Advanced Synthesis & Catalysis

Accepted Article

Title: Coenzyme A-Conjugated Cinnamic Acids – Enzymatic Synthesis of a CoA-Ester Library and Application in Biocatalytic Cascades to Vanillin Derivatives

Authors: Martin Dippe, Anne-Katrin Bauer, Andrea Porzel, Evelyn Funke, Anna Müller, Jürgen Schmidt, Maria Beier, and Ludger A. Wessjohann

To be cited as: *Adv. Synth. Catal.* 10.1002/adsc.201900892

Link to VoR: http://dx.doi.org/10.1002/adsc.201900892

10.1002/adsc.201900892

COMMUNICATION

DOI: 10.1002/adsc.201((will be filled in by the editorial staff))

Coenzyme A-Conjugated Cinnamic Acids – Enzymatic Synthesis of a CoA-Ester Library and Application in Biocatalytic Cascades to Vanillin Derivatives

Martin Dippe, Anne-Katrin Bauer, Andrea Porzel, Evelyn Funke, Anna O. Müller, Jürgen Schmidt, Maria Beier and Ludger A. Wessjohann*

Leibniz-Institute of Plant Biochemistry, Department of Bioorganic Chemistry, Weinberg 3, D-06120 Halle, Germany

Tel.: +49-345-55821301, Fax: +49-345-55821309, email: wessjohann@ipb-halle.de

Received: ((will be filled in by the editorial staff))

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/adsc.201######.

Abstract. We present a bioorthogonal method for the ligation of coenzyme A (CoA) with cinnamic acids. The reaction, which is the initial step in the biosynthesis of a multitude of bioactive secondary metabolites, is catalyzed by a promiscuous plant ligase and yields CoA conjugates with different functionalization in high purity and without formation of by-products. Its applicability in biosynthetic cascades is shown for the direct transformation of cinnamic acids into natural benzaldehydes (like vanillin) or synthetic derivatives (e.g. ethylvanillin).

Keywords: Aldehydes; Biocatalysis; Carboxylic acids; Ligases; Lyases

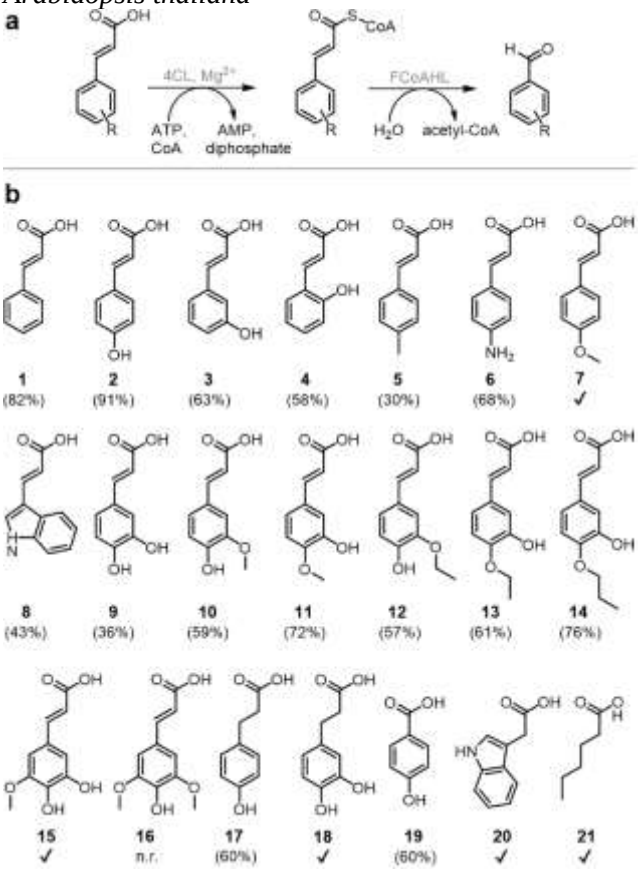
Coenzyme A (CoA) serves as a carrier and activator for carboxylic acids in more than 100 types of biochemical reactions.[1] Over 4 % of all known enzymes catalyze transformations which involve such CoA conjugates, e.g. the oxidoreductases acting in fatty acid catabolism (such as beta-oxidation to acetyl-CoA[2] and the subsequent oxidation of this intermediate in the Krebs cycle)[3], in the reduction of acids to aldehydes[4], or in hydroxylation of aromatic acids.[5] The activation of the carboxylic acid substrate by the reactive thioester bond of CoA esters is also indispensable for acylation reactions on small molecules and proteins, e.g. in gene-regulatory histone acetylation.[6] In particular, a number of primary and secondary metabolites are synthesized via repeated condensation of CoA conjugates. These reactions are the basis for several pathways to polyketide natural products, such as antibiotics[7], certain alkaloids and quinones[8] or flavonoids.[9] The starting point for the biosynthesis of the latter class of compounds is the ATP-dependent conversion of *trans*-4-coumaric acid (**2**, Scheme 1b) to its CoA ester (4-coumaroyl-CoA) by 4-coumarate:CoA ligase (4CL, E.C. 6.2.1.12) (Sche-me 1a).

