Molecular cloning of a 10-deacetylbaccatin III-10-*O*-acetyl transferase cDNA from *Taxus* and functional expression in *Escherichia coli*

Kevin Walker and Rodney Croteau*

Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340

Contributed by Rodney Croteau, November 17, 1999

The cDNA clone for a 10-deacetylbaccatin III-10-*O*-acetyl transferase, which catalyzes formation of the last diterpene intermediate in the Taxol biosynthetic pathway, has been isolated from *Taxus cuspidata*. By using consensus sequences from an assembly of transacylases of plant origin and from many deduced proteins of unknown function, a homology-based PCR cloning strategy was employed to amplify initially a 911-bp gene fragment of the putative taxane C-10 hydroxyl acetyl transferase from *Taxus*. This amplicon was used to screen a cDNA library constructed from mRNA isolated from methyl jasmonate-induced *Taxus* cells, from which the full-length 10-deacetylbaccatin III-10-*O*-transacetylase sequence was obtained. Expression of the ORF from pCWori^{+} in *Escherichia coli* JM109 afforded a functional enzyme, as determined by ^{1}H-NMR and MS verification of the product baccatin III derived from 10-deacetylbaccatin III and acetyl CoA. The full-length cDNA has an ORF of 1,320 bp corresponding to a deduced protein of 440 residues with a calculated molecular weight of 49,052, consistent with the size of the operationally soluble, monomeric, native acetyl transferase. The recombinant acetyl transferase has a pH optimum of 7.5, has K_m values of 10 μM and 8 μM for 10-deacetylbaccatin III and acetyl CoA, respectively, and is apparently regiospecific toward the 10-hydroxyl group of the taxane ring. Amino acid sequence comparison of 10-deacetylbaccatin III-10-*O*-acetyl transferase with taxadienol-5-*O*-acetyl transferase and with other known acyl transferases of plant origin indicates a significant degree of similarity between these enzymes (80% and 64–67%, respectively).

Taxol biosynthesis | paclitaxel | baccatin III | *Taxus cuspidata*

Taxol[†] is one of the structurally more complex members of the taxoid family of natural products characterized by the tricyclic diterpene taxane ring system. There are at least 12 distinct enzymatic reactions involved in Taxol biosynthesis (1, 2). Although the complete course of the Taxol biosynthetic pathway may involve metabolites that are transiently acylated and deacylated for the purpose of organellar targeting or flux regulation (thereby increasing the number of pathway steps), the acyl groups derived from five acyl transferase reactions are present in the final product, Taxol. The first acyl transferase in the Taxol pathway seems to be taxadien-5α-ol-*O*-acetyl transferase (3), which converts the intermediate taxa-4(20),11(12)-dien-5α-ol to taxa-4(20),11(12)-dien-5α-yl acetate (Fig. 1*A*). The gene encoding this 5α-ol transacetylase has been cloned from *Taxus cuspidata* and heterologously expressed in *Escherichia coli* (4). The 5α-yl acetate ester is considered to be the third specific intermediate of the Taxol biosynthetic pathway, because it can be further transformed to a series of advanced polyhydroxylated taxa-4(20),11(12)-diene metabolites in *Taxus* microsomal preparations optimized for cytochrome P450 reactions (5). Moreover, the 4(20)-ene-5α-acetoxy functional grouping present in taxadien-5α-yl acetate is postulated to proceed through a 4(20)-epoxy-5α-acetoxy intermediate, which undergoes intramolecular rearrangement to the 4-acetoxy-4(5)-oxetane found in advanced taxoid metabolites (ref. 6; Fig. 1*B*).

A survey of all taxoid metabolites characterized to date (7) suggests that the second acylation reaction in the Taxol pathway is the benzoylation of the taxane C-2 hydroxyl group of an advanced intermediate (2), whereas the third acylation reaction is considered to be the acetylation of the C-10 hydroxyl group of the advanced metabolite 10-DAB to yield baccatin III (ref. 2; Fig. 1*A*), the immediate diterpenoid precursor of Taxol. Evidence for the latter enzymatic transformation was obtained first by using crude cell-free extracts from needles of *Taxus baccata* (8), and a purified acetyl CoA:10-hydroxytaxane *O*-acetyl transferase has now been prepared from cultured *Taxus chinensis* cells (9). Herein, we describe the isolation and sequence of the gene encoding a DBAT from *T. cuspidata* and report the properties of the recombinant enzyme.

