Pergamon

*Tetrahedron*, Vol. 52, No. 27, pp. 8957-8968, 1996

0040-4020/96 $15.00 + 0.00

PII: S0040-4020(96)00444-9

# NMR and Molecular Modeling Study of Paclitaxel Putative Precursors

**Yvan Boulanger*, Abdesslem Khiat,**

INRS-Santé, Université du Québec, 245 boul. Hymus, Pointe-Claire, Québec H9R 1G6

**Zhen-Hua Zhou, Gaétan Caron and Lolita O. Zamir ***

Institut Armand-Frappier, Université du Québec, 531 boul. des Prairies, Laval, Québec H7N 4Z3, Canada

**Abstract:** Isolation and purification procedures for the bicyclic oxygenated taxane canadensene, obtained from *Taxus canadensis* are reported. Molecular modeling studies of canadensene, baccatin III (the core structure of paclitaxel), paclitaxel, taxa-4(5),11(12)-diene (the biosynthetic precursor of paclitaxel) and putative derivatives were performed. NMR constraints were obtained for canadensene and baccatin III. A striking U-shape for canadensene was obtained in the 3D-model. Biosynthetic pathways for the formation of canadensene are proposed. Copyright © 1996 Elsevier Science Ltd

The 1990's can be characterized by a huge effort from chemists, biochemists and biologists to investigate many facets of research on paclitaxel (Taxol®) **1** (Fig. 1), the amazing anti-cancer drug which was recently approved in the United States and in Canada. The core structure of paclitaxel is also a natural product found in yews: baccatin III **2** (Fig. 1). Some of the key discoveries involve the unusual mode of action of paclitaxel[1], the discovery of the active semi-synthetic docetaxel (Taxotère®)[2], two elegant total synthetic schemes[3], and the identification of a cell-free preparation from *Taxus brevifolia* which established the first tricyclic intermediate to taxanes: taxa-4(5),11(12)-diene **3**.[4,5] This last impressive work proves unambiguously that in the Pacific yew (*T. brevifolia*), the three main cycles of the taxane skeleton (A, B and C) are formed before the oxygenations of this olefin intermediate. The identification of this precursor also ruled out the intermediacy of taxa-4(20),11(12)-diene, which had been previously suggested.[6] Prior to the publication of this important biosynthetic finding, we had reported in a preliminary communication[7] the isolation from the Canadian yew (*Taxus canadensis*) of a taxane canadensene **4** (Fig. 1), which was extensively oxygenated *but with rings B and C not yet cyclized.* In addition, the stereochemistry of the oxygenated substituents at the positions 2, 7, 10 and 13 was identical to paclitaxel. Attempts to synthetically cyclize canadensene failed.[8] Many explanations could be given for the detection of canadensene as a natural product: 1) it could be an alternate biosynthetic pathway to taxanes more prevalent in the Canadian yew, 2) it could be derived from opening of the cyclized oxygenated taxane or 3) it could be a dead end metabolite originating from oxygenations of the uncyclized bicyclic hydrocarbon.

The first step in trying to delineate which of these hypotheses occur in nature involved the understanding of the energetic requirements for the uncyclized canadensene-like versus the cyclized taxadiene-like moieties.

Subsequent to the discovery of canadensene **4**, two very similar structures (one with an additional acetate at C20 and the other at C5) were isolated from *Taxus chinensis*.[9] The NMR and molecular modeling studies reported relate to paclitaxel, docetaxel and some analogues in hydrophilic and hydrophobic solvents.[10-15] Under all conditions, the baccatin core was found to adopt a U-shaped conformation. The studies revealed that only the structure of the A-ring side chain of paclitaxel was solvent dependent. The structures of paclitaxel determined by NMR[11,14,15] were very similar to the docetaxel structure determined by X-ray diffraction.[16]