Compared to other acyl donors involved in enzymatic reactions (e.g. mixed anhydrides such as phosphoacylates), CoA esters are much more stable against hydrolytic cleavage under physiological conditions.[10] This stability and the importance of the compounds in a multitude of enzyme reactions facilitated the development of chemosynthetic strategies for their production, which are in general based on the acylation of CoA or precursors with reactive acyl chlorides, carbodiimide adducts[11] or imidazolides.[10,12] Although high yields of the thioesters can be achieved in these reactions, they cannot be directly applied in bioorthogonal (enzymatic) syntheses or in synthetic biology due to the cross-reactivity of the used acyl donors with reactants and enzymes. Thus, enzyme-mediated syntheses seem to be the most convenient strategy for *in situ* or *in vivo* generation of CoA conjugates in biocatalytic cascades.

Ideally, a cascading reaction including a thioester formation step should rely on a CoA ligase with relaxed substrate specificity. The use of a promiscuous enzyme offers the possibility to react a broad range of carboxylic acid substrates, providing the basis for both the synthesis of natural or new-to-nature derivatives in subsequent reaction steps. In the search for a candidate with suitable substrate scope,

10.1002/adsc.201900892

the 4CL isoform 2 (4CL2) from the model plant *Arabidopsis thaliana*



Scheme 1. Synthesis of CoA esters in the biocatalytic cascade to substituted benzaldehydes. a) Enzymatic activation of cinnamic acids (left) yielding the corresponding cinnamoyl-CoA thioesters (middle) by 4CL, followed by FCoAHL-catalyzed hydration and retro-aldol reaction to substituted benzaldehydes (right). b) Selected acids used as 4CL substrates in this study. cinnamic acid (**1**), 4-coumaric acid (**2**), 3-coumaric acid (**3**), 2-coumaric acid (**4**), 4-methyl-cinnamic acid (**5**), 4-aminocinnamic acid (**6**), 4-methoxy-cinnamic acid (**7**), 3-indoleacrylic acid (**8**), caffeic acid (**9**), ferulic acid (**10**), isoferulic acid (**11**), homoferulic acid (**12**), 4-ethoxy-3-hydroxycinnamic acid (**13**), 3-hydroxy-4-*n*-propoxycinnamic acid (**14**), 5-hydroxyferulic acid (**15**), sinapic acid (**16**), dihydro-4-coumaric acid (**17**), dihydrocaffeic acid (**18**), 4-hydroxybenzoic acid (**19**), indole-3-acetic acid (**20**), hexanoic acid (**21**). The yield (in %, calculated on the basis of CoA as the most expensive component) in preparative transformations is given in brackets. AMP, adenosine-5´-monophosphate; CoA, coenzyme A; n.r., not reactive; ✓, reactive but not used in preparative reaction due to low conversion rates (see Tab. S1) or high susceptibility to oxidation. *General conditions*: 50 mM Tris/HCl (pH 7.5), 2.5 mM $MgCl_2$, 0.265 mM CoA, 2 mM acid substrate, 1.25 mM ATP, 25 µg ml^{-1} 4CL2. Reaction time: 20 h. Fresh ligase enzyme (25 µg ml^{-1} 4CL2) and cosubstrates (0.265 mM CoA, 1.25 mM ATP) were added five hours after initiation of the reactions. Products were purified by SPE (see Supporting Information).

