Taxol Biosynthesis: An Update

Mehri Hezari1 and Rodney Croteau1,2
1 Institute of Biological Chemistry, Washington State University, Pullman, Washington 99164-6340, USA
2 Address for correspondence

Received: November 26, 1996; Accepted: February 22, 1997

Abstract: The novel diterpenoid taxol (paclitaxel) is now well-established as a potent chemotherapeutic agent. Total synthesis of the drug is not commercially feasible and, in the foreseeable future, the supply of taxol and its synthetically useful progenitors must rely on biological methods of production. The first three steps of taxol biosynthesis have been defined and the responsible enzymes described. These are the cyclization of the universal diterpenoid precursor geranylgeranyl diphosphate to taxa-4(5),11(12)-diene, the cytochrome P450-catalyzed hydroxylation of this olefin to taxa-4(20),11(12)-dien-5a-ol, and the acetyl CoA-dependent conversion of the alcohol to the corresponding acetate ester. Demonstration of these early steps of taxol biosynthesis suggests that the complete pathway can be defined by a systematic, stepwise approach at the cell-free enzyme level. When combined with in vivo studies to determine contribution to pathway flux, slow steps can be targeted for gene isolation and subsequent overexpression in Taxus to improve the yield of taxol and related compounds.

Key words: Taxol, paclitaxel, biosynthesis, taxadiene synthase, taxadiene hydroxylase, taxadienol acetyltransferase.

Introduction

Very limited supply and early difficulties with formulation did not prevent the development of the diterpenoid natural product taxol* as one of the most useful anticancer drugs now available (1, 2). Taxol has proven effective in the treatment of ovarian and breast cancer, and shows promise for a variety of other cancers such as head and neck, lung, gastrointestinal, and bladder (2, 3). Unlike the Vinca alkaloids, colchicine, and other antimitotic agents that act by preventing the polymerization of tubulin into microtubules, taxol promotes microtubule assembly and suppresses depolymerization (5, 6).

Originally, taxol was obtained by extraction from the peeled bark of the Pacific yew (Taxus brevifolia Nutt.; Taxaceae), a small, slow-growing evergreen tree sparsely dispersed along the coastal and interior mountains of the Pacific Northwest (7). The harvest is destructive and the yields are low; consequently, alternative sources of the drug were sought (8). Nicolaou (9, 10), Holton (11), and others (12) have achieved the total synthesis of taxol by two very elegant routes; however, because of the high costs involved, this approach is not commercially feasible. Currently, taxol and its analog taxotere are being produced semi-synthetically via the acylation of 10-deacetyl-baccatin III (see Fig. 1) and related compounds obtained from the needles of Taxus (13-16). This source of starting material,

Fig. 1 Cyclization of geranylgeranyl diphosphate to taxa-4(5),11(12)-diene and elaboration of the olefin to taxol, cephalomannine or baccatin III. OPP denotes the diphosphate moiety, Bz the benzoyl group, and Ac the acetyl group.

* Paclitaxel is the generic name for Taxol, which is now a registered trademark of Bristol-Myers Squibb. Because of the far greater familiarity with the word taxol, we use it in this paper in lieu of paclitaxel. Docetaxel is the generic name for Taxotere, which is a registered trademark of Rhone-Poulenc Rorer. The structural differences between taxol (Fig. 1) and taxotere are a tert-butoxycarbonyl group instead of a benzoyl group on the nitrogen atom at C-3' of the C-13 side-chain and a hydroxy function instead of an acetate at the C-10 position of the taxane 5-ring.
although still low yielding, is at least renewable and can be practiced on the plantation scale. *Taxomyces andreanae* and other microorganisms have been reported to produce taxol (17, 18), and several groups have been successful in producing taxol in reasonable titers in Taxus cell culture (19, 20) but these production alternatives have not yet proved to be commercially viable. Given the projected increase in taxol usage and the reliance on biological methods of production (8), it is important to understand the biosynthesis of taxol, especially the rate-limiting steps of the pathway which may be manipulated to improve drug yields. This review summarizes recent developments in the study of taxol biosynthesis.

**Previous Work on Taxol Biogenesis**

The structure of taxol is quite complex in bearing an unusual diterpene carbon skeleton, eight o xo-functional groups, and an assortment of appended side-chains, for a total of eleven stereocenters (21) (Fig. 1). Most proposals for taxol formation posit the initial generation of an olefin with the taxane skeleton [as proposed first by Lythgoe (22)], followed by a series of oxygenation steps and, ultimately, acylation reactions (23, 24). Based on precedent in the metabolism of other terpenoid natural products (25), the pathway to taxol from the universal diterpenoid precursor geranylgeranyl diphosphate must involve over a dozen individual enzymatic steps. Given this complexity, the low yields and the operational limitations of the biological sources, deciphering the biogenetic origin of taxol represents a considerable challenge.

