Biochemical characterization of acyl activating enzymes for side chain moieties of Taxol and its analogs

Received for publication, January 22, 2020, and in revised form, February 14, 2020 Published, Papers in Press, February 20, 2020, DOI 10.1074/jbc.RA120.012663

Narayanan Srividya, Iris Lange, Michael Hartmann, Qunrui Li, Maryam Mirzaei, and Bernd Markus Lange

From the Institute of Biological Chemistry and M. J. Murdock Metabolomics Laboratory, Washington State University, Pullman, Washington 99164-6340

Edited by Joseph M. Jez

Taxol (paclitaxel) is a very widely used anticancer drug, but its commercial sources mainly consist of stripped bark or suspension cultures of members of the plant genus Taxus. Taxol accumulates as part of a complex mixture of chemical analogs, termed taxoids, which complicates its production in pure form, highlighting the need for metabolic engineering approaches for high-level Taxol production in cell cultures or microbial hosts. Here, we report on the characterization of acyl-activating enzymes (AAEs) that catalyze the formation of CoA esters of different organic acids relevant for the N-substitution of the 3-phenylisoserine side chain of taxoids. On the basis of similarities to AAE genes of known function from other organisms, we identified candidate genes in publicly available transcriptome data sets obtained with Taxus × media. We cloned 17 AAE genes, expressed them heterologously in Escherichia coli, purified the corresponding recombinant enzymes, and performed in vitro assays with 27 organic acids as potential substrates. We identified TmAAE1 and TmAAE5 as the most efficient enzymes for the activation of butyric acid (Taxol D side chain), TmAAE13 as the best candidate for generating a CoA ester of ticlic acid (Taxol B side chain), TmAAE3 and TmAAE13 as suitable for the activation of 4-methylbutyric acid (N-debenzoyl-N-(2-methylbutyryl)taxol side chain), TmAAE15 as a highly efficient candidate for hexanoic acid activation (Taxol C side chain), and TmAAE4 as suitable candidate for esterification of benzoic acid with CoA (Taxol side chain). This study lays important groundwork for metabolic engineering efforts aimed at improving Taxol production in cell cultures.

This work was supported by the National Institutes of Health Grants RC2GM092561-01 and R01AT010593-01 and McIntire–Stennis Capacity Grant WNP 00811 (to B. M. L.). The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This article contains Tables S1, S2, and S3 and Figs. S1–S3.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) MN961504, MN961505, MN961506, MN961507, MN961508, MN961509, MN961510, MN961511, MN961512, MN961513, MN961514, MN961515, MN961516, MN961517, MN961518, MN961519, MN961520, MN961521, and MN961522.

1Present address: Dept. of Molecular Ecophysiology of Plants, University of Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany.

2Present address: CP Pharmaceutical Group, Jia 3 Yonggangdongli, Jiangguoneuai Daijie, Beijing, China 100022.

3Present address: Tarbiat Modares University, 14115–111 Tehran, Iran.

4To whom correspondence should be addressed: Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340. Tel.: 509-335-3794; Fax: 509-335-7643; E-mail: lange-m@wsu.edu.

Paclitaxel (Taxol® or Taxol A) was originally extracted from the bark of the Pacific yew tree (Taxus brevifolia Nutt.) as part of a program, performed by the U.S. Department of Agriculture under contract to the National Cancer Institute in the 1960s, that was focused on the discovery of novel chemostatic active principles from plants (1). The isolated compound was subsequently demonstrated, during the 1970s, to be highly effective for the treatment of breast and ovarian cancers and was formulated into a drug that received approval from the Food and Drug Administration for the treatment of ovarian cancer in 1992 (2). It later became one of the best-selling anti-cancer drugs of all time. Because Taxol is only accumulated in specialized phloem parenchyma cells of the bark (3), the amounts obtainable from stripped bark of the yew tree were extremely low, and large numbers of this slow-growing tree species had to be sacrificed to collect sufficient amounts of Taxol for clinical trials (4).

Research initiatives designed to find alternative strategies demonstrated the viability of extracting abundant 10-deacetylba-catin III from needles of the European yew (Taxus baccata L.), followed by a four-step chemical synthesis to the desired end product (5). This semisynthetic process was used commercially by the pharmaceutical company Bristol–Myers Squibb for decades but has been phased out because of the emergence of an even more cost- and resource-effective process developed by Phyton Biotech (now part of DBF Pharmaceuticals), which utilizes Taxus cell suspension cultures as a direct source of Taxol (6). Despite the above-mentioned successes with Taxus cell cultures, the yield of Taxol is still fairly low, and the target metabolite needs to be separated from abundant off-pathway taxoids (7, 8). It would thus be desirable to employ metabolic engineering approaches for the high-level Taxol production in Taxus cell cultures or microbial hosts. However, such efforts are currently limited by an insufficient understanding of the highly branched biosynthetic pathways leading toward Taxol and other taxoids (9). One important issue pertains to the activation of different precursors for the N-substitution of the 3-phenylisoserine side chain of taxoids. Taxus cell suspension cultures have been reported to produce Taxol (containing a benzoyl moiety in the side chain) at 6.8%, Taxol B (tigloyl moiety) at 2.9%, Taxol C (hexanoyl moiety) at 6.5%, Taxol D (butanoyl moiety) at 7.5%, and N-debenzoyl-N-(2-methylbutyryl)taxol (2-methylbutyryl moiety) at 1.3% of the total taxoids (8) (Fig. 1). Although these values vary substantially depending on culture conditions (10), it is clear that significant flux is diverted from the desired end product (Taxol) to other structurally related...
Acyl-activating enzymes for Taxol biosynthesis

Figure 1. Side chain moieties of Taxol and its analogs. A, baccatin III scaffold with 3-phenylisoserine side chain (on gray background) at C-13. The position where the side chain is further modified is indicated by an arrow. B, organic acid residues incorporated into Taxol and its analogs.

taxoids. From a metabolic engineering point of view, it would thus be desirable to up-regulate the acyl activating enzymes (AAEs)\(^5\) responsible for conjugating CoA with benzoic acid (required for Taxol biosynthesis) and/or down-regulate the AAEs for the conjugation of CoA with other aromatic or aliphatic acids (leading to other taxoids).

Several AAEs with relevance for the activation of organic acids with relevance for the biosynthesis of functionalized terpenoids or related plant natural products have been described. Enzyme activities capable of activating benzoic acid were reported to occur in *Clarkia breweri*, *Petunia hybrida*, and *Arabidopsis thaliana* (11–13). Hexanoic acid was demonstrated to be activated by an AAE from *Cannabis sativa* and branched-chain organic acids were converted to the corresponding CoA esters by AAEs from *Humulus lupulus* (14, 15). To begin to address organic acid activation potentially related to Taxol biosynthesis, we searched publicly available data sets obtained with methyl jasmonate-treated *Taxus × media* cv. *Hicksii* cell suspension cultures for putative orthologs of these previously characterized AAEs. Candidate AAE genes were then cloned from *T. media*, expressed in *Escherichia coli*, the corresponding recombinant proteins were purified, and activities were determined in *in vitro* assays against a wide range of different organic acid substrates. This approach resulted in the identification and functional characterization of the AAEs required for the activation of side chain precursors involved in the biosynthesis of all major taxoids with an \(N\)-substituted 3-phenylisoserine side chain. Our discoveries can therefore serve as a valuable resource for the development of metabolic engineering strategies for Taxol production.