proved to be beneficial. The enzyme is characterized by high activity and stability and – unlike other 4CL enzymes which are prone to aggregation in heterologous expression – can be produced in high yield in *Escherichia coli* as production host. [14] In particular, the ability of this ligase to convert the three cinnamates 4-coumaric acid (**2**), caffeic acid (**9**) and ferulic acid (**10**)[13] suggests a potential promiscuity towards carboxylic acids. This tempted us to probe the substrate scope of the enzyme.

Thus, the activity of 4CL2 was systematically tested on a set of cinnamates, substituted either with one hydroxyl (**2** – **4**), methyl (**5**), amino (**6**) or methoxyl (**7**) group, or with several hydroxyl (**9**) and/or alkoxyl substituents (**10** – **16**). Other compounds included in the screening differed from the native substrate **2** in the substitution of the phenyl group by the heterocycle indole (**8**), or in saturation (**17**, **18**) or absence of the vinyl moiety (**19**). Strikingly, conversion into the thioesters – which was detected spectroscopically by depletion of the co-substrate CoA – could be observed for all substrates except for the highly substituted sinapic acid (**16**). Even benzoic and arylacetic acids were accepted by the enzyme (reaction rates at a substrate concentration of 600 µM are given in brackets): 4CL2 transformed 4-hydroxybenzoic acid (**19**) (27 ± 1 nmol min^{-1} mg^{-1}) and indole-3-acetic acid (**20**) (2 ± 1 nmol min^{-1} mg^{-1}). Among nine fatty acids with different chain lengths, only hexanoic acid (**21**) was converted (86 ± 1 nmol min^{-1} mg^{-1})*. The kinetic constants of 4CL2 (see Fig. S1 in the Supporting Information) reflect its broad substrate specificity but

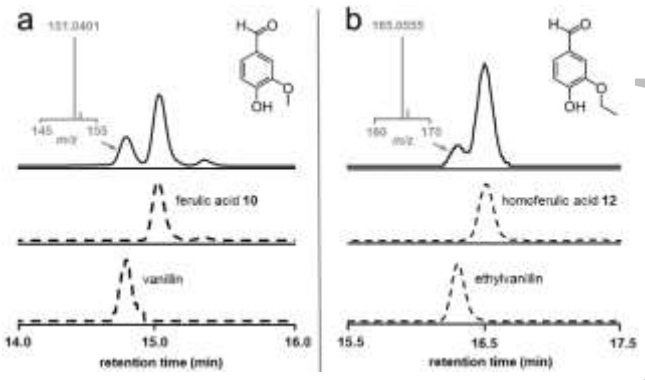


Figure 1. Conversion of cinnamic acids into vanillin and related benzaldehydes by the 4CL2/FCoAHL cascade. The chromatograms show the enzymatic formation of vanillin (a) or ethylvanillin (b) from the corresponding cinnamic acids (top row). Standards of the acid substrates and vanillin products (structures depicted at top row) are shown at the center and in the lowermost row (chromatograms with dashed lines). The products of the enzymatic reactions were identified by HR-MS/MS (gray insets, see Supporting Information for methods). *General conditions*: 50 mM Tris/HCl (pH 7.5), 2.5 mM $MgCl_2$, 0.265 mM CoA, 2 mM acid substrate, 1.25 mM ATP, 25

µg ml^{-1} 4CL2. Reaction time: 20 h. Fresh ligase enzyme (25 µg ml^{-1} 4CL2) and cosubstrates (0.265 mM CoA, 1.25 mM ATP) were added five hours after initiation of the reaction. The transformation into vanillins was started by addition of FCoAHL (85 µg ml^{-1}, incubation for 14 h).

