos G. *Proc Natl Acad Sci U S A.* 2012; 109:13538–13543. [PubMed: 22869698]
49. Lin Y, Sun X, Yuan Q, Yan Y. *Metab Eng.* 2014; 23:62–69. [PubMed: 24583236]
50. Vick JE, Johnson ET, Choudhary S, Bloch SE, Lopez-Gallego F, Srivastava P, Tikh IB, Wawrzyn GT, Schmidt-Dannert C. *Appl Microbiol Biotechnol.* 2011; 92:1275–1286. [PubMed: 22033566]
51. Knight T. MIT Synthetic Biology Working Group Technical Reports. 2003
52. Choi O, Wu CZ, Kang SY, Ahn JS, Uhm TB, Hong YS. *J Ind Microbiol Biotechnol.* 2011; 38:1657–1665. [PubMed: 21424580]
53. Lin Y, Yan Y. *Microb Cell Fact.* 2012; 11:42. [PubMed: 22475509]
54. Zhang H, Stephanopoulos G. *Appl Microbiol Biotechnol.* 2013; 97:3333–3341. [PubMed: 23179615]
55. Bernard N, Johnsen K, Ferain T, Garmyn D, Hols P, Holbrook JJ, Delcour J. *Eur J Biochem.* 1994; 224:439–446. [PubMed: 7925358]
56. Watts KT, Lee PC, Schmidt-Dannert C. *ChemBiochem.* 2004; 5:500–507. [PubMed: 15185374]
57. Ehrling J, Buttner D, Wang Q, Douglas CJ, Somssich IE, Kombrink E. *Plant J.* 1999; 19:9–20. [PubMed: 10417722]
58. Berger A, Meinhard J, Petersen M. *Planta.* 2006; 224:1503–1510. [PubMed: 17047986]

59. della-Cioppa G, Garger SJ, Sverlow GG, Turpen TH, Grill LK. *Biotechnology* (N Y). 1990; 8:634–638. [PubMed: 1367455]
60. Lagunas-Munoz VH, Cabrera-Valladares N, Bolivar F, Gosset G, Martinez A. *J Appl Microbiol.* 2006; 101:1002–1008. [PubMed: 17040223]
61. Chavez-Bejar MI, Balderas-Hernandez VE, Gutierrez-Alejandre A, Martinez A, Bolivar F, Gosset G. *Microb Cell Fact.* 2013; 12:108. [PubMed: 24225202]
62. Yao YF, Wang CS, Qiao J, Zhao GR. *Metab Eng.* 2013; 19:79–87. [PubMed: 23774671]
63. Tokuda C, Ishikura Y, Shigematsu M, Mutoh H, Tsuzuki S, Nakahira Y, Tamura Y, Shinoda T, Arai K, Takahashi O, Taguchi H. *J Bacteriol.* 2003; 185:5023–5026. [PubMed: 12897026]
64. Ishikura Y, Tsuzuki S, Takahashi O, Tokuda C, Nakanishi R, Shinoda T, Taguchi H. *J Biochem.* 2005; 138:741–749. [PubMed: 16428303]
65. Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KL, Keasling JD. *Nat Biotechnol.* 2009; 27:753–759. [PubMed: 19648908]
66. Khlebnikov A, Datsenko KA, Skaug T, Wanner BL, Keasling JD. *Microbiology.* 2001; 147:3241–3247. [PubMed: 11739756]

Figure 1. Plasmids assembled for donor substrate, acceptor substrate and HCE biosynthesis
 All plasmids were assembled using the BioBrick™ system previously designed and built by our lab [50]. $P_{lac'}$, a modified constitutive *lac* promoter; P_{araBAD} , arabinose inducible promoter; *araC*, gene encoding the repressor of P_{araBAD} ; pUCBB, high copy number plasmid backbone; pCDFBB, medium copy number plasmid backbone; pACBB, low copy number plasmid backbone. Plasmids that constitute the functional three-module RA 7 biosynthetic pathway are boxed.