Results

**Cloning and categorization of candidate AAEs**

We previously performed metabolite analyses as part of a capacity-building project for medicinal plants, which also included samples from *T. media* cell suspension cultures (16). Because these cultures were available to us as a source for candidate genes, we downloaded publicly available whole transcriptome shotgun sequencing (RNA-Seq) data sets acquired with this species (National Center for Biotechnology Information Bioproject numbers PRNJA169654, PRJNA341288, and PRNJA497542) (17, 18), with an emphasis on suspension cells treated with the elicitor methyl jasmonate, which is commonly used to up-regulate the formation of taxoids (19). We then searched in these publicly available data sets for gene sequences with homology to previously characterized AAEs of varying specificity (for details, see “Experimental procedures”). Using this approach, a total of 19 transcripts were initially selected for further evaluation (Fig. S1). The sequences of the corresponding genes were translated to peptide sequence and compared with previously characterized AAEs from plants. A phylogenetic comparison indicated that the *T. media* candidates fell into the AAE clades I, IV, V, VI, and VII, as previously defined for AAEs of *A. thaliana* (20) (Fig. 2 and Fig. S2). Two of the *T. media* candidates (TmLACS1 and TmLACS2) clustered with long-chain fatty acyl-CoA synthetases (LACSs) (clade I) involved in lipid biosynthesis (AtLACS8 and AtLACS9; localized to the ER and plastids, respectively) or fatty acid \(β\)-oxidation (AtLACS6 and AtLACS7); localized to peroxisomes) (21, 22). Because these functions are unrelated to taxoid biosynthesis, a further characterization of TmLACS1 and TmLACS2 was not attempted. Four candidates (TmAAE7, TmAAE8, TmAAE9, and TmAAE10) clustered with previously characterized malonate-CoA ligase (AtAAE13) (23) or oxaloate-CoA ligase (AtAAE3) (24) of clade VII. Although these activities were not of direct relevance for taxoid biosynthesis, these candidates could be promiscuous (given that the enzymes accept small dicarboxylic acids), and we therefore decided to include them in functional assays. Functional assays indeed demonstrated that oxalic acid (TmAAE9) and malonic acid (TmAAE7 and TmAAE8) were converted to the corresponding CoA esters (Fig. S3). TmAAE10 was not active under these assay conditions.

**Separation and identification of AAE assay products**

The AAE-mediated conjugation of various organic acid substrates with CoA was anticipated to result in the formation of thioesters for which authentic standards were not readily available from commercial sources. It was therefore imperative to develop a protocol for the separation and identification of enzyme assay products. We obtained cDNA clones for three previously characterized AAEs that act on

\(^5\) The abbreviations used are: AAE, acyl activation enzyme; HPLC-QTOF-MS, high-performance liquid chromatography–quadrupole time-of-flight–mass spectrometry; LACS, long-chain fatty acyl-CoA synthetase.
substantially different substrates (4-coumaric acid, *A. thaliana* At1g51680; *β*-phenylalanine, *Penicillium chrysogenum* Pc21g30650 (A312G mutant); and 12-oxo-phytodienoic acid, *A. thaliana* At5g63380) (25–27), expressed these genes in *E. coli*, purified the corresponding recombinant proteins, performed enzyme assays with appropriate substrates, and tested conditions for subsequent analysis by high-performance liquid chromatography–quadrupole time-of-flight–mass spectrometry (HPLC-QTOF-MS). Separation of the enzyme assay substrates and products was achieved on a reversed-phase column with a linear gradient (for details, see “Experimental procedures”). The assay co-substrates ATP and CoA eluted at 3.3 and 3.6 min, respectively, whereas the co-product AMP was detected at 2.9 min (Fig. 3, A–C). The quasi-
molecular ions for the CoA ester products, detected at high mass accuracy (mass error generally below 2 ppm), enabled the calculation of molecular formulas that were consistent with expectations (Fig. 3, D–F). The fragmentation patterns of enzymatically produced CoA ester products, obtained by MS/MS, were also consistent with those reported in the literature (26, 28, 29) (Fig. 3, G–I), indicating that our analytical platform was suitable for the analysis of CoA ester conjugates of a wider range of organic acids.

**Screening candidate AAEs for substrate preference**

Open reading frames corresponding to the candidate AAEs from *T. media* were amplified by PCR from first-strand cDNA that represents transcripts expressed in methyl jasmonate-treated suspension cells. The resulting amplicons were subcloned separately into the pET32b expression vector, and these plasmids transformed into *E. coli* as a heterologous host. Recombinant proteins produced after inducing transgene expression in transformed *E. coli* cells were purified by Ni²⁺ affinity chromatography and enzymes assayed with 27 different organic acids, including aliphatic organic acids of different chain lengths and branching patterns, various aromatic organic acids, and aromatic amino acid derivatives, to assess substrate utilization. The thioesters formed by enzymatic catalysis were identified by HPLC-QTOF-MS (Table 1 and Table S1). In addition, the depletion of the co-substrate ATP and emergence of AMP as a co-product (Fig. 3B) were monitored to quantify conversion rates of AAEs.

Of the 17 candidate AAEs from *T. media* to be further characterized, five were highly active (> 10 nmol h⁻¹ mg protein⁻¹ with at least one substrate; TmAAE3, TmAAE5, TmAAE11, TmAAE12, and TmAAE15), four had activities in the medium high range (2–8 nmol h⁻¹ mg protein⁻¹ with at least one substrate; TmAAE1, TmAAE2, TmAAE13, and TmAAE16), four were poorly active (0.03–0.7 nmol h⁻¹ mg protein⁻¹ with at least one substrate; TmAAE4, TmAAE7, TmAAE8, and TmAAE10), and four had no detectable activity with any of the...
Acyl-activating enzymes for Taxol biosynthesis

Table 1
HPLC-QTOF-MS properties of CoA ester products formed in enzyme assays with T. media AAE candidates (ordered by $R_{t}$ of product)

| Substrate                  | $R_{t}$ | Molecular formula               | Molar mass | [M-H]− | Theoretical mass | Experimental mass | Mass error | ppm |
|----------------------------|---------|---------------------------------|------------|--------|-----------------|------------------|------------|-----|
| Phenylalanine              | 8.1     | C$_{9}$H$_{8}$N$_{2}$O$_{3}$P$_{5}$S | 914.18     | 913.18 | 913.177         | 3.395            | 713.18     |     |
| Propionic acid             | 9.7     | C$_{3}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 823.141    | 822.13 | 822.1341        | 0.608            | 822.13     |     |
| Crotonic acid              | 9.7     | C$_{3}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 834.13     | 834.13 | 834.1341        | 0.599            | 834.13     |     |
| Caffeic acid               | 12.8    | C$_{9}$H$_{8}$N$_{2}$O$_{3}$P$_{5}$S | 929.147    | 928.14 | 928.1369        | 2.370            | 928.14     |     |
| Butyric acid               | 13.2    | C$_{3}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 837.157    | 836.15 | 836.1501        | 0.957            | 836.15     |     |
| Benzoic acid               | 15.0    | C$_{7}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 871.141    | 870.13 | 870.1339        | 0.345            | 870.13     |     |
| Tiglic acid                | 15.1    | C$_{3}$H$_{6}$O$_{10}$P$_{5}$S   | 849.157    | 848.15 | 848.1507        | 1.651            | 848.15     |     |
| 4-Coumaric acid            | 15.7    | C$_{7}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 913.152    | 912.14 | 912.1439        | 0.328            | 912.14     |     |
| Valeric acid               | 16.0    | C$_{3}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 851.173    | 850.16 | 850.1630        | 2.235            | 850.16     |     |
| 2-Methylbutyric acid       | 16.2    | C$_{3}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 851.173    | 850.16 | 850.1645        | 0.470            | 850.16     |     |
| Furalic acid               | 16.4    | C$_{3}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 943.163    | 942.15 | 942.1524        | 2.550            | 942.15     |     |
| trans-Cinnamic acid        | 21.6    | C$_{7}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 897.157    | 896.15 | 896.1471        | 2.455            | 896.15     |     |
| Hexanoic acid              | 21.8    | C$_{3}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 865.188    | 864.18 | 864.1802        | 0.463            | 864.18     |     |
| Heptanoic acid             | 28.4    | C$_{7}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 879.204    | 878.2  | 878.1966        | 0.455            | 878.2      |     |
| Octanoic acid              | 30.5    | C$_{7}$H$_{6}$N$_{2}$O$_{3}$P$_{5}$S | 892.230    | 891.21 | 891.2123        | 0.448            | 891.21     |     |
| Decanoic acid              | 31.2    | C$_{11}$H$_{18}$N$_{2}$O$_{3}$P$_{5}$S | 921.251    | 920.24 | 920.2449        | 1.847            | 920.24     |     |