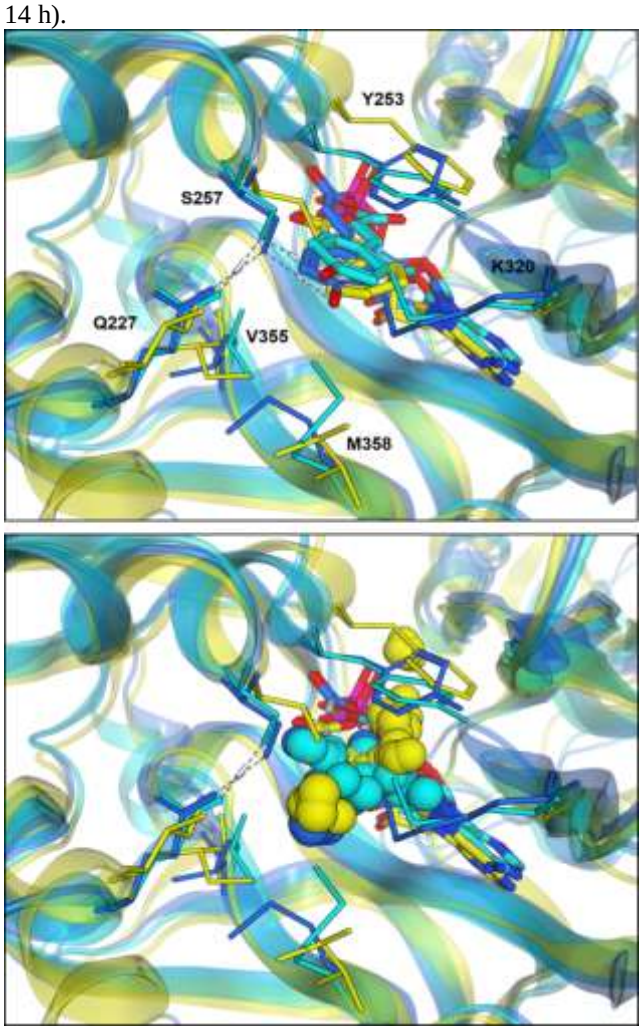


Figure 2. Active sites of plant 4CL enzymes. Modelled structures of 4CL2 from *A. thaliana* (dark blue) and cinnamic acid-CoA ligase from *S. baicalensis* (yellow) were aligned with the crystal structure of the tobacco enzyme[18] (cyan). Upper panel: Substrate binding sites occupied by the reaction intermediate 4-coumaroyl adenosine 5´-mono-phosphate. Amino acids involved in binding (numbering of positions indicated for the *Arabidopsis* enzyme) are shown in stick representation. Lower panel: Sphere model of the cavity accommodating the aromatic acid moiety. The coloration of the amino acid residues, bound molecules and sphere models is the same as for the protein scaffolds.

also hints to a distinct selectivity towards the different cinnamic acids. The maximum reaction rate (v_{max}) specifically depends on the substitution pattern of the aromatic ring, e.g. on the type of the substituent (*p*-OH > *p*-H = *p*-Me (*p*-OH > *m*-OH = *o*-OH) or alkoxy residues (*p*-OMe in **11** > *m*-OMe in **10**)) and on the size of the attached alkyl chain (*p*-OMe > *p*-OEt = *p*-OPr) (see Tab. S1 in the Supporting Information).