BIOCHEMISTRY

Materials and Methods

Substrates. [20-^{3}H]Taxa-4(20),11(12)-dien-5α-ol (6 Ci/mol) was synthesized as described (10). Unlabeled CoA and [2-^{3}H]acetyl CoA were obtained from Sigma and DuPont/NEN, respectively. Authentic 10-DAB and baccatin III were generously provided by David Bailey of Hauser Chemical Research (Boulder, CO). Methyl jasmonate was obtained from Bedoukian Research (Danbury, CT).

Bacterial Strains, Plasmid Constructs, and Cloning. To facilitate the isolation of clones encoding transacylases involved in Taxol biosynthesis, a cDNA library was constructed from mRNA derived from *Taxus* cell suspension cultures induced with methyl jasmonate to increase Taxol production (4, 11). Transacylase sequences were PCR-amplified from this enriched *T. cuspidata* cell library cDNA as template with the previously described transacylase-specific, 72-fold degenerate reverse primer [AT-REV1, 3′-CT(A/G) AA(A/G) CC(I/C/A) ACC CC(I/C/A) TT(T/C) GG-5′; ref. 4] and a 144-fold degenerate forward primer [AT-FOR3, 5′-TT(C/T) TA(T/C) CCI TT(C/T) GC(I/C/A) GG(I/C/A) AG-3′] that was designed based on a consensus sequence element (FYPFAGR) noted in an assembly of seemingly related deduced proteins containing few defined transacylase sequences of plant origin and many sequences of unknown function (4). The *Taq* polymerase catalyzed PCRs were performed at relatively low annealing temperatures (40–45°C) and with $MgCl_2$ concentrations varying between 1 and 5 mM to optimize the amplification of potential transacylase sequences,

Abbreviations: 10-DAB, 10-deacetylbaccatin III; DBAT, 10-deacetylbaccatin III-10-*O*-acetyl transferase.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF193765).

*To whom reprint requests should be addressed. E-mail: croteau@mail.wsu.edu.

[†]Paclitaxel is the generic name for Taxol, which is now a registered trademark of Bristol-Myers Squibb. Because of the greater familiarity of the word Taxol, we use it in this paper instead of paclitaxel.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