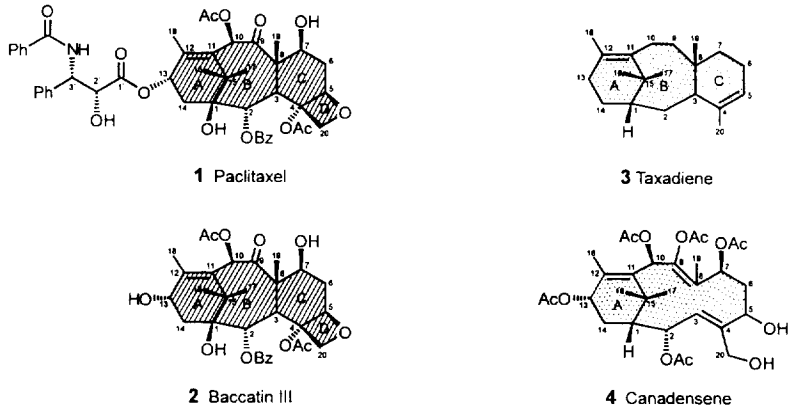



**Fig. 1**. Natural taxanes investigated in this study.

In this publication, we report the isolation and purification procedures of canadensene **4** derived from *Taxus canadensis*. In addition, a detailed study using NMR and molecular modeling enabled us to gain insight into the structural and energetic properties of paclitaxel putative precursors: 1) the bicyclic oxygenated taxanes such as canadensene and other hypothetical bicyclic hydrocarbons, 2) the tricyclic hydrocarbons (taxadienes) and 3) the tetracyclic core of paclitaxel (baccatin III). On the basis of these data, a biosynthetic scheme could be postulated.

## EXPERIMENTAL

*Isolation of Canadensene*

Ground dried needles and small branches (1.4 kg) from *T. canadensis* were suspended in 12 l of methanol/dichloromethane (1 : 1 v/v) and shaken for 3 days at room temperature. The ground plants were filtered and extracted again with fresh solvent for another 3 days. The combined organic extracts were evaporated under reduced pressure. Water (700 ml) was added and lipids were removed by stirring the mixture with hexane (6 x 500 ml). The aqueous phase and the insoluble material were transferred to a separatory funnel and extracted with dichloromethane (3 x 700 ml). The combined dichloromethane extracts were dried over magnesium sulfate, filtered and evaporated yielding a dark green extract (39 g).

*Purification of Canadensene*

The extract was dissolved in dichloromethane / methanol (2 : 1 v/v; 300 ml). Celite (170 g) was added and the mixture evaporated under reduced pressure. The solid was separated on a silica gel column (40 cm x 5 cm) eluting with the following solvent mixtures: hexane/ dichloromethane (4 : 1; 2 l); hexane/ dichloromethane (3 : 2; 2 l); hexane/ dichloromethane/ ethyl acetate (6 : 3 : 1; 2 l); hexane/ dichloromethane/ ethyl acetate (3 : 1 : 1; 2 l); hexane/ dichloromethane/ ethyl acetate (5 : 2 : 3; 2 l); hexane/ dichloromethane/ ethyl acetate (5 : 1 : 4; 2 l); hexane/ ethyl acetate/ acetone (45 : 45 : 10; 2 l) and hexane/ ethyl acetate/ acetone (1 : 3 : 1; 2 l). The products of the last fraction (5.2 g) were separated on a reverse phase column (ODS-2 MAG-20, 2.2 cm x 50 cm from Whatman) using a preparative Waters Delta 3000 HPLC instrument coupled to a variable wavelength detector (Model 481 from Waters) at 227 nm. The products were eluted with a 70 min linear gradient of acetonitrile (25 to 100%) in water at a flow rate of 18 ml/min. In these conditions, 9-dihydro-13-acetylbaccatin III, and canadensene were eluted and collected together at a retention time of 30.1 min yielding a total of 380 mg for the two products. At this point, most of the 9-dihydro-13-acetylbaccatin III, could be discarded by precipitation in cold methanol. Canadensene was then separated by chromatography on silica gel (30 cm x 5 cm) eluting with a gradient of ethyl acetate (30 to 100%) in hexane (canadensene $R_f = 0.42$; 9-dihydro-13-acetylbaccatin III $R_f = 0.50$ with 100% ethyl acetate). A final HPLC purification on reverse phase column (ODS-2 MAG-9, 9.4 x 500 mm from Whatman) eluting with an isocratic mobile phase of isopropanol (33%) in water at a flow rate of 6 ml/min afforded canadensene (26 mg; 0.002%) as a white solid (canadensene: $R_t = 38$ min; 9-dihydro-13-acetylbaccatin III: $R_t = 52$ min).