Early studies on the biogenesis of taxol involved in vivo feeding experiments with radiolabeled precursors that seemed to indicate that acetate, mevalonate, and phenylalanine represented the building blocks of taxol (26). Recently, however, Zenk and his associates (27) have shown conclusively that mevalonate is not a precursor of the diterpenoid taxane skeleton but rather have implicated the origin of the key intermediate isopentenyl diphosphate by an alternate pathway for isoprenoid biosynthesis proposed by Rohmer (28). In exploring later pentenyl diphosphate by an alternate pathway for isoprenoid rather than a precursor of the diterpenoid taxane ring system but the endocyclic olefin was demonstrated to be present in *T. brevifolia* bark at low levels (5–10 μg per kg dry wt.), whereas the exocyclic isomer was not detected (31). Both taxa-4(5),11(12)-diene and the 4(20)-isomer were subsequently synthesized by Rubenstein and Williams (32).

A cell-free system from *T. brevifolia* sampling stem was shown to catalyze the divalent metal ion dependent conversion of geranylgeranyl diphosphate to taxadiene as essentially the only olefinic product (33). This operationally soluble enzyme was localized mainly in the bark and adhering cambium cells, and a 600-fold conventional purification, by chromatographic and electrophoretic means, yielded the pure protein which the combination of gel permeation chromatography and SDS-PAGE showed to be a monomer of ~79 kDa (33). This taxadiene synthase has been characterized with regard to pH optimum, kinetic constants, and possible active site residues, and shown to prefer Mg²⁺ (Kₘ ~0.16 mM) to Mn³⁺ and other divalent metal ions as the required cofactor.

The cyclization reaction has been examined in some detail and shown to occur without detectable free intermediates or the preliminary formation of enzyme-bound taxa-4(20),11(12)-diene with subsequent isomerization, and to involve an intramolecular hydrogen transfer from C-11 of a verticillene intermediate to the C-ring of taxa-4(5),11(12)-diene (34). From these results, a stereochemical mechanism has been proposed for the taxadiene synthase reaction involving the initial cyclization of geranylgeranyl diphosphate to a transient verticillene cation intermediate, with transfer of the C-11 α-proton to C-7 to initiate transannular B-C-ring closure to the taxenyl cation, followed by protonation at C-5 to yield the taxa-4(5),11(12)-diene product directly (Fig. 2).

The low levels of taxadiene synthase activity in *Taxus* stem tissue (140 pmol·h⁻¹·g⁻¹), the very low abundance of taxadiene in bark, and the efficient conversion of the olefin to

**First Step in Taxol Biosynthesis: Taxadiene Formation**

The committed step in the biosynthesis of taxol and related taxoids is the cyclization of the universal diterpenoid precursor geranylgeranyl diphosphate to taxa-4(5),11(12)-diene (Fig. 1) (31), not the exocyclic isomer, taxa-4(20),11(12)-diene, as earlier proposed based on the abundance of taxane metabolites with this structure (23, 24). In vivo experiments with *T. brevifolia* stem sections, using labeled taxa-4(5),11(12)-diene, demonstrated the very efficient incorporation (30%) of the olefin into taxol and closely related taxoids, and the endocyclic olefin was demonstrated to be present in *T. brevifolia* bark at low levels (5–10 μg per kg dry wt.), whereas the exocyclic isomer was not detected (31). Both taxa-4(5),11(12)-diene and the 4(20)-isomer were subsequently synthesized by Rubenstein and Williams (32).

The taxadiene synthase has been characterized with regard to pH optimum, kinetic constants, and possible active site residues, and shown to prefer Mg²⁺ (Kₘ ~0.16 mM) to Mn³⁺ and other divalent metal ions as the required cofactor.

The cyclization reaction has been examined in some detail and shown to occur without detectable free intermediates or the preliminary formation of enzyme-bound taxa-4(20),11(12)-diene with subsequent isomerization, and to involve an intramolecular hydrogen transfer from C-11 of a verticillene intermediate to the C-ring of taxa-4(5),11(12)-diene (34). From these results, a stereochemical mechanism has been proposed for the taxadiene synthase reaction involving the initial cyclization of geranylgeranyl diphosphate to a transient verticillene cation intermediate, with transfer of the C-11 α-proton to C-7 to initiate transannular B-C-ring closure to the taxenyl cation, followed by protonation at C-5 to yield the taxa-4(5),11(12)-diene product directly (Fig. 2).