The acceptance of the naturally occurring cinnamic acids **2**, **9** and **10** was also reported for other 4CL enzymes which are involved in plant lignin biosynthesis, such as the isoenzymes 4CL1 and 4CL3 from *A. thaliana*[13], the orthologs from poplar[15], rowanberry[16] and mulberry[17], or the structurally characterized enzyme from tobacco[18]. In order to gain insights into substrate recognition by these promis-cuous enzymes, the active site architecture of the tobacco protein was compared with structural models of *Arabidopsis* 4CL2 and of the CoA ligase from *Scutellaria baicalensis*. The substrate range of the latter enzyme is restricted to unsubstituted cinnamic acid **1** only.[19] The aromatic acid substrates are bound in a cavity which is conserved in the three homologs (Fig. 2). Interestingly, the binding sites of the individual enzymes rather differ in amino acid composition than in size (relative volume of the cavity in the enzymes from *A. thaliana*/tobacco/*S. baica-lensis*: 1/1.03/1.17). In the *Arabidopsis* enzyme, the phenolic side chain of the carboxylic acid substrate seems to be bound via a combination of a serine, glutamine and tyrosine residue (S257, Y253, Q227, see Fig. 2).[14a] These amino acids are also present in the tobacco enzyme and other broad-spectrum CoA ligases in homologous positions (see Fig. S4 in the Supporting Information). The binding cavity of *S. baicalensis* contrasts by being more hydrophobic, with the respective residues being replaced by valine, phenylalanine and isoleucine. This hydrophobization seems to be responsible for the enzyme's selectivity towards unsubstituted cinnamic acid (**1**), excluding more hydrophilic carboxylic acids such as coumarates.[19] As reported in previous studies, subtle differences in the active sites strongly influence the reactivity of 4CL enzymes[18] – a fact which suggests that both rational engineering and activity screening of yet uncharacterized proteins will facilitate access to tailored ligase biocatalysts in the future.

To confirm formation of the thioesters in the spectrophotometric assays mentioned above, preparative conversions of the cinnamic acid derivatives **1** – **6**, **8** – **15**, and **18** were performed. The one-pot synthetic reactions were incubated for a prolonged reaction time (20 hours) and contained a 7.5-fold excess of the acid substrate to achieve complete consumption of the added CoA (2.65 µmol). The CoA esters were subsequently separated from other components present in the samples in a single simple purification step. The crude reaction mixtures containing the thioester product were applied to solid phase extraction (SPE) cartridges and buffer salts, the

co-product adenosine-5´-mono-phosphate and 4CL2 were removed by a washing step. Pure CoA esters were recovered from the SPE matrix in excellent purity (see Supporting Information) and in good to moderate yield (Scheme 1b). The identity of the products was confirmed by HR-MS and NMR. Formation of side products was not observed in any of the enzymatic reactions.

A distinguished advantage of the incorporation of CoA ligases in enzymatic multistep reactions is the fact that their labile reaction products will be directly transformed into follow-up products without the need of purification or the risk of hydrolysis.[20] Even combinatorial approaches, i.e. precursor-directed biosyntheses using e.g. non-natural substrates, are possible if the consecutive steps are catalyzed by promiscuous biocatalysts. In order to demonstrate the feasibility of such a cell-free system in a proof-of-concept study, we combined 4CL2 with feruloyl-CoA hydratase/lyase (FCoAHL, E.C. 4.2.1.101), a bacterial hydratase which catalyzes the Michael addition of water to the acrylate moiety of feruloyl-CoA. The intermediate β-hydroxypropionic thioester then undergoes retro-aldol reaction to vanillin (structure depicted in Fig. 1a) and the coupled by-product acetyl-CoA.[21] Vanillin is produced when FCoAHL is added to a 4CL2 reaction containing ferulic acid **10** (Fig. 1a). As the hydratase showed activity not only on naturally occurring CoA esters (e.g. 4-coumaroyl- and feruloyl-CoA) but also towards artificial derivatives (see Fig. S2 in Supporting Information), we exploited the enzyme cascade for direct synthesis of non-natural benzaldehydes from the corresponding cinnamic acids. Indeed, the formation of unprotected ethylvanillin – an important artificial aroma compound – was achieved in samples con-taining **12** as starting material via homoferulate-CoA (Fig. 1b).