Proffered substrates (TmAAE6, TmAAE9, TmAAE14, and TmAAE17) (Table 2 and Table 3).

High specific activities for short-chain and branched chain organic acids were observed with TmAAE11 (preference for butyric, propionic, and valeric acids), TmAAE3 (preference for 2-methylbutyric acid), TmAAE5 (active with almost all proffered short-chain and branched-chain organic acids), TmAAE2 (preference for valeric, butyric, and propionic acids), TmAAE7 (preference for 2-methylbutyric and propionic acids), TmAAE1 (preference for butyric and propionic acid), and TmAAE13 (preference for valeric and crotonic acids but active with all proffered short-chain and branched-chain organic acids) (Table 2).

Additional activities with medium-chain fatty acids were detected with four of these candidates (TmAAE11, TmAAE3, TmAAE2, and TmAAE13). Activities with aromatic and/or phenylpropanoid acids as substrates were observed with five of the above-mentioned candidates (TmAAE5, TmAAE13, TmAAE11, TmAAE3, and TmAAE2), thus leaving TmAAE1 as the only apparent candidate with high selectivity for short-chain and branched organic acids (Table 2).

High specific activities with hexanoic acid were detected for three candidates (TmAAE13, TmAAE15, and TmAAE12) (Table 2). All of these candidates also accepted other medium-chain fatty acids, aromatic acids, and phenylpropanoid acids as substrates. TmAAE8 was the only candidate that appeared to be selective for medium chain fatty acids, but it had only a very low specific activity with lauric acid as a substrate (0.03 nmol h$^{-1}$ mg protein$^{-1}$). Expectedly, based on sequence identity scores when compared with AAEs with established functions, none of the candidates investigated here were active with long-chain fatty acids (Table 2).

Benzoic acid was converted to the corresponding CoA ester with high activity by TmAAE5, TmAAE13, and TmAAE15, but as mentioned above, these candidates were also active with a range of other substrates, including short-chain, branched-chain, medium-chain, and phenylpropanoid acids (Table 2). TmAAE4 appeared to be the only candidate with high specificity for benzoic acid, and TmAAE10 was active only with phenylpropanoid acids.

Kinetic characterization of T. media AAEs with potential relevance for the biosynthesis of Taxol and its analogs

Kinetic analyses were then performed with seven AAEs that were selected because they had a particularly high specific activity and/or apparent selectivity for the CoA ester-based activation of specific organic acids with relevance for the decoration of the 3-phenylisoserine side chain of taxoids. When assayed with butyric acid as a substrate (required for Taxol D side chain), significant differences in binding affinities were determined, as indicated by $K_{m}$ values of 68 µM for TmAAE5, 81 µM for TmAAE3, 86 µM for TmAAE1, and 483 µM for TmAAE13 (value too high for accurate determination with TmAAE16 and TmAAE15) (Table 3). The $k_{cat}$ values for TmAAE1, TmAAE3, TmAAE5, and TmAAE13 were relatively low (1.06, 0.68, 0.88, and 0.11 s$^{-1}$, respectively), but the catalytic efficiency for this substrate was comparatively high for TmAAE1 and TmAAE5 ($k_{cat}/K_{m} = 12,357$ and $12,885$ M$^{-1}$ s$^{-1}$, respectively).

Appreciable binding affinities for tiglic acid (Taxol B side chain) were determined for TmAAE13 and TmAAE1 ($K_{m}$ values of 85 and 99 µM, respectively), with TmAAE5 exerting a much lower affinity toward this substrate ($K_{m} = 368$ µM) (other selected AAEs had no detectable activity) (Table 3). The turnover numbers with tiglic acid were low for TmAAE5 and TmAAE1 ($k_{cat}$ of 0.71 and 0.19 s$^{-1}$, respectively) and slightly higher for TmAAE13 ($k_{cat} = 1.75$ s$^{-1}$). Only TmAAE13 had a high catalytic efficiency with this substrate ($k_{cat}/K_{m} = 20,615$ M$^{-1}$ s$^{-1}$).

Productive binding of 4-methylbutyric acid (N-debenzoyl-N-(2-methylbutyl)taxol side chain) was observed with TmAAE3 and TmAAE13 ($K_{m}$ values of 78 and 140 µM, respectively), with fairly low $k_{cat}$ values (0.56 and 0.49 s$^{-1}$, respectively) and a decent catalytic efficiency only for TmAAE3 ($k_{cat}/K_{m} = 7,131$ M$^{-1}$ s$^{-1}$) (very low binding affinity for all other candidates) (Table 3).

MS/MS data are provided in Table 3s.
### Table 2

Specific activity of *T. media* CoA ligase candidates (expressed as nmol h\(^{-1}\) mg protein\(^{-1}\)) with various organic acid substrates

The data are included for candidate AAEs that were shown to have activity with at least one substrate. Full data are available in Table S2. The standard deviation given is based on *n* = 3. NQ, not quantifiable.