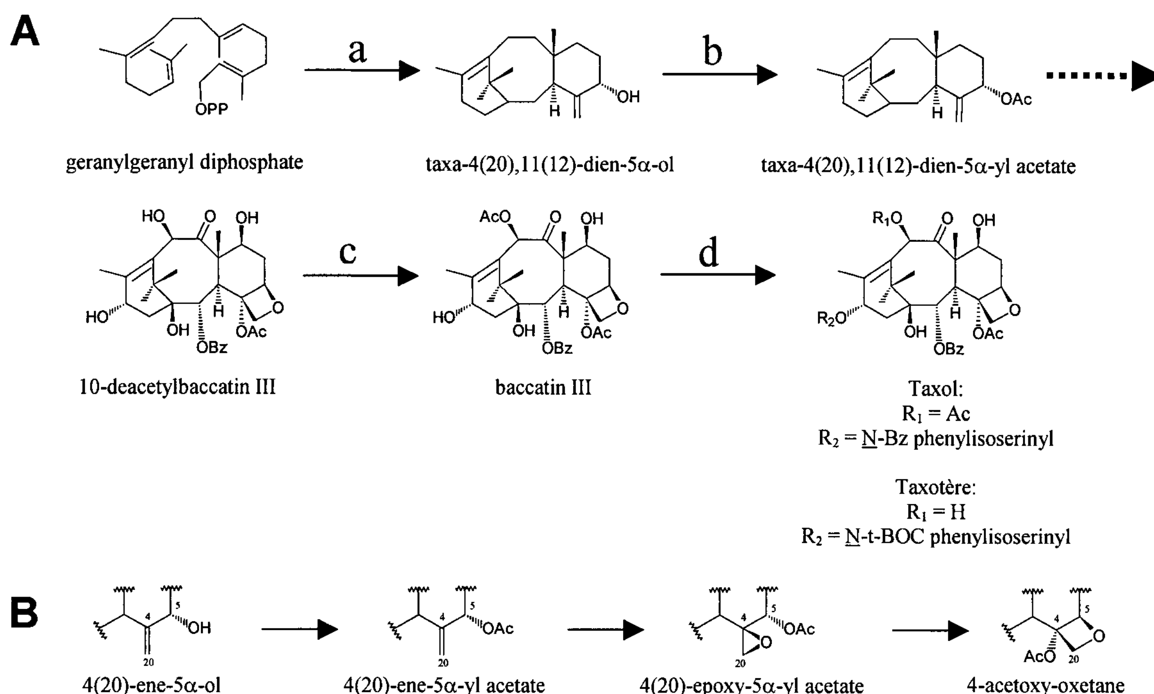

Fig. 1. (*A*) Outline of the Taxol biosynthetic pathway. The cyclization of geranylgeranyl diphosphate to taxadiene by taxadiene synthase and the hydroxylation to taxadien-5α-ol by taxadiene 5α-hydroxylase (*a*), the acetylation of taxadien-5α-ol by taxa-4(20),11(12)-dien-5α-ol-*O*-acetyl transferase (*b*), the conversion of 10-deacetylbaccatin III (10-DAB) to baccatin III by 10-deacetylbaccatin III-10-*O*-acetyl transferase (DBAT) (*c*), and the side-chain attachment to baccatin III to form Taxol (*d*) are highlighted. The broken arrow indicates several as yet undefined steps. (*B*) Postulated biosynthetic scheme for the formation of the oxetane, present in Taxol and related late-stage taxoids, in which the 4(20)-ene-5α-ol is converted to the 4(20)-ene-5α-yl acetate, followed by epoxidation to the 4(20)-epoxy-5α-acetoxy group and then intramolecular rearrangement to the 4-acetoxy oxetane moiety. Ac, acetyl; Bz, benzoyl; *t*-BOC, tertiary-butoxycarbonyl; OPP, diphosphate.

while minimizing production of nonspecific amplicons. A 911-bp fragment (designated pPRB2) was amplified by these means (PCR conditions: 94°C for 4 min; 32 cycles at 94°C for 45 s, 45°C for 1.2 min, and 72°C for 2 min; and, finally, 74°C for 5 min; each reaction contained 3 units of *Taq* polymerase and 2 mM $MgCl_2$) and was found, by cloning and sequencing, to have significant nucleotide-level homology (80%) to the *T. cuspidata* taxadien-5α-ol-*O*-acetyl transferase gene (4), thus suggesting that pPRB2 may represent a partial cDNA sequence of a related taxoid acetyl transferase. By high-stringency screening of the *T. cuspidata* cDNA library (6.0×10^4 plaques) with a ^{32}P-labeled probe derived from pPRB2 (designated probe 2), one putative full-length clone was obtained by standard hybridization and purification procedures (12). This full-length clone was 99% identical to probe 2, indicating that the probe had located its cognate. A subsequently designed primer set (293NDEF, 5′-GGGAATTCC-ATATGGCAGGCTCAACAGAATTTGTGG-3′; 293XBAR, 3′-GTTTATACATTGATTCGGAACTAGATCTGATC-5′) was employed to amplify the full-length acetyl transferase gene and to incorporate *Nde*I and *Xba*I restriction sites at the 5′ and 3′ termini, respectively, for directional ligation into vector pCWori$^+$ (13). This construct was used to transform *E. coli* JM109 by established methods (4).