The low levels of taxadiene synthase activity in *Taxus* stem tissue (140 pmol·h⁻¹·g⁻¹), the very low abundance of taxadiene in bark, and the efficient conversion of the olefin to
taxoids in vivo suggested that the cyclization of geranylgeranyl diphosphate to taxadiene was slow relative to downstream steps in the biosynthetic pathway to taxol. The potential role of this first committed step in the control of pathway flux prompted a search for the cDNA encoding taxadiene synthase. Attempts to obtain amino acid sequence information from the enzyme, for use as a cloning tool, were unsuccessful, and led to the adoption of a PCR-based strategy built upon homologous sequences of other known terpenoid cyclases of plant origin. One set of primers, of the many tested, yielded a small DNA fragment that was "cyclase-like" in sequence and that was subsequently employed as a hybridization probe to screen a cDNA library from T. brevifolia stem. Several of the cDNA clones so isolated were full-length and were functionally expressed in Escherichia coli, thereby providing confirmation by the conversion of geranylgeranyl diphosphate to taxadiene (35). The sequence specifies an open reading frame of 2586 nucleotides, and the complete deduced polypeptide contains 862 amino acid residues, including a long plastidial targeting sequence, corresponding to a molecular weight of about 98,000. Sequence comparison with other terpenoid synthases from higher plants showed a significant degree of similarity (especially with those of other gymnosperms), suggesting that hydroxylation of taxoids with the 4(20)-ene-5-oxy functional grouping. Since an assessment of the possible role of taxadiene synthase in the control of pathway flux to taxol could not be easily carried out in intact Taxus stem tissue, heterotrophic suspension cell cultures of the Canadian yew (Taxus canadensis) were employed as an experimental system. The enzyme from this source was isolated and shown to be chromatographically, electrophoretically, and kinetically identical to the taxadiene synthase of T. brevifolia stem (20). Results from the analysis of enzyme activity levels during the time-course of taxol accumulation in developing cell cultures indicated a rise in taxadiene synthase activity before taxol accumulation began, which persisted into stationary phase, and also that synthase activity in vitro exceed the maximum rate of taxol accumulation in vivo (20). Combined, these results suggest that, although taxadiene synthase represents a slow biosynthetic step, rate-limiting transformations lay further down the pathway in this system.

Second Step in Taxol Biosynthesis: Taxadienol Formation

No oxygenated taxoids bearing a 4(5)-double bond are known, whereas taxoids with the 4(20)-ene-5-oxy functional grouping are exceedingly common (36), suggesting that hydroxylation of taxa-4(5),11(12)-dien-5a-ol. Membranous fractions were inactive in the acetylation reaction. The hydroxylase, found mainly in the light membrane fraction, fulfilled all of the expected requirements of a cytochrome P450 monoxygenase (heme thiolate protein), including CO inhibition and blue light reversal, and other modes of oxygenation could be eliminated (39). Many terpenoid hydroxylases are of the cytochrome P450 type (40). Whether the mechanism of conversion of the 4(5),11(12)-dien to the 4(20),11(12)-dien-5a-ol involves an epoxide intermediate or direct oxygen insertion at C-5 is not yet known. Nevertheless, the structure of this first oxygenated intermediate on the taxol pathway establishes that the hydroxylation reaction proceeds with an unusual double bond migration that constrains the mechanistic possibilities for subsequent elaboration at this site of the oxetane moiety of taxol (see below).

The role of taxa-4(20),11(12)-dien-5a-ol as a pathway intermediate was confirmed by in vivo conversion (in the 10–15% incorporation range) into advanced taxoids, including 10-deacetylbaccatin III, cephalomannine, and taxol, in T. brevifolia stem disks (39). Radiochemically-guided fractionation of an extract of T. brevifolia dried bark additionally demonstrated that the alcohol (at 5–10 µg·kg\(^{-1}\)) and its esters (at 25–50 µg·kg\(^{-1}\)) were present as naturally occurring metabolites in the relevant tissue. Both of these observations based on in vivo studies suggest that this first oxygenation step in taxol biosynthesis is slow relative to subsequent metabolic transformations and, thus, a worthy candidate for gene isolation.

Third Step in Taxol Biosynthesis: Acylation of Taxadienol

The abundance of naturally occurring taxoids that are acetylated, or otherwise acylated, at the C-5 position (36) suggested that the acylation of taxa-4(20),11(12)-dien-5a-ol could be the next step in taxol biosynthesis. Additionally, the 4(20)-ene-5a-acetoxy functional group could play a role in the elaboration of the oxetane ring of taxol by a mechanism involving epoxidation of the 4(20)-double bond followed by intramolecular acetoxy migration with oxirane ring expansion (Fig. 3) (23, 24, 39). These considerations, as well as studies on subsequent oxygenation steps (see below), prompted an evaluation of the enzymatic acylation of taxa-4(20),11(12)-dien-5a-ol.