In conclusion, broad-spectrum ligases like 4CL2 are indispensable tools in the synthesis of activated carboxylates, which are key intermediates in natural or artificial pathways to various valuable naturally produced products. In subsequent studies, the reaction will be extended to multi-enzyme sequences for the biosynthesis of non-natural products by adding compatible catalysts such as cyclizing enzymes (e.g. polyketide synthases) or oxidases (e.g. coumarin biosynthetic enzymes). Moreover, such cascade reactions can be used either for structural diversification of intermediates and products (if the reactions include promiscuous downstream enzymes) or in the screening for biocatalysts acting on CoA conjugates.

Footnotes

[#] Saturated fatty acids having an even number (2 – 16) of carbon atoms were used in the activity assays.

* This strict selectivity is probably caused by steric hindrance which prevents occupation of long-chain fatty acids in the active site. The calculated length of a hexanoic acid molecule (8.90 Å) is similar to that of the natural substrate 4-coumaric acid (**2**) (10.0 Å).

Experimental Section

All reagents were of the highest quality available and were purchased from Sigma-Aldrich (St. Louis, USA) (cinnamic acids, Ellman´s reagent) or Applichem (Darmstadt, Germany) (CoA). 4CL2 and FCoAHL were produced as recombinant proteins (as *N*-terminal fusion to maltose-binding protein / with *C*-terminal decahistidine tag (4CL2), or with an *N*-terminal hexahistidine tag (FCoAHL)) in *E. coli* and were purified by metal affinity chromatography. For a detailed description of experimental procedures see Section "Supporting methods" in the Supporting Information.

Acknowledgements

We thank Anja Ehrlich and Julia Christke for excellent technical assistance, and Annegret Laub and Dr. Andrej Frolov (IPB Halle) for their help with the MS/MS measurements. Parts of the research leading to these results have received funding from the European Union´s Seventh Framework Program FP7/2007-2013 under grant agreement nº 266025 (BIONEXGEN, MD), from the BMBF (P34 project Biokatalyse 2021, AKB), from the Leibniz Research Cluster "Bio/synthetic multifunctional micro production units", and from the ScienceCampus Halle - Plant-Based Bioeconomy, cofinanced by the Land Sachsen-Anhalt and the Leibniz Association.

References

[1] R. Leonardi, Y. M. Zhang, C. O. Rock, S. Jackowski, *Prog. Lipid Res*. **2005**, *44*, 125-153.

[2] S. M. Houten, R. J. A. Wanders, *J. Inherit. Metab. Dis.* **2010**, *33*, 469-477.

[3] a) J. M. Berg, J. L. Tymoczko, L. Stryer, *in Biochemistry, 5th ed.* (Eds.: J.M. Berg, J.L. Tymoczko, L. Stryer), W. H. Freeman, New York, **2002**, pp. 501-527; b) D. Voet, J. G. Voet, *in Biochemistry, 4th ed.*, John Wiley & Sons, Hoboken, **2011**, pp. 789-823.

[4] W. Boerjan, J. Ralph, M. Baucher, *Annu. Rev. Plant Biol.* **2003**, *54*, 519-46.

[5] K. Kai, M. Mizutani, N. Kawamura, R. Yamamoto, M. Tamai, H. Yamaguchi, *Plant J.* **2008**, *55*, 989-999.

[6] a) S. M. Houten, R. J. A. Wanders, *J. Inherit. Metab. Dis.* **2010**, *33*, 469-477; b) M. D. Resh, *Biochim. Biophys. Acta - Mol. Cell Res*. **1999**, *1451*, 1-16.