Expression, DBAT Assay, and Product Identification. For the enzyme preparation, *E. coli* cells transformed with the pCWori$^+$ vector harboring the putative DBAT gene were grown overnight at 37°C in 5 ml of Luria–Bertani medium supplemented with ampicillin, and 1 ml of this inoculum was added to and grown in 100 ml of Terrific Broth culture medium supplemented with 1 mM isopropyl-β-D-thiogalactopyranoside, 1 mM thiamine HCl, and 50 μg/ml ampicillin. After 24 h, the bacteria were harvested by centrifugation, resuspended in 20 ml of assay buffer (25 mM Mopso, pH 7.4), and then disrupted by sonication at 0–4°C. The resulting homogenate was centrifuged at 15,000 × *g* to pellet debris. After determining by SDS/PAGE that a protein of the appropriate size (≈50 kDa) was expressed in operationally soluble form, a 1-ml aliquot of the supernatant was incubated with 10-DAB (400 μM) and [2-^{3}H]acetyl CoA (0.45 μCi; 400 μM) for 1 h at 31°C. The reaction mixture was then extracted with ether, and the solvent was concentrated *in vacuo*. The crude product (pooled from five such assays) was purified by silica gel TLC [70:30 (vol/vol) ethyl acetate:hexane]. The band comigrating with authentic baccatin III (R_f = 0.45 for the standard) was isolated, dissolved in 100 μl of acetonitrile, and analyzed by radio-HPLC on a Perkin–Elmer HPLC ISS 200 system equipped with a Packard A100 radioactivity detector (see Fig. 2). The sample was loaded onto an Alltech Econosil C_{18} column (5 μ; 250 × 4.6 mm), eluted at 1 ml/min with a linear gradient from 95:5 (vol/vol) H_2O:CH_3CN to 10:90 (vol/vol) H_2O:CH_3CN over 10 min, then eluted to 100% CH_3CN with a concave gradient over 30 min, and held at 100% CH_3CN for 10 min. The TLC-purified product was also analyzed by combined liquid chromatography-MS on a Hewlett–Packard Series 1100 MSD system in the atmospheric pressure chemical ionization mode. The sample was dissolved in acetonitrile, loaded onto a Phenomenex Curosil PFP column (Belmont, CA; 5 μ; 4.6 × 250 mm), and eluted with 50:50 (vol/vol) acetonitrile:water at 1 ml/min, with the effluent directed to the atmospheric pressure chemical ionization mass spectrometer (see Fig. 3). Proton ^{1}H-NMR spectra were recorded on a Varian Mercury 300, and samples were dissolved in deuterated chloroform as internal standard.

Partial Purification and Characterization of Recombinant *T. canadensis* DBAT. *E. coli* JM109 cells were grown and harvested as described above. After centrifugation (15,000 × *g*), the supernatant was

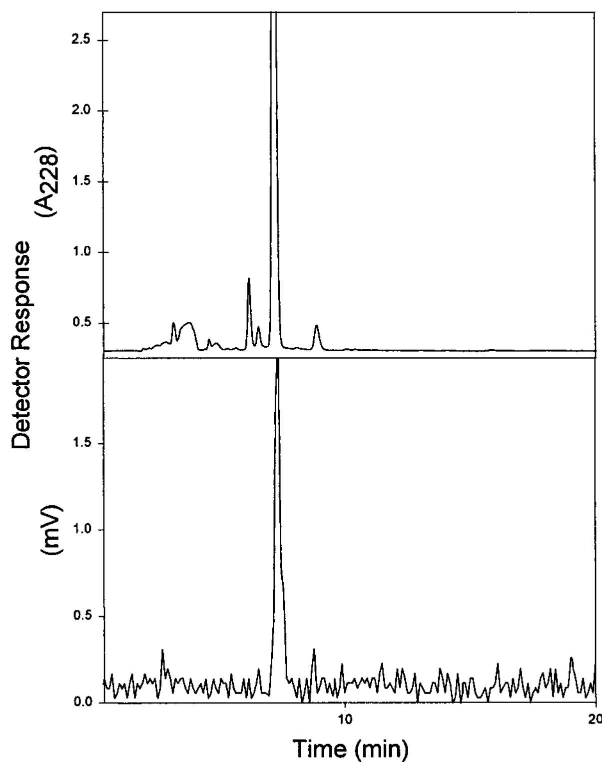

Fig. 2. Radio-HPLC analysis of the biosynthetic product (retention time = 7.0 ± 0.1 min) generated from 10-DAB and [2-^{3}H]acetyl CoA by the recombinant acetyl transferase. The UV profile (*Upper*) and the coincident radioactivity profile (*Lower*) coincide with the retention time of authentic baccatin III. For column, solvent, flow rates, and counting details, see *Materials and Methods*.