| Substrate                  | TmAEE1 | TmAEE2 | TmAEE3 | TmAEE4 | TmAEE5 | TmAEE7 | TmAEE8 | TmAEE10 | TmAEE11 | TmAEE12 | TmAEE13 | TmAEE15 | TmAEE16 |
|----------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Propionic acid (C3, saturated) | 1.48 ± 0.24 | 1.79 ± 0.16 | 1.18 ± 0.07 | NQ | 13.37 ± 1.74 | 0.24 ± 0.05 | NQ | NQ | 6.06 ± 1.40 | 0.38 ± 0.05 | 0.14 ± 0.06 | NQ | 0.14 ± 0.01 |
| Butyric acid (C4, saturated) | 2.32 ± 0.27 | 1.84 ± 1.18 | 0.31 ± 0.14 | NQ | 6.51 ± 1.10 | 0.04 ± 0.01 | NQ | NQ | 10.01 ± 2.13 | 0.41 ± 0.05 | 0.83 ± 0.06 | NQ | 0.78 ± 0.22 |
| Crotonic acid (C5, unsaturated) | 0.12 ± 0.02 | NQ | 0.56 ± 0.09 | NQ | 4.02 ± 0.96 | 0.16 ± 0.04 | NQ | NQ | NQ | NQ | 5.69 ± 2.76 | 5.11 ± 2.19 | 1.63 ± 0.11 |
| Fumaric acid (C4, unsaturated, diac) | NQ | NQ | 0.04 ± 0.01 | NQ | NQ | NQ | NQ | NQ | NQ | 0.17 ± 0.04 | NQ | 0.07 ± 0.04 |
| Valeric acid (C5, saturated) | 0.06 ± 0.01 | 2.15 ± 0.42 | 0.07 ± 0.01 | NQ | 7.55 ± 0.06 | NQ | NQ | NQ | 2.38 ± 0.06 | 0.06 ± 0.01 | 7.16 ± 0.91 | 0.20 ± 0.09 | 1.78 ± 0.47 |
| 2-Methylbutyric acid (C5, branched, saturated) | NQ | 0.61 ± 0.24 | 12.30 ± 1.70 | NQ | 17.18 ± 0.57 | 0.64 ± 0.05 | NQ | NQ | 0.55 ± 0.22 | NQ | 0.24 ± 0.04 | NQ | 0.13 ± 0.01 |
| Tigliic acid (C5, branched, unsaturated) | 0.69 ± 0.01 | NQ | 0.37 ± 0.15 | NQ | 10.87 ± 0.57 | NQ | NQ | NQ | NQ | 1.3 ± 0.29 | NQ | NQ |
| Hexanoic acid (6:0) | NQ | 0.30 ± 0.02 | 0.68 ± 0.2 | NQ | NQ | NQ | NQ | NQ | 2.19 ± 0.22 | 4.87 ± 0.70 | 4.88 ± 1.20 | 15.00 ± 4.34 | 1.18 ± 0.15 |
| Heptanoic acid (7:0) | NQ | 0.31 ± 0.21 | 1.72 ± 0.24 | NQ | NQ | NQ | NQ | NQ | 2.48 ± 0.22 | 10.39 ± 0.82 | 6.06 ± 2.94 | 2.75 ± 0.84 | 2.69 ± 1.31 |
| Octanoic acid (8:0) | NQ | NQ | 3.37 ± 1.27 | NQ | NQ | NQ | NQ | NQ | 0.19 ± 0.07 | 1.04 ± 0.06 | NQ | 4.42 ± 1.33 | 1.55 ± 0.35 |
| Decanoic acid (10:0) | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | 5.86 ± 0.72 | NQ | 1.57 ± 0.08 |
| Lauric acid (12:0) | NQ | NQ | NQ | NQ | NQ | 0.03 ± 0.01 | NQ | NQ | NQ | NQ | 0.13 ± 0.03 | 0.53 ± 0.02 | NQ |
| Myristic acid (14:0) | NQ | NQ | NQ | NQ | NQ | 0.06 ± 0.01 | NQ | NQ | NQ | NQ | 0.04 ± 0.01 | 0.34 ± 0.03 | NQ |
| Palmitic acid (16:0) | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ |
| Stearic acid (18:0) | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ |
| Oleic acid (18:1, cis-9) | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ |
| Linoleic acid (18:2, n-6) | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ |
| 12-oxo-Phytodienoic acid | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ |

### Aromatic and phenylpropionic acids

| Substrate                  | TmAEE1 | TmAEE2 | TmAEE3 | TmAEE4 | TmAEE5 | TmAEE7 | TmAEE8 | TmAEE10 | TmAEE11 | TmAEE12 | TmAEE13 | TmAEE15 | TmAEE16 |
|----------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Benzoic acid (C7) | NQ | NQ | NQ | 0.31 ± 0.02 | 11.19 ± 0.46 | NQ | NQ | NQ | NQ | 4.78 ± 0.89 | 22.51 ± 8.92 | NQ |
| Cinnamic acid (C9) | NQ | 0.01 ± 0.01 | 5.80 ± 1.81 | NQ | 0.05 ± 0.01 | NQ | NQ | NQ | 1.62 ± 0.75 | 0.28 ± 0.1 | 1.87 ± 0.43 | 0.32 ± 0.03 | 0.98 ± 0.12 |
| 4-Coumaric acid (C9) | NQ | 0.54 ± 0.15 | 2.86 ± 1.51 | NQ | NQ | 0.18 ± 0.03 | NQ | 0.67 ± 0.17 | 1.85 ± 0.76 | 1.65 ± 0.45 | 2.10 ± 0.27 | NQ | 2.14 ± 0.61 |
| Caffeic acid (C9) | NQ | 2.34 ± 0.55 | NQ | NQ | 0.32 ± 0.01 | 0.28 ± 0.01 | NQ | NQ | NQ | 0.55 ± 0.31 | 0.57 ± 0.09 | 6.78 ± 1.42 | 1.02 ± 0.18 | 3.70 ± 0.47 | 0.10 ± 0.01 |
| Ferulic acid (C9) | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | 0.17 ± 0.01 | 0.18 ± 0.04 | 2.60 ± 0.88 | 0.14 ± 0.07 | 4.65 ± 1.94 | 1.83 ± 0.25 |

### Aromatic amino acid derivatives

| Substrate                  | TmAEE1 | TmAEE2 | TmAEE3 | TmAEE4 | TmAEE5 | TmAEE7 | TmAEE8 | TmAEE10 | TmAEE11 | TmAEE12 | TmAEE13 | TmAEE15 | TmAEE16 |
|----------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| L-Phenylalanine | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ |
| D-Phenylalanine | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ |
| d-Phenylalanine | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ |
| Phenylisoserine | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ | NQ |
Hexanoic acid (Taxol C side chain) was bound with fairly high affinity by TmAAE16, TmAAE15, and TmAAE13 (K_m values of 38, 190, and 207 μM, respectively). The turnover number with hexanoic acid (349 s⁻¹) and catalytic efficiency (k_cat/K_m = 1,839,557 M⁻¹ s⁻¹) of TmAAE15 were exceptionally high with hexanoic acid (Table 3).

Table 3

| CoA ligase candidate | Butyric acid | 2-Methylbutyric acid | Tiglic acid | Hexanoic acid | Benzoic acid |
|----------------------|--------------|----------------------|-----------|---------------|-------------|
| TmAAE1               | 85.8         | 98.8                 |           | 0.19          |             |
| K_m                  |              |                      |           |               |             |
| k_cat/K_m            | 12,358       |                      | 1931      |               |             |
| TmAAE3               | 81.13        | 77.59                | 928.9     | 0.20          |             |
| K_m                  |              |                      |           |               |             |
| k_cat/K_m            | 8385         | 7131                 | 213       |               |             |
| TmAAE4               |              |                      |           | 172.7         |             |
| K_m                  |              |                      |           | 181.28        |             |
| k_cat/K_m            |              |                      |           | 1,049,678     |             |
| TmAAE5               | 68.3         | 648.0                | 368.1     | 246.6         |             |
| K_m                  |              |                      |           |               |             |
| k_cat/K_m            | 1,049,678    |                      |           |               |             |
| TmAAE13              | 482.7        | 140.2                | 84.9      | 207.2         |             |
| K_m                  |              |                      |           |               |             |
| k_cat/K_m            | 3461         | 20,615               | 10,397    |               |             |
| TmAAE15              |              |                      |           | 190.0         |             |
| K_m                  |              |                      |           | 594.0         |             |
| k_cat/K_m            | 1,839,557    |                      |           |               |             |
| TmAAE16              |              |                      | 37.7      | 53,113        | 2.00        |
| K_m                  |              |                      |           |               |             |
| k_cat/K_m            |              |                      |           |               |             |