[7] S. Ma, L. B. Pickens, Y. Tang, *in Encyclopedia of Industrial Biotechnology, 1st ed.* (Eds.: M. C. Flickinger, S. W. Drew, R. E. Spier), John Wiley &

Sons, Hoboken, **2009**, pp. 1-11.
[8] a) R. B. Herbert, *Nat. Prod. Rep*. **1999**, *16*, 199-208; b) A. J. Humphrey, D. O'Hagan, *Nat. Prod. Rep*. **2001**, *18*, 494-502; c) K. Springob, R. Lukačin, C. Ernwein, I. Gröning, U. Matern, *Eur. J. Biochem.* **2000**, *267*, 6552-6559; d) M. F. Roberts, D. Strack, M. Wink, *in Annual Plant Reviews, vol. 40* (Ed.: M. Wink), Wiley-VCH, Weinheim, **2008**, pp. 20-91. e) M. Petersen, J. Hans, U. Matern, *in Ann. Plant Rev., vol. 40* (Ed.: M. Wink), Wiley-VCH, Weinheim, **2008**, pp. 182-257.
[9] T. Vogt, *Mol. Plant* **2010**, *3*, 2–20.
[10] V. Agarwal, S. Diethelm, L. Ray, N. Garg, T. Awakawa, P. C. Dorrestein, B.S. Moore, *Org. Lett.* **2015**, *17*, 4452-4455.
[11] a) P. K. Mishra, D. G. Drueckhammer, *Chem. Rev.* **2000**, *100*, 3283-3310; b) J. Stöckigt, M. H. Zenk, *Z. Naturforschung. C, A. J. Biosci*. **1975**, *30*, 352-358.
[12] a) J. E. Bishop, A. K. Hajra, *Anal. Biochem.* **1980**, *106*, 344-350; b) A. Kawaguchi, T. Yoshimura, S. Okuda, *J. Biochem.* **1981**, *89*, 337-339.
[13] a) J. Ehlting, D. Büttner, Q. Wang, C. J. Douglas, I. E. Somssich, E. Kombrink, *Plant J.* **1999**, *19*, 9-20; b) M. Pietrowska-Borek, H.-P. Stuible, E. Kombrink, A. Guranowski, *Plant Physiol.* **2003**, *131*, 1401-1410.
[14] a) K. Schneider, K. Hövel, K. Witzel, B. Hamberger, D. Schomburg, E. Kombrink, H.-P. Stuible, *Proc. Natl. Acad. Sci. USA*. **2003**, *100*, 8601-8606; b) H. P. Stuible, E. Kombrink, *J. Biol. Chem*. **2001**, *276*, 26893-26897.
[15] Y. Hu, Y. Gai, L. Yin, X. Wang, C. Feng, L. Feng, D. Li, X.-N. Jiang, D.-C. Wang, *Plant Cell* **2010**, *22*, 3093-3104.
[16] M. M. Gaid, H. Scharnhop, H. Ramadan, T. Beuerle, L. Beerhues, *J. Plant Physiol*. **2011**, *168*, 944-951.
[17] C.-H. Wang, J. Yu, Y. X. Cai, P.-P. Zhu, C.-Y. Liu, A.-C. Zhao, R.-H. Lü, M.-J. Li, F.-X. Xu, M.-D. Yu, PLoS ONE 2016, 11, e0155814.
[18] Z. Li, S. K. Nair, *Structure* **2015**, *23*, 2032-2042.
[19] Q. Zhao,Y. Zhang, G. Wang, L. Hill, J. K. Weng, X.Y. Chen, H. Xue, C. Martin, *Plant Biochem*. **2016**, *2*, e1501780.
[20] J. M. Luengo, *J. Antibiot. (Tokyo)* **1995**, *48*, 1195-1211.
[21] A. Mitra, Y. Kitamura, M. J. Gasson, A. Narbad, A. J. Parr, J. Payne, M. J. Rhodes, C. Sewter, N. J. Walton, *Arch. Biochem. Biophys*. **1999**, *365*, 10-16.

10.1002/adsc.201900892

COMMUNICATION

Coenzyme A-Conjugated Cinnamic Acids – Enzymatic Synthesis of a CoA-Ester Library and Application in Biocatalytic Cascades to Vanillin Derivatives

Adv. Synth. Catal. **Year**, *Volume*, Page – Page

Martin Dippe, Anne-Katrin Bauer, Andrea Porzel, Evelyn Funke, Anna O. Müller, Jürgen Schmidt, Maria Beier and Ludger A. Wessjohann *