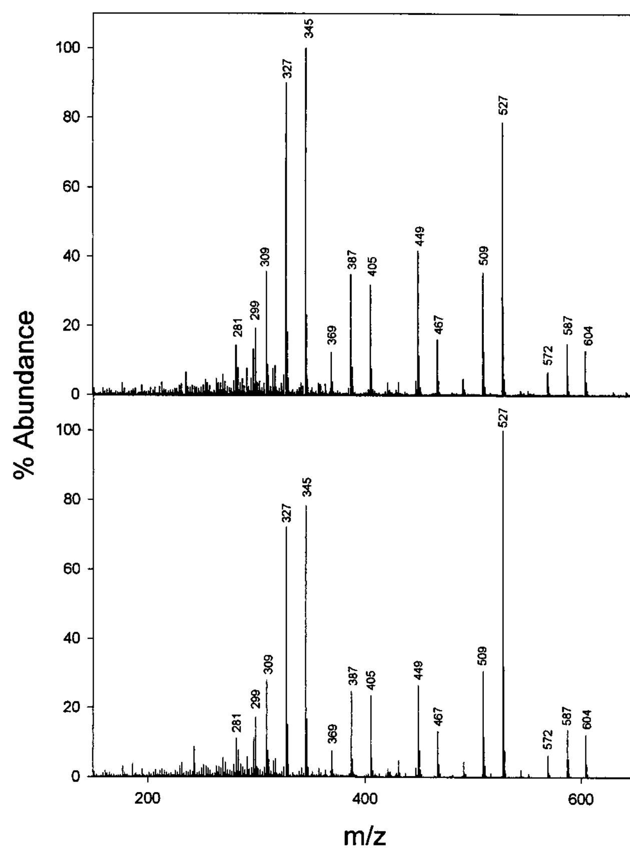

Fig. 3. Combined reverse-phase HPLC-atmospheric pressure chemical ionization MS analysis of the biosynthetic product (retention time = 8.6 ± 0.1 min) generated by the recombinant acetyl transferase with 10-DAB and acetyl CoA as cosubstrates (*Upper* spectrum) and of authentic baccatin III (retention time = 8.6 ± 0.1 min; *Lower* spectrum). The diagnostic mass spectral fragment ions are at m/z 605 ($P + NH_4^+$), 587 (PH^+), 572 ($PH^+ - CH_3$), 527 ($PH^+ - CH_3COOH$), and 509 ($PH^+ - (CH_3COOH - H_2O)$). For column, solvent, flow rates, and instrument details, see *Materials and Methods*.

applied to a column of *O*-diethylaminoethylcellulose (2.8 × 10 cm, Whatman DE-52) that was previously washed with 25 mM Mopso buffer (pH 7.4) containing 3 mM DTT and 1 M NaCl and then equilibrated with this buffer (without NaCl). After removal of unbound material, protein was eluted with a linear gradient of 0 to 200 mM NaCl in equilibration buffer (total volume of 125 ml at 3 ml/min). Fractions containing DBAT activity (eluting at ≈120 mM NaCl) were combined and subjected to dilution and ultrafiltration (Amicon Diaflo YM 10 membrane, Millipore) to decrease the salt concentration to ≈5 mM as well as to remove selectively proteins of molecular weight lower than that of DBAT.

For kinetic evaluations, linearity with respect to protein concentration and time was first established, and the concentrations of 10-DAB (0.5–200 μM) and [2-^{3}H]acetyl CoA (0.5–200 μM) were then independently varied while the remaining reactant was maintained at saturation. Double reciprocal plots were constructed for each data set, and the equation of the best-fit line ($R^2 = 0.97$) was determined (KALEIDAGRAPH, version 3.08, Synergy Software, Reading, PA). To examine the influence of pH on transacetylase activity, enzyme preparations (2 ml) were dialyzed at 4°C for 24 h in Mes (pH 6), Mopso (pH 7.4), glycine (pH 8), Capso (pH 9), and Caps (pH 10 and 11), all at 25 mM containing 3 mM DTT, before the assay.