The exceptionally high turnover number of TmAAE4 (181 s⁻¹), the catalytic efficiency TmAAE4 with benzoic acid and specificity (reactions of TmAAE4 with other substrates were unproductive), indicates that TmAAE4 would appear to be an excellent candidate for the activation of a side chain aryl acid precursor relevant for Taxol biosynthesis. It should also be noted that a benzoyl moiety is not only contained in the N-benzoyl-(2R,3S)-phenylisoserine side chain of Taxol but is also present in the C2 position of all other taxoids with a baccatin III structural backbone (Fig. 1). Ph-4CL1 from P. hybrida showed low binding affinity for benzoic acid (K_m value > 9 mM), and AtAAE20 (At1g65880) from A. thaliana likely has considerable cinnamate CoA ligase activity (in addition to the activity with benzoic acid as a substrate) (12, 13). A partially purified protein fraction obtained from C. breweri flower extracts was demonstrated to exert CoA ligase activity with benzoic acid (K_m value of 45 μM), but a kinetic characterization with other acyl and aryl acids was not reported (11), and an evaluation of specificity is therefore not possible. Therefore, TmAAE4 seems to be the plant CoA ligase with the highest specificity and catalytic efficiency for benzoic acid as a substrate characterized to date.

**Discussion**

Activation of benzoic acid, which is required for the biosynthesis of the Taxol side chain but also for the decoration of the taxa-4,11-diene core of many other taxoids

Several AAEs converted benzoic acid to the corresponding CoA ester, most notably TmAAE15, TmAAE5, TmAAE13, and TmAAE4. The binding affinity for benzoic acid was low for TmAAE15 and TmAAE13, whereas desirable lower K_m values were determined for TmAAE4 and TmAAE5. Although we recognize the limitations with directly comparing k_cat/K_m ratios (30), the exceptionally high catalytic efficiency TmAAE4 with benzoic acid and specificity (reactions of TmAAE4 with other substrates were unproductive), indicates that TmAAE4 would appear to be an excellent candidate for the activation of a side chain aryl acid precursor relevant for Taxol biosynthesis. It should also be noted that a benzoyl moiety is not only contained in the N-benzoyl-(2R,3S)-phenylisoserine side chain of Taxol but is also present in the C2 position of all other taxoids with a baccatin III structural backbone (Fig. 1). Ph-4CL1 from P. hybrida showed low binding affinity for benzoic acid (K_m value > 9 mM), and AtAAE20 (At1g65880) from A. thaliana likely has considerable cinnamate CoA ligase activity (in addition to the activity with benzoic acid as a substrate) (12, 13). A partially purified protein fraction obtained from C. breweri flower extracts was demonstrated to exert CoA ligase activity with benzoic acid (K_m value of 45 μM), but a kinetic characterization with other acyl and aryl acids was not reported (11), and an evaluation of specificity is therefore not possible. Therefore, TmAAE4 seems to be the plant CoA ligase with the highest specificity and catalytic efficiency for benzoic acid as a substrate characterized to date.

Activation of hexanoic acid, which is required for the biosynthesis of Taxol C

Hexanoic acid, a precursor for the assembly of the side chain of Taxol C (taxuyunnanine), was converted at appreciable levels by seven T. media AAE candidates. A kinetic evaluation indicated that three AAEs had a relatively high binding affinity for hexanoic acid as a substrate, but TmAAE15 had the by far greatest turnover rate and catalytic efficiency. Because this enzyme did not show a significant kinetic preference for any other substrate, TmAAE15 is an excellent candidate for a high efficiency hexanoate:CoA ligase. CsAAE3 from C. sativa, which is involved in the activation of hexanoic acid for the incorporation into the olivetolic acid precursor of cannabinoids, had kinetic values (high binding affinity for hexanoic acid and high turnover rate) comparable with those of TmAAE15 (14), indi-
Acyl-activating enzymes for Taxol biosynthesis

cating a considerable chain-length specificity of these plant CoA ligases.

Activation of short-chain and branched organic acids with relevance for taxoid biosynthesis

Butyric acid (linear) and 2-methylbutyric acid (branched) are saturated short-chain organic acids that need to be activated prior to their integration as side chain moieties of Taxol D (tax-adultin) and N-debenzoyl-N-(2-methylbutyryl)taxol, respectively. A tigloyl moiety, the 2,3-unsaturated derivative of 2-methylbutyric acid, is found in the side chain of Taxol B (cephalomannine). Nine AAE candidates had detectable activities with at least one of these organic acids, of which six were further investigated in kinetic assays. A fairly high binding affinity for butyric acid ($K_a$ value $< 100 \mu M$) was observed with TmAAE1, TmAAE3, and TmAAE5. $k_{cat}$ values for butyric acid were highest for TmAAE1 and TmAAE5. 2-Methylbutyric acid was converted to the corresponding CoA ester with appreciable activation of tiglic acid was catalyzed with high binding affinity and turnover rate by TmAAE3. The CoA-based activation of tiglic acid was catalyzed with high binding affinity and turnover rate by TmAAE13. Only one AAE candidate (TmAAE1) showed relatively high specificity for the activation of butyric acid over all other short-chain organic acids. TmAAE3 had very similar kinetic values for butyric and 2-methylbutyric acid as substrates. Tiglic acid was the preferred substrate converted by TmAAE13 ($k_{cat}/K_m$ value approximately twice as high as with hexanoic acid). AAEs involved in the activation of short-chain, branched organic acids were previously characterized as part of efforts to investigate bitter acid biosynthesis in hop ($H. lupulus$) (15). 3-Methylbutyric acid was converted by HicCL4 with kinetic values similar to those reported here. A second CoA ligase from hop, HicCL2, was productive only with 3-methylbutyric acid (also referred to as isovaleric acid) as a substrate. AtAAE1 and AtAAE2 from $A. thaliana$ were demonstrated to be active with a range of short-chain organic acids, but their specificity was not evaluated in kinetic assays (15). To the best of our knowledge, the current study on $T. media$ AAEs is the first report of an enzyme (TmAAE13) with high catalytic efficiency toward the activation of tiglic acid.

Did we miss an AAE for the activation of $\beta$-phenylalanine?

We tested 17 AAEs from $T. media$ in $in vitro$ assays with 27 organic acid substrates, but none of these candidates showed detectable activity with $\beta$-phenylalanine or 3-phenylisoserine, which are moieties constituting the C-13 side chain of Taxol and its close analogs. The sequence of TmAAE17 is 99% identical to that of a purported $\beta$-phenylalanine-CoA ligase from $T. baccata$, which had been characterized by expressing the corresponding gene in Nicotiana benthamiana as a heterologous host and searching for new products in assays with crude enzyme extracts (31). We therefore performed assays with $P. chrysogenum$ Pc21g30650 (A312G mutant), which had been characterized more thoroughly in kinetic assays with the purified, recombinant protein (26). Our routine assays consistently showed the presence of $\beta$-phenylalaninyl-CoA in assays with this positive control clone, indicating that the experimental conditions were appropriate. More work beyond the scope of the present study will be needed to confirm whether the side chain assembly of Taxol involves an AAE with activity on $\beta$-phenylalanine or possibly another form of activation.