Figure 2. *De novo* biosynthesis of 4-HPL by *E. coli* BW27784 cells transformed with pUCBB-hdhA

Transformants were cultured in minimal media at 30 °C; culture samples were taken over time, cleared of cells, and extracted with ethyl acetate for HPLC analysis. An analytical standard of 4-HPL **2** was used to confirm peak identity and quantification. No 4-HPL **2** was detected in cultures of cells transformed with the empty pUCBB vector.

Figure 3. Bioconversion of *p*-coumaric acid and 4-HPL to caffeic acid and 3,4-DHPL, respectively, by *E. coli* BW27784 transformants of pUCBB-hpaBC

Transformants were cultured in minimal media at 30 °C, supplemented with 1 mM *p*-coumaric acid or 1 mM 4-HPL. 2 culture and samples were taken over time and processed for HPLC analysis. Samples were compared to standard compounds in order to identify compound peaks. Solid trace, pUCBB-hpaBC transformants; dotted trace, empty pUCBB transformants (control).

Figure 4. *De novo* biosynthesis of HCEs in *E. coli*

E. coli BW27784 was co-transformed with pUCBB-ABC containing the acceptor substrate module, pCDFBB-TBC4 containing the acceptor substrate module, and pACBB-4cBR containing the hydroxycinnamoyl transfer module (Figure 1). Transformants were cultured in minimal media at 30 °C, and *at4CL2* and *cbRAS* expression was induced after 24 h growth. Culture samples were taken over time and processed for HPLC and LC-MS analysis. An analytical standard of RA **7** was used to confirm peak identity and mass fragmentation of RA **7**, and MS/MS mass spectra confirm the presence of both RA **7** and IA **8** in culture media. Asterisks denote key mass fragments in the identification of the acceptor substrate moieties of the products.

Figure 5. Limiting factors and leaks in the *E. coli* HCE biosynthetic pathway

E. coli BW27784 containing the modular HCE biosynthesis plasmids were cultured in minimal media at 30 °C; after 24h, *at4CL2* and *cbRAS* expression was induced and various supplements were added to a final concentration of 1 mM. Culture samples were taken over time and analyzed by HPLC. Asterisks denote significantly different concentrations at $p < 0.05$. A) HCE production is increased in the presence of different feeding precursors 48h after induction of *at4CL2* and *cbRAS*. B) The total concentrations of donor substrates and HCE products remain stable over time. C) The total concentrations of acceptor substrates and HCE products significantly decreases after 24 h.

Figure 6. HCE biosynthesis in *E. coli* cultures expressing three orthologous RAS enzymes in the context of the redesigned biosynthetic pathway

Two new hydroxycinnamoyl transfer module plasmids were constructed containing *laRAS* and *moRAS*. *E. coli* BW27784 were co-transformed with empty pACBB, pACBB-4cbR, pACBB-4laR or pACBB-4moR, pUCBB-ABC and pCDFBB-TBC and analyzed for HCE production 72 h after induction of the *at4CL2* and RAS-encoding genes. Asterisks denote significantly different concentrations at $p < 0.01$.

Scheme 1. Redesigned chimeric RA biosynthetic pathway

The D-hydroxyisocaproate dehydrogenase HdhA from *L. delbrueckii* subsp. *bulgaricus* catalyzes the conversion of endogenous 4-HPP **1** to 4-HPL **2**, which is subsequently meta-hydroxylated by hydroxylase complex HpaBC cloned from *E. coli* to yield 3,4-DHPL **3**, the intended acceptor substrate for rosmarinic acid synthase (RAS). Donor substrate biosynthesis is achieved through the conversion of L-tyrosine **4** to p-coumaric acid by a tyrosine ammonia lyase from *R. sphaeroides* (RsTAL) and hydroxylation of p-coumaric acid **5** to caffeic acid **6** by HpaBC. A 4-coumaroyl-CoA ligase from *Arabidopsis thaliana* (At4CL2) generates the CoA-ester of the donor substrate, allowing ester formation catalyzed by RAS. Both 3,4-DHPL **3** and 4-HPL **2** may be accepted by RAS to form both RA **7** and IA **8**, respectively.