*NMR Measurements*

NMR samples were prepared by dissolving 3 mg of baccatin III **2** and 1 mg of canadensene **4** in 0.5 ml DMSO-$d_6$ or DMSO-$d_6$: $H_2O$ 1:1 for final concentrations of 6.0 and 2.2 mM, respectively. All NMR spectra were recorded on a Bruker AMX2 500 spectrometer (Bruker Spectrospin, Milton, Canada) operating at 500.13 MHz and at 300 K. One-dimensional $^1H$ and $^1H$-decoupled $^{13}C$ NMR spectra as well as two-dimensional DQFCOSY, TOCSY, NOESY (mixing times 400-600 ms) and HMQC spectra with time-proportional phase incrementation were recorded. NMR spectra were processed with FELIX 2.30 software (Biosym Technologies, San Diego CA) operating on a Silicon Graphics Indigo workstation (Silicon Graphics Inc., Montreal, Canada).

*Molecular Modeling*

Molecular modeling calculations were performed on the Silicon Graphics workstation with Biosym softwares (INSIGHT II, DISCOVER, DGII). Chemical structures were constructed using the Biosym consistent valence forcefield and submitted to a conformational search consisting of a $10^5$ fs molecular dynamics trajectory at 1000 K with sampling at every 100 fs. Each sampled conformation was energy minimized for 1000 steps using the conjugate gradients algorithm. Cluster analysis revealed the existence of three families of conformers, two U-shaped structures with the extremities pointing up or down, and the third structure being more planar. The lowest energy conformer of each family was submitted to the following procedures. The structures of paclitaxel **1**, baccatin III **2** and canadensene **4** were calculated by distance geometry using the NMR constraints to produce ten structures, followed by 10,000 steps of simulated annealing (maximal temperature 1200 K) and conjugate gradients energy minimization with a root-mean-square gradient deviation (RMSD) of 0.001 Å. The NMR

constraints for paclitaxel **1** consisted of 39 previously published NOE distances for the molecule in $CDCl_3$.[10] Distance constraints from NOESY spectra were classified as short, medium and long with maximal values of 3.0, 4.0 and 5.0 Å, respectively. When equivalent protons were present on a same heteroatom, a pseudoatom was created and a distance correction of 1 Å was applied. The structure of canadensene **4** was calculated by the same method using our NMR constraints which consisted of 26 NOE distances and eight dihedral angles derived from vicinal coupling constants.

The molecular models of canadensene and taxadiene analogues were calculated using 1000 steps of steepest descent energy minimization with a RMS gradient of 0.001 Å, the molecules were then solvated using the periodic boundary condition model in a box (20 x 20 x 20 $Å^3$) of 227 water molecules and the whole system was submitted to 2000 steps of conjugate gradients energy minimization with a RMS gradient of 0.001 Å.