Assessment of $T. media$ AAE sequence properties in the context of catalytic specificity

AAEs are composed of a larger N-terminal domain and a smaller C-terminal domain. Structural studies identified two catalytically relevant conformations—one for the synthesis of an acyl-adenylate intermediate and a second for thioester formation—with a $140^\circ$ domain rotation between these half-reactions (Fig. 4A). The hinge residue that facilitates this movement is usually an aspartate (32), which was conserved in TmAAE1 through TmAAE6 and TmAAE14 (Fig. 4B). The sequences of TmAAE11, TmAAE12, TmAAE13, TmAAE15, TmAAE16, and TmAAE17 had a glutamate in this position, which has occasionally been observed in other AAEs (33) (TmAAE7 through TmAAE10 (clade VII) had activities that were not of direct relevance to taxoid biosynthesis). Analyses of bacterial AAEs indicated the presence of two conserved AMP-binding motifs (AMP1 and AMP2) and a fatty acid CoA-ligase signature (34, 35), which were also present in $T. media$ AAEs.

Interestingly, sequence divergence across $T. media$ AAE sequences indicated that they fell into two groups. The AMP1 motif in TmAAE1 through TmAAE6 (group 1; members of AAE clade VI) was L(L/P)YSSGTTLG(K)/XKGVYXTHXX, whereas in TmAAE11 through TmAAE17 (group 2; members of AAE clades IV and V), there was an L(N/G)YTSQurtlesPK-GV/V/L)X(S/C)HRG motif (Fig. 4B). The AMP2 motif consisted of HXYGL(T/S)ET in group 1 and QGY(L/M)TEX in group 2. Group 1 AAEs had a conserved fatty acid CoA-ligase signature (GW(F/H)(T/S)GD(L/V)(A/G)XHPD-GY/I/L)EIKDR(S/A)K) that was distinguishable from that of group 2 sequences (GW(L/X)TGDX(C/G)Y(I/F)DXXXXXXX/I/V)VDRXK) (Fig. 4B). The A8 motif, which is likely involved in binding both an acyl-adenylate and CoA in AAEs such as the short/medium-chain CoA ligase FadK of $E. coli$ (34), was also found to be divergent in $T. media$ group 1 (D(I/V)ISGGEN(I/V)) and group 2 ((E/D)(L/M)JK(Y/C)(K/N)/(G/A)XQV) sequences. Group-specific divergence was again observed for the C-terminal A10 motif (PKT(S/A)TGK in group 1 sequences and PKXXXGK in group 2 sequences) (Fig. 4B), which contains a strictly conserved lysine that is present at the active site of the ANL superfamily of adenylation enzymes (36). Finally, $T. media$ AAEs contained an A4 motif, featuring a highly conserved histidine residue that caps a helix in the acyl-binding pocket, that was distinguishable in group 1 (H(A/C)NGW(C/S)(F/Y)) and group 2 sequences (H(I/V)(Y/G/S)/(L/F)X) (Fig. 4B).

There was no obvious correlation of an AAE group with sub- strate specificity, indicating that different specificities may have evolved independently from different progenitors, but this hypothesis will need further testing beyond the scope of this study. Some of the AAEs characterized as part of the efforts presented here are remarkably specific and operate at high catalytic efficiency and should thus be an excellent starting point for metabolic engineering efforts. As a next step, it will also be desirable to develop mutant AAEs with enhanced specificity and turnover characteristics, which bears the potential to
increase the yields of highly functionalized taxoid targets (and possibly other natural products with side chains that require activation) (37).

**Experimental procedures**

**Cloning and heterologous expression of candidate AAE genes**

*T. media* cell suspension cultures were maintained, and taxoid accumulation was induced as described previously (38). Total RNA was isolated using the RNeasy plant kit (Qiagen) according to the manufacturer’s instructions. RNA integrity was determined using a BioAnalyzer 2100 (Agilent Technologies, Santa Clara, CA). First-strand cDNA was prepared from RNA with the SuperScript III first-strand synthesis kit (Invitrogen) with random hexamer oligonucleotides. Open reading frames for CoA ligases were amplified using primers (Table S3) designed based on AAE sequences identified in publicly available *T. media* transcriptome data sets (National Center for Biotechnology Information Bioproject numbers PRJNA169654, PRJNA341288, and PRJNA497542) (17, 18). Amplicons were ligated into the pET-32b vector featuring an encoded 6× histidine tag (EMD Millipore, Burlington, MA) and, following sequence confirmation, transformed into *E. coli* BL21(DE3) cells (New England Biolabs, Ipswich, MA). Cells from individual colonies, selected on carbenicillin plates (50 μg/ml), were transferred to 5 ml of LB broth (with the same concentration of antibiotic) and maintained at 37 °C (250 rpm) until the cultures reached an optical density of 0.8. The culture was then diluted to 1 liter with additional LB broth and once again grown under the above-mentioned conditions to an optical density of 0.8. The expression of the target gene was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (final concentration, 0.5 mM), and the cultures were maintained at 16 °C for 16 h.

**Purification and biochemical characterization of recombinant proteins**

The cultures expressing the target gene were transferred to plastic containers, and the cells were precipitated by centrifugation at 3,500 × g and 4 °C for 15 min (Sorvall RC3C; ThermoFisher). The supernatant was removed, and the cell pellet

**Figure 4. Sequence features of *T. media* AAEs.** A, two-step reaction path for the adenylation and thioesterification of organic acid substrates, with conformations adopted by AAEs indicated by different background colors (light gray, adenylate-forming; dark gray, thioester-forming) (TmAAE7 through TmAAE10 are not included in the analysis because these enzymes catalyze CoA transfers to substrates that are not relevant for taxoid biosynthesis). CoA-SH, CoA; PPi, pyrophosphate. B, multiple sequence alignment of residues constituting motifs with suggested functions in AAE catalysis. The conserved aspartate/glutamate residue located at the interdomain hinge position (facilitating a domain rotation caused by a change in the main chain torsional angle) is pointed out by an upward arrow.
Acyl-activating enzymes for Taxol biosynthesis

was resuspended in extraction buffer (100 mM Tris buffer, pH 7.5, containing 20% glycerol, 5 mM DTT, and 1 mM phenylmethylsulfonyl fluoride). The mixture was ground to a fine powder with mortar and pestle under liquid nitrogen and, after allowing the contents to warm up for 5 min at 23 °C, centrifuged at 13,000 × g for 15 min at 4 °C (Sorvall RC3C; ThermoFisher). The supernatant was subjected to a buffer exchange to 50 mM phosphate buffer using Vivaspin centrifugal concentrators (Sartorius, Bohemia, NY) according to the manufacturer’s instructions. Enzyme extracts were then passed through a 0.22-μm cellulose acetate filter and loaded onto a 1-ml nickel column (GE Healthcare) that had been equilibrated with binding buffer (50 mM phosphate buffer containing 20 mM imidazole). After washing the column with 10 bed volumes of binding buffer, the retained protein was eluted with 50 mM phosphate buffer containing 200 mM imidazole. The fractions (1 ml each) were collected, and their protein compositions were compared with that of column flow-through and crude extracts based on separation by denaturing PAGE (SDS-PAGE; staining with Colloidal Blue; ThermoFisher; destaining with 7 (v/v) % acetic acid in 20% aqueous methanol). Fractions corresponding to SDS-PAGE lanes with only one major band of the appropriate size (calculated based on the sequence of the CoA ligase candidate) were desalted over a PD10 column (GE Healthcare) according to the manufacturer’s instructions.