Results and Discussion

Cloning and Heterologous Expression of DBAT from *T. cuspidata*. The Taxol pathway is presumed to proceed through the late-stage, oxetane-bearing metabolite 10-DAB, which is regiospecifically 10-*O*-acetylated to the penultimate pathway intermediate, baccatin III (ref. 2; Fig. 1). In an effort to isolate the gene responsible for the acyl group transfer from acetyl CoA to 10-DAB, a homology-based PCR cloning strategy was used to amplify transacylase sequences from enriched library cDNA made from mRNA isolated from *T. cuspidata* cells treated with methyl jasmonate to induce Taxol production. Sequence analysis of the resulting 911-bp amplicon revealed a partial cDNA (designated pPRB2) that bore 80% nucleotide-level homology to the *T. cuspidata* taxadien-5α-ol-*O*-acetyl transferase gene (4), suggesting that the full-length version of pPRB2 may encode another taxoid transacetylase.

By using ^{32}P-labeled pPRB2 as a probe to screen the library described above, the full-length cDNA was obtained, transferred into vector pCWori$^+$, and expressed in *E. coli* JM109 to test for the function of this putative transacetylase. After preliminary determination that a protein of the appropriate size (≈50 kDa, determined by SDS/PAGE) was expressed in operationally soluble form, preparative-scale cultures of the transformed bacteria were generated, and the resulting soluble enzyme fraction was partially purified by anion-exchange chromatography to remove competing activities and assayed under standard conditions (4) with 10-DAB (400 μM) and [2-^{3}H]acetyl CoA (0.45 μCi, 400 μM) as cosubstrates. This enzyme preparation yielded a single reaction product on reversed-phase radio-HPLC analysis, with a retention time of 7.0 ± 0.1 min (coincident radio and UV traces) corresponding exactly to that of authentic baccatin III (Fig. 2). Extracts of *E. coli* transformed with empty