## RESULTS AND DISCUSSION

*Taxus canadensis* differs from other yews by its taxane composition. Indeed, it is the only yew reported to accumulate 9-dihydro-13-acetylbaccatin III[17-18] (5-7 times the amount of paclitaxel) in its needles. This taxane was

**Table 1.** Proton NMR Data for Canadensene **4** in Different Solvents

| Proton | δ (ppm); J (Hz) | | |
|---|---|---|---|
| | $CDCl_3$[7] | DMSO-$d_6$ | DMSO-$d_6$: $H_2O$ 1:1 |
| 1 | 1.76; $J^{1,2} = 4.9$; $J^{1,14a} = 6.9$ | 1.66; $J^{1,2} = 4.4$; $J^{1,14a} = 7.0$ | 1.69; $J^{1,2} = 5.6$; $J^{1,14a} = 7.7$ |
| 2 | 5.82; $J^{2,3} = 11.7$ | 5.72; $J^{2,3} = 11.6$ | 5.69; $J^{2,3} = 11.3$ |
| 3 | 6.14 | 5.95 | 5.97 |
| 5 | 4.55; $J^{5,6a} = 2.5$ | 4.37; $J^{5,6a} = 2.9$ | 4.39; $J^{5,6a} = 2.8$ |
| OH-5 | | 4.95 | 5.11 |
| 6a | 2.79; $J^{6a,6b} = 12.7$; $J^{6a,7} = 9.3$ | 2.68; $J^{6a,6b} = 14.6$; $J^{6a,7} = 9.8$ | 2.63; $J^{6a,6b} = 16.1$; $J^{6a,7} = 10.0$ |
| 6b | 1.99 | 1.82 | 1.90 |
| 7 | 5.30 | 5.23 | 5.11 |
| 10 | 7.31 | 7.19 | 7.13 |
| 13 | 5.29; $J^{13,14a} = 9.1$ | 5.17; $J^{13,14a} = 9.6$ | 5.16; $J^{13,14a} = 9.8$ |
| 14a | 2.55; $J^{14a,14b} = 17.1$ | 2.44; $J^{14a,14b} = 16.5$ | 2.44; $J^{14a,14b} = 17.3$ |
| 14b | 2.12 | 1.94 | 1.94 |
| 16 | 1.13 | 1.02 | 0.99 |
| 17 | 1.26 | 1.17 | 1.13 |
| 18 | 2.05 | 1.99 | 1.95 |
| 19 | 1.61 | 1.51 | 1.48 |
| 20a | 4.54; $J^{20a,20b} = 12.7$ | 4.29; $J^{20a,20b} = 12.9$ | 4.29; $J^{20a,20b} = 12.3$ |
| 20b | 3.67 | 3.57 | 3.51 |
| OAc#1 | 2.20 | 2.16 | 2.15 |
| OAc#2 | 2.13 | 2.08 | 2.06 |
| OAc#3 | 2.00 | 1.94 | 1.92 |
| OAc#4 | 1.97 | 1.95 | 1.92 |
| OAc#5 | 1.92 | 1.86 | 1.87 |

also reported in *Taxus chinensis*[19] but only as a minor metabolite isolated from the bark. Due to its large amount in *T. canadensis*, it was necessary to crystallize most of it out, prior to the tedious purification of canadensene by multiple HPLC runs. In the same chromatographic fraction as that abundant metabolite, a very minor taxane (0.002 %) canadensene was found to co-elute.

*NMR Analysis*

In order to obtain the NMR constraints for molecular modeling studies, baccatin III **2** and canadensene **4** were investigated in different solvents. In paclitaxel analogues[11], the core ring structure was unaffected and only the side chain on the A ring showed significant conformational changes upon varying the polarity of the medium. This result was confirmed with baccatin III **2** which lacks the side chain on the A ring. In order to determine if the flexible bicyclic conformation of canadensene **4** is solvent dependent, its NMR spectra were recorded in $CDCl_3$, DMSO-$d_6$ and DMSO-$d_6$: $H_2O$ 1:1. The $^1H$ chemical shifts and $^1H$-$^1H$ coupling constants of canadensene