Activity assays were carried out with 50 μg of purified enzyme in a buffer containing 50 mM Tris (pH 7.5), 3 mM ATP, 10 mM CoA, 5 mM MgCl₂, and 2.5 mM substrate (saturating conditions; final volume of 200 μl). The reaction was carried out at 31 °C for 60 min and then stopped by the addition of 2 μg ml⁻¹ solution of 2,4-dichlorophenoxyacetic acid (internal standard for subsequent chromatography) in glacial acetic acid. The reaction mixture was passed through 0.2-μm AcroPrep Advance filter plates (Pall Corp., New York, NY). Kinetic assays were performed under the same conditions as routine activity assays (protein concentration of 50 μg and reaction at 31 °C for 60 min, which had been demonstrated to fall into the linear range of the candidate enzymes, with less than 20% of substrate being converted). Kinetic parameters were determined by varying substrate concentrations while maintaining other reactants at saturation (same concentrations as in routine activity assays). Kinetic constants (Kₘ and kₗₜₜ) were calculated by nonlinear regression analysis (Origin 8; OriginLab Corp., Northampton, MA).

The separation of enzyme assay substrates and products was performed using a Zorbax SB-Aq C18 rapid resolution HT column (2.1 × 50 mm, 1.8-μm particle size) integrated into a 1290 HPLC system (Agilent Technologies). The mobile phase consisted of solvent A (0.1% formic acid) and solvent B (0.1% formic acid in acetonitrile). The gradient elution procedure was as follows: 0 min, 5% (v/v) B; 0 to 25 min, linear gradient from 0 to 30% of B; 25 to 35 min, 30 to 95% B; 35 to 45 min, 95 to 98% B; and then re-equilibration at starting conditions for 4 min. The flow rate was 0.3 ml/min. Traces and spectra in the UV and visible range were recorded with an in-line diode array detector (Agilent Technologies) and mass spectra with a 6530 QTOF-MS (Agilent Technologies) for m/z values of 100 to 1,500 at a scan rate of 1.1 s⁻¹. The instrument was operated with an electrospray ion source in negative polarity mode. The nebulizer gas temperature was set to 325 °C, the nebulizer gas flow rate was set to 5 liters/min, the nebulizer pressure was set to 25 ps.i., the sheath gas temperature was set to 325 °C, the sheath gas flow rate was set to 7.5 liters/min, the capillary voltage was set to 3500 V, the nozzle voltage was set to 2000 V, the fragmentor voltage was set to 175 V, the voltage on the skimmer was set to 65 V, and the octupole RF peak voltage was set to 750 V. A reference mass solution (containing hexakis(1H,1H,3H-H-tetrafluorophenyl)diphosphazene (250 nm) and purine (300 nm) in methanol) was infused into the ion source using an isocratic pump at a flow rate of 0.1 ml/min. The fragmentor voltage was set to 30 eV for MS/MS experiments.

Author contributions—N. S. and B. M. L. supervision; N. S. validation; N. S., I. L., M. H., Q. L., and M. M. investigation; N. S. visualization; N. S., I. L., M. H., Q. L., and M. M. methodology; B. M. L. conceptualization; B. M. L. funding acquisition; B. M. L. writing—original draft.

Acknowledgment—Amber N. Parrish is thanked for maintaining T. media cell suspension cultures. We would also like to thank Dr. Dick B. Janssen for providing a vector containing the Pc21g20650 gene of Penicillium chrysogenum.

References
1. Schein, Z. (1976) History of the National Cancer Institute and the plant screening program. Cancer Treat. Rep. 60, 975–977 Medline
2. Arbus, S. G., Christian, M. C., Fisherman, J. S., Genes, L. A., Sarosy, G., Saffin, M., Adams, J., Canetta, R., Cole, K. E., and Friedman, M. A. (1993) Clinical development of Taxol. J. Natl. Cancer Inst. Monogr. 11–24 Medline
3. Russin, W. A., Ellis, D. D., Gottwald, J. R., Zeldin, E. L., Brodhagen, M., and Evert, R. F. (1995) Immunocytochemical localization of Taxol in Taxus–Cuspidata. Int. J. Plant Sci. 156, 668–678 CrossRef
4. Cragg, G. M. (1998) Paclitaxel (Taxol): a success story with valuable lessons for natural product drug discovery and development. Med. Res. Rev. 18, 315–331 CrossRef Medline
5. Holton, R. A., Biediger, R. J., and Boatman, P. D. (1995) Semisynthesis of Taxol and taxotere. In Taxol: Science and Applications (Suffness, M., ed) pp. 97–121, CRC Press, New York
6. Kingston, D. G. I. (2001) Taxol, a molecule for all seasons. Chem. Commun. 10, 867–880 CrossRef
7. Fischdeck, J. T., Johnson, S. R., Ketchum, R. E., Crouse, R. B., and Lange, B. M. (2015) NMR spectroscopic search module for Spectraris, an online resource for plant natural product identification: Taxane diterpenoids from Taxus × media cell suspension cultures as a case study. Phytochemistry 113, 87–95 CrossRef Medline
8. Ketchum, R. E. B., and Crouse, R. B. (2006) The Taxus metabolome and the elucidation of the Taxol (R) biosynthetic pathway in cell suspension cultures. in Plant Metabolomics. Biotechnology in Agriculture and Forestry (Saito, K. D. R. A., and Willmitzer, L., eds) pp. 291–309, Springer, Berlin, Heidelberg, Germany
9. McElroy, C., and Jennewein, S. (2018) Taxol® biosynthesis and production: from forests to fermenters. In Biotechnology of Natural Products (Schwab, W., Lange, B. M., and Wüst, M., eds) pp. 145–185, Springer International Publishing, Heidelberg, Germany
10. Sabater-Jara, A. B., Tudela, L. R., and López-Pérez, A. J. (2010) In vitro culture of Taxus sp.: strategies to increase cell growth and taxoid production. Phytochem. Rev. 9, 343–356 CrossRef
11. Beuerle, T., and Pichersky, E. (2002) Purification and characterization of benzoate:coenzyme A ligase from Clarksia brevire. Arch. Biochem. Biophys. 400, 258–264 CrossRef Medline
12. Klempien, A., Kaminaga, Y., Qualely, A., Nagegowda, D. A., Widhalm, J. R., Orlova, I., Shasany, A. K., Taguchi, G., Kish, C. M., Cooper, B. R., D’Auria, J. C., Rhodes, D., Pichersky, E., and Dudareva, N. (2012) Contribution of CoA ligases to benzenoid biosynthesis in petunia flowers. Plant Cell 24, 2015–2030 CrossRef Medline

13. Kliebenstein, D. J., D’Auria, J. C., Behere, A. S., Kim, J. H., Gunderson, K. L., Breen, J. N., Lee, G., Gershenzon, J., Last, R. L., and Jander, G. (2007) Characterization of seed-specific benzyloxyglucosinolate mutations in Arabidopsis thaliana. Plant J. 51, 1062–1076 CrossRef Medline

14. Stout, J. M., Boubakir, Z., Ambrose, S. J., Purves, R. W., and Page, J. E. (2012) The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in Cannabis sativa trichomes. Plant J. 71, 353–365 Medline