A soluble enzyme preparation from T. canadensis suspension cultured cells was shown to catalyze the acetyl coenzyme A dependent acylation of taxa-4(20),11(12)-dien-5a-ol. Membranous fractions were inactive in the acetylation reaction. The soluble transacylase has been partially purified and characterized as a 50 kDa monomeric protein (with \(pI \approx 4.7\)) that accepts acetyl CoA, but not benzoyl CoA, as a cosubstrate.
Subsequent Oxygenation Steps

Subsequent reactions on the taxol pathway include additional oxygenations, hydroxy group acylations, oxidation to a ketone, and generation of the oxetane ring system. The large number of naturally occurring taxoid metabolites with a broad range of oxygenation and acylation patterns provides little guidance in predicting the order of functionalization beyond taxa-4(20),11(12)-dien-5α-yl acetate. Consideration of the relative abundance of taxoids bearing an oxygen function at each carbon of the taxane ring has suggested the order of oxygenation to be C-5 followed by C-10, then C-2 and C-9 (24, 41). Metabolites bearing an oxygen at C-13 are next in abundance, thus suggesting a sequence for the conversion of taxadiene to the level of a pentaol (Fig. 4). Oxygenations at C-7 and C-1 of the taxane nucleus are considered to be late introductions, possibly occurring after oxetane ring formation. However, oxetane formation, itself a relatively late-stage transformation, seemingly must precede acylation at C-13 and oxidation of the C-9 hydroxy to a carbonyl (cf. Fig. 1) (24, 41).

Microsomal preparations of *T. canadensis* cell cultures, when incubated with taxa-4(20),11(12)-dien-5α-ol in the presence of NADPH and O₂, yield an enzymatic product more polar than the substrate, with chromatographic properties and mass spectrum (by HPLC-MS) consistent with a taxadiene diol. The reaction, like the hydroxylation of taxadiene, is catalyzed by a cytochrome P450 hydroxylase, suggesting that the remaining oxygenation steps in taxol biosynthesis may involve similar enzymes. By utilizing the microsomal system that had been carefully optimized to sustain cytochrome P450-type reactions (38), both taxa-4(5),11(12)-diene and taxa-4(20),11(12)-dien-5α-ol as substrates gave rise to products that were tentatively identified by HPLC-MS analysis as the corresponding diol, triol, and tetraol; the precise structures of these products are as yet uncertain. Incubation of the optimized microsomes with taxa-4(20),11(12)-dien-5α-yl acetate as substrate yielded a single major product that, based on HPLC-MS evaluation, appeared to be a pentaol monooacetate. Taxa-4(20),11(12)-dien-5α-yl benzoate was not detectably oxygenated by this system, indicating that the efficiency of the acetate ester as precursor was not due simply to increased lipophilicity but rather to improved recognition of this substrate by this set of microsomal hydroxylases that appear to operate sequentially. This successive series of hydroxylations is now being systematically characterized with regard to substrates and the structural elucidation of the corresponding products, and with consideration to the possibility of intervention of additional acylation steps in the sequence.

Conclusion

Improvement of the biological production of taxol, and of related taxoid metabolites of use in drug semi-syntheses (42), should be based on a thorough understanding of the complex target pathway and its rate limiting steps. A systematic, stepwise approach should reveal the sequence of oxygenation and acylation steps, the timing of the dehydrogenation at the C-9 hydroxy and the formation of the unique oxetane ring, and the coordination of these processes with the assembly of the N-benzoyl phenylisoserine side chain (29, 30). The role of each enzymatic step in the control of pathway flux can then be assessed by *in vivo* studies, the slow steps revealed, and suitable cloning strategies devised to acquire the corresponding genes. This approach will ultimately lead to engineering of *Taxus* or relevant microorganisms that overexpress key genes to increase flux through the corresponding enzyme catalyzed steps in order to improve production yields of taxol.

Fig. 4 Proposed sequence for the hydroxylation of taxa-4(5),11(12)-diene to the level of a pentao based on the relative abundances of naturally occurring taxoids. The reactions are probably catalyzed by cytochrome P450 oxygenases. Acetylation at C-5 before further hydroxylation seems likely, but whether acylations of the various other hydroxy groups are required steps of the oxygenation sequence is unknown.

Acknowledgements

The research by the authors was supported in part by U.S. National Institutes of Health Grant CA-55254, McIntire-Stennis Project 0967 from the Agricultural Research Center, Washington State University, and a grant-in-aid from Cytoclinal Pharmaceuticals Inc.

References