**Table 2**. Carbon-13 Chemical Shifts (ppm) for Canadensene **4** in Different Solvents

| Carbon | $CDCl_3$[7] | DMSO-$d_6$ | DMSO-$d_6$: $H_2O$ 1:1 | Carbon | $CDCl_3$ | DMSO-$d_6$ | DMSO-$d_6$ $H_2O$ 1:1 |
|---|---|---|---|---|---|---|---|
| 1 | 48.04 | 46.51 | 47.06 | 16 | 33.94 | 33.66 | 34.51 |
| 2 | 70.96 | 69.14 | 70.85 | 17 | 25.41 | 25.06 | 25.61 |
| 3 | 122.98 | 120.85 | 122.87 | 18 | 17.48 | 16.90 | 17.86 |
| 4 | 142.19 | 142.36 | 142.24 | 19 | 12.48 | 12.48 | 13.08 |
| 5 | 68.03 | 65.96 | 67.39 | 20 | 58.62 | 57.36 | 58.10 |
| 6 | 38.34 | 37.96 | 37.98 | OAc#1-$CH_3$ | 20.12 | 20.57 | 21.29 |
| 7 | 68.61 | 67.45 | 68.65 | OAc#1-CO | 170.10 | 168.37 | 171.56 |
| 8 | 125.53 | 124.72 | 125.85 | OAc#2-$CH_3$ | 21.00 | 21.30 | 22.11 |
| 9 | 144.69 | 143.13 | 144.07 | OAc#2-CO | 172.18 | 170.53 | 173.96 |
| 10 | 70.38 | 68.95 | 70.83 | OAc#3-$CH_3$ | 20.71 | 21.10 | 22.14 |
| 11 | 137.19 | 136.02 | 137.41 | OAc#3-CO | 171.35 | 169.71 | 173.18 |
| 12 | 136.36 | 135.02 | 135.98 | OAc#4-$CH_3$ | 21.00 | 21.62 | 22.14 |
| 13 | 70.67 | 69.05 | 70.83 | OAc#4-CO | 171.35 | 169.01 | 172.31 |
| 14 | 26.29 | 25.55 | 26.03 | OAc#5-$CH_3$ | 20.71 | 20.93 | 21.71 |
| 15 | 36.80 | 36.03 | 36.89 | OAc#5-CO | 169.68 | 167.78 | 171.03 |

**4** in these three solvents are presented in Table 1; the $^{13}C$ chemical shifts are given in Table 2. Very little difference is seen in the chemical shifts and coupling constants when going from the apolar solvent $CDCl_3$ to the polar solvents DMSO-$d_6$ and DMSO-$d_6$: $H_2O$ 1:1, indicating that no significant conformational change occurs upon varying the polarity of the medium. We can therefore conclude that the bicyclic conformation of canadensene **4** is not very flexible and behaves like the tricyclic core structure of paclitaxel (baccatin III **2**) whose NMR data (Table 3) was also not affected by solvent change. The $^1H$ and $^{13}C$ chemical shifts and the $^1H$- $^1H$ coupling constants of baccatin III **2** are listed in Table 3. Except for the atoms in the vicinity of position 13 where the paclitaxel side chain is lacking, the chemical shifts and coupling constants are practically identical to those of paclitaxel under the same conditions.[10-12,14]

**Table 3.** Proton and Carbon-13 NMR Data for Baccatin III **2** in $CDCl_3$

| Atom | δ $^1$H (ppm); J (Hz) | δ $^{13}$C (ppm) | Atom | δ $^1$H (ppm); J (Hz) | δ $^{13}$C (ppm) |
|---|---|---|---|---|---|
| 1 | | 79.3 | 16 | 1.13 | 21.1 |
| 2 | 5.64; $J^{2,3} = 7.1$ | 75.1 | 17 | 1.13 | 27.2 |
| 3 | 3.90 | 46.3 | 18 | 2.07 | 15.7 |
| 4 | | 81.1 