15. Xu, H., Zhang, F., Liu, B., Huhman, D. V., Sumner, L. W., Dixon, R. A., and Wang, G. (2013) Characterization of the formation of branched short-chain fatty acid-CoA for bitter acid biosynthesis in hop glandular trichomes. Mol. Plant 6, 1301–1317 CrossRef Medline

16. Cuthbertson, D. J., Johnson, S. R., Piljac-Zegarac, J., Kappel, J., Schäfer, S., Wüst, M., Ketchum, R. E., Croteau, R. B., Marques, J. V., Davin, L. B., Lewis, N. G., Rolf, M., Kutchan, T. M., Soejarto, D. D., and Lange, B. M. (2013) Accurate mass-time tag library for LC/MS-based metabolite profiling of medicinal plants. Phytochemistry 91, 187–197 CrossRef Medline

17. Sun, G., Yang, Y., Xie, F., Wen, J. F., Wu, J., Wilson, I. W., Tang, Q., Liu, H., and Qiu, D. (2013) Deep sequencing reveals transcriptome re-programming of Taxus x media cells to the elicitation with methyl jasmonate. PLoS One 8, e62865 CrossRef Medline

18. Zhou, T., Luo, X., Yu, C., Zhang, C., Zhang, L., Song, Y. B., Dong, M., and Chen, C. (2019) Transcriptome analyses provide insights into the expression pattern and sequence similarity of several Taxol biosynthesis-related genes in three Taxus species. BMC Plant Biol. 19, 33 CrossRef Medline

19. Lange, B. M. (2018) Commercial-scale tissue culture for the production of plant natural products: successes, failures and outlook. In Biotechnology of Natural Products (Schwab, W., Lange, B. M., and Wüst, M., eds) pp. 189–218, Springer International Publishing, Heidelberg, Germany

20. Shockey, J. M., Fulda, M. S., and Browse, J. (2003) Arabidopsis contains a large superfamily of acyl-activating enzymes: phylogenetic and biochemical analysis reveals a new class of acyl-coenzyme a synthetases. Plant Physiol. 132, 1065–1076 CrossRef Medline

21. Fulda, M., Schnurr, J., Abbadi, A., Heinz, E., and Browse, J. (2004) Peroxisomal Acyl-CoA synthetase activity is essential for seedling development in Arabidopsis thaliana. Plant Cell 16, 394–405 CrossRef Medline

22. Zhao, L., Katavic, V., Li, F., Haughn, G. W., and Kunst, L. (2010) Insertional mutational analysis reveals that long-chain acyl-CoA synthetase 1 (LACS1), but not LACS8, functionally overlaps with LACS9 in Arabidopsis seed oil biosynthesis. Plant J. 64, 1048–1058 CrossRef Medline

23. Chen, H., Kim, H. U., Weng, H., and Browse, J. (2011) Malonyl-CoA synthetase, encoded by acyl activating enzyme 13, is essential for growth and development of Arabidopsis. Plant Cell 23, 2247–2262 CrossRef Medline

24. Foster, J., Kim, H. U., Nakata, P. A., and Browse, J. (2012) A previously unknown oxalyl-CoA synthetase is important for oxalate catabolism in Arabidopsis. Plant Cell 24, 1217–1229 CrossRef Medline

25. Ehlling, J., Büttner, D., Wang, Q., Douglas, C. J., Somsisch, I. E., and Kombrik, E. (1999) Three 4-oxoararic:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. Plant J. 19, 9–20 CrossRef Medline

26. Koetsier, M. J., Jekel, P. A., Wijima, H. J., Bovenberg, R. A., and Janssen, D. B. (2011) Aminoacyl-coenzyme A synthetase catalyzed by a CoA ligase from Penicillium chrysogenum. FEBS Lett. 585, 893–898 CrossRef Medline

27. Koo, A. J., Chung, H. S., Kobayashi, Y., and Howe, G. A. (2006) Identification of a peroxisomal acyl-activating enzyme involved in the biosynthesis of jasmonic acid in Arabidopsis. J. Biol. Chem. 281, 33511–33520 CrossRef Medline

28. Liu, T., Yao, R., Zhao, Y., Xu, S., Huang, C., Luo, J., and Kong, L. (2017) Cloning, functional characterization and site-directed mutagenesis of 4-coumarate: coenzyme A ligase (4CL) involved in coumarin biosynthesis in Peucedanum praeruptorum Dunn. Front. Plant Sci. 8, 4 Medline

29. Schneider, K., Kienow, L., Schmelzer, E., Colby, T., Bartsch, M., Miersch, O., Wasternack, C., Kombrink, E., and Stuible, H. P. (2005) A new type of peroxisomal acyl-coenzyme A synthetase from Arabidopsis thaliana has the catalytic capacity to activate biosynthetic precursors of jasmonic acid. J. Biol. Chem. 280, 13962–13972 CrossRef Medline

30. Eisenhthal, R., Danson, M. J., and Hough, D. W. (2007) Catalytic efficiency and kcat/Km: a useful comparator? Trends Biotechnol. 25, 247–249 CrossRef Medline

31. Ramírez-Estrada, K., Ahtaballa, T., Onrubia, M., Moyano, E., Notredame, C., Osuna, L., Vanden Bossche, R., Goossens, A., Casido, R. M., and Palaizon, J. (2016) Transcript profiling of jasmonate-elicited Taxus cells reveals a β-phenylalanine–CoA ligase. Plant Biotechnol. J. 14, 85–96 CrossRef Medline

32. Gulick, A. M., Starai, V. J., Horswill, A. R., Homick, K. M., and Escalante-Semerena, J. C. (2003) The 1.75 Å crystal structure of acetyl-CoA synthetase bound to adenosine-5’-propyl phosphate and coenzyme A. Biochemistry 42, 2866–2873 CrossRef Medline

33. Watkins, P. A., Maiguel, D., Jia, Z., and Pevsner, J. (2007) Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genome. J. Lipid Res. 48, 2736–2750 CrossRef Medline

34. Morgan-Kiss, R. M., and Cronan, J. E. (2004) The Escherichia coli faaK (ydiD) gene encodes an aerobically regulated short chain acyl-CoA synthetase. J. Biol. Chem. 279, 37324–37333 CrossRef Medline

35. Weimar, J. D., DiRusso, C. C., Delio, R., and Black, P. N. (2002) Functional role of fatty acyl-coenzyme A synthetase in the transmembrane movement and activation of exogenous long-chain fatty acids. Amino acid residues within the ATP/AMP signature motif of Escherichia coli FadD residues are required for enzyme activity and fatty acid transport. J. Biol. Chem. 277, 29369–29376 CrossRef Medline

36. Gulick, A. M. (2009) Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. ACS Chem. Biol. 4, 811–827 CrossRef Medline

37. Loncaric, C., Ward, A. F., and Walker, K. D. (2007) Expression of an acetyl-CoA synthase and a CoA-transferase in Escherichia coli to produce modified taxanes in vivo. Biotechnol. J. 2, 266–274 CrossRef Medline

38. Ketchum, R. E., Rithner, C. D., Qiu, D., Kim, Y. S., Williams, R. M., and Croteau, R. B. (2003) Taxus metabolomics: methyl jasmonate preferentially induces production of taxoids oxygenated at C-13 in Taxus x media call cultures. Phytochemistry 62, 901–909 CrossRef Medline

39. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282 Medline

40. Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791 CrossRef Medline

41. Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874 CrossRef Medline