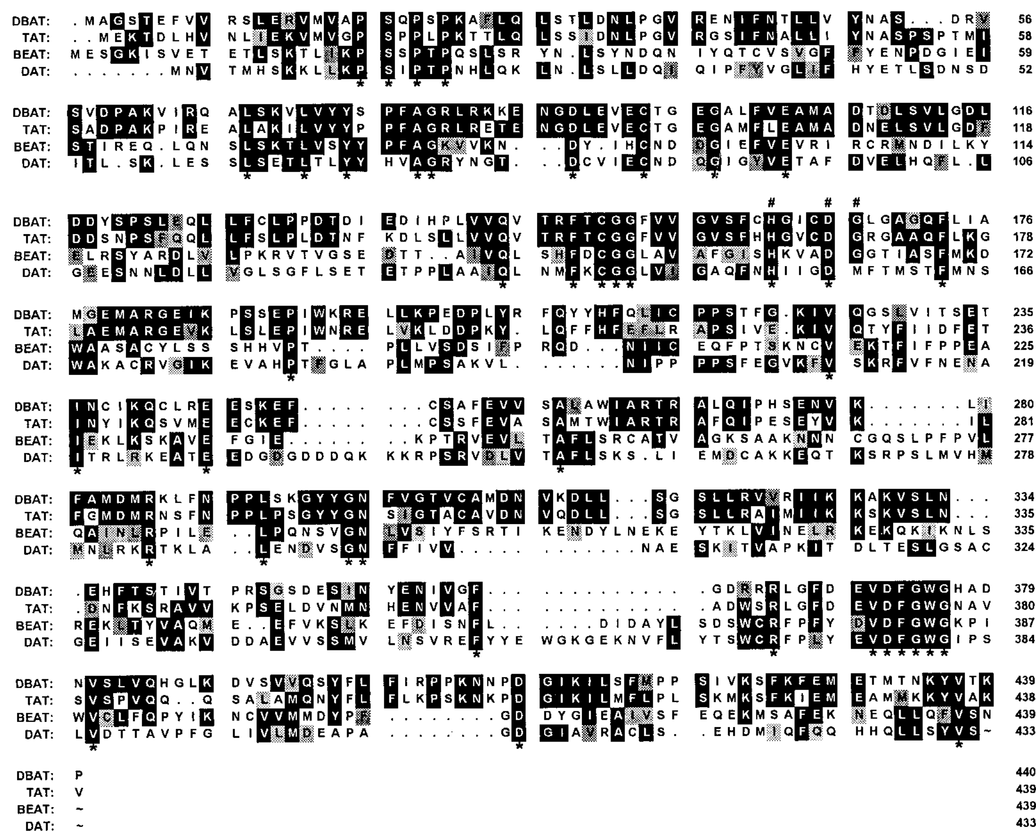

Fig. 4. Deduced amino acid sequence comparison of DBAT (accession no. AF193765) from *T. cuspidata*, taxadien-5α-ol-*O*-acetyl transferase (TAT, accession no. AF190130) from *T. cuspidata*, benzylalcohol acetyl transferase (BEAT, accession no. AF043464) from *Clarkia breweri*, and deacetylvindoline-4-*O*-acetyl transferase (DAT, accession no. AF053307) from *Catharanthus roseus*. Residues boxed in black indicate positional identity for at least two of the compared sequences; similar amino acids are indicated by gray shading. Asterisks (*) indicate conserved residues in all acetyl transferase sequences of plant origin. Pound signs (#) indicate a putative acyl group transfer motif (HXXXDG) present in three of the four sequences. The alignment was created with the PILEUP program (Wisconsin Package Version 9.0; Genetics Computer Group, Madison, WI).

vector controls did not yield detectable product when assayed by identical methods (data not shown).

The identity of the biosynthetic product was further verified as baccatin III by combined liquid chromatography-MS analysis (Fig. 3), which demonstrated the identical retention time (8.6 ± 0.1 min) and mass spectrum for the product and for the authentic standard. Finally, a sample of the biosynthetic product, purified by silica gel TLC, gave a ^{1}H-NMR spectrum identical to that of authentic baccatin III, confirming that the enzyme was DBAT and that the corresponding gene had been isolated.

Characterization of Heterologously Expressed DBAT. The recombinant enzyme was determined to have a pH optimum of 7.4, with half-maximal velocities at pH 6.4 and 7.8. The K_m values for 10-DAB and acetyl CoA were determined to be 10 μM and 8 μM, respectively, by Lineweaver–Burk analysis (for both plots, $R^2 = 0.97$). These kinetic constants for DBAT are comparable to those of the taxa-4(20),11(12)-dien-5α-ol acetyl transferase (possessing K_m values for taxadienol and acetyl CoA of 4 μM and 6 μM, respectively). The DBAT seems to acetylate the 10-hydroxyl group of taxoids with a high degree of regioselectivity, because the enzyme does not acetylate the 1β-, 7β-, or 13α-hydroxyl groups of 10-DAB, nor does it acetylate the 5α-hydroxyl group of taxa-4(20),11(12)-dien-5α-ol.

Sequence Analysis. The DBAT cDNA contains an ORF of 1,320 nucleotides (GenBank accession no. AF193765) encoding a deduced protein of 440 amino acid residues (Fig. 4) with a calculated molecular weight of 49,052. This size is consistent with that of the native DBAT protein, ≈50 kDa, determined by gel permeation chromatography, indicating the protein to be a functional monomer; this size is also very similar to that of the related, monomeric taxadien-5α-ol transacetylase (molecular weight = 49,079). The acetyl CoA:DBAT from *T. cuspidata* seems substantially different in size from the acetyl CoA:10-hydroxytaxane-*O*-acetyl transferase recently isolated from *T. chinensis* and reported at a molecular weight of 71,000 (9).

The deduced amino acid sequence of DBAT resembles that of taxadien-5α-ol acetyl transferase from *T. cuspidata* (ref. 4; 64% identity; 80% similarity) and the sequences of other acetyl transferases (56–57% identity; 65–67% similarity) involved in different pathways of secondary metabolism in plants (14, 15). Additionally, DBAT possesses the HXXXDG motif (residues H162, D166, and G167, respectively) found in other acyl transferases (refs. 16–19; Fig. 4); this sequence element has been suggested to function in acyl group transfer from acyl CoA to the substrate alcohol (15).

As the demand for Taxol use in cancer treatment increases, the long-term supply of this antineoplastic drug will continue to be an important issue. The supply of Taxol is currently largely sustained by semisynthetic means in which 10-DAB, isolated from yew needles, is used to produce Taxol and the closely related analog Taxotère (see Fig. 1). The semisynthetic procedure for Taxol production involves protection of the 7-hydroxyl of 10-DAB, chemical acetylation of the 10-hydroxyl to afford 7-*O*-protected baccatin III, synthetic attachment of the benzoyl

phenylisoserine side chain at the 13-hydroxyl, and, finally, deprotection to yield Taxol. The related procedure for Taxotère preparation involves protection of the 7- and 10-hydroxyls of 10-DAB, acylation of the 13-hydroxyl with the *N-t*-butoxycarbonyl phenylisoserine side chain, and, lastly, deprotection.

With increased understanding of the enzymology, molecular genetics, and regulation of the Taxol pathway, it may be feasible to augment these semisynthetic approaches and, ultimately, to replace them entirely with organisms genetically engineered for enhanced Taxol production. In addition to the confirmed DBAT cDNA and taxadienol-5-*O*-acetyl transferase cDNA (4), there are at least three additional transacylation steps in the Taxol biosynthetic pathway, represented by the 2-debenzoyl baccatin III-2-*O*-benzoyl transferase, the baccatin III-13-*O*-phenylisoserine transferase, and the debenzoyltaxol-*N*-benzoyl transferase. The general cloning strategy used to obtain taxadienol-5-*O*-acetyl transferase (4) and DBAT is likely to yield these remaining transacylase genes.

We thank C. Sanchez, G. Munske, and D. Pouchnik for nucleotide sequencing and primer synthesis. This investigation was supported in part by Grant CA-55254 from the National Institutes of Health, by Cytoclonal Pharmaceutics, and by McIntire–Stennis Project 0967 from the Washington State University Agricultural Research Center.

1. Croteau, R., Hefner, J., Hezari, M. & Lewis, N. G. (1995) in *Phytochemicals and Health*, eds. Flores, H. E. & Gustine, D. L. (Am. Soc. Plant Physiol., Rockville, MD), pp. 94–104.
2. Floss, H. G. & Mocek, U. (1995) in *Taxol*: *Science and Applications*, ed. Suffness, M. (CRC, Boca Raton, FL), pp. 191–208.
3. Walker, K., Ketchum, R. E. B., Hezari, M., Gatfield, D., Golenowski, M., Barthol, A. & Croteau, R. (1999) *Arch. Biochem. Biophys.* **364**, 273–279.
4. Walker, K., Schoendorf, A. & Croteau, R. (2000) *Arch. Biochem. Biophys.*, in press.
5. Hezari, M. & Croteau, R. (1997) *Planta Med.* **63,** 291–295.
6. Guéritte-Voegelein, F., Guénard, D. & Potier, P. (1987) *J. Nat. Prod.* **50,** 9–18.
7. Baloglu, E. & Kingston, D. G. I. (1999) *J. Nat. Prod.* **62,** 1448–1472.
8. Zocher, R., Weckwerth, W., Hacker, C., Kammer, B., Hornbogen, T. & Ewald, D. (1996) *Biochem. Biophys. Res. Commun.* **229,** 16–20.
9. Menhard, B. & Zenk, M. H. (1999) *Phytochemistry* **50,** 763–774.
10. Hefner, J., Rubenstein, S. M., Ketchum, R. E. B., Gibson, D. M., Williams, R. M. & Croteau, R. (1996) *Chem. Biol.* **3,** 479–489.
11. Ketchum, R. E. B., Gibson, D. M., Croteau, R. B. & Shuler, M. L. (1999) *Biotechnol. Bioeng.* **62,** 97–105.
12. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
13. Barnes, H. J. (1996) *Methods Enzymol.* **272,** 3–14.
14. Dudareva, N., D'Auria, J. C., Nam, K. H., Raguso, R. A. & Pichersky, E. (1998) *Plant J.* **14,** 297–304.
15. St-Pierre, B., Laflamme, P., Alarco, A.-M. & De Luca, V. (1998) *Plant J.* **14,** 703–713.
16. Brown, N. F., Anderson, R. C., Caplan, S. L., Foster, D. W. & McGarry, J. D. (1994) *J. Biol. Chem.* **269,** 19157–19162.
17. Carbini, L. A. & Hersh, L. B. (1993) *J. Neurochem.* **61,** 247–253.
18. Hendle, J., Mattevi, A., Westphal, A. H., Spee, J., de Kok, A., Teplyakov, A. & Hol, W. G. (1995) *Biochemistry* **34,** 4287–4298.
19. Lewendon, A., Murray, I. A., Shaw, W. V., Gibbs, M. R. & Leslie, A. G. (1994) *Biochemistry* **33,** 1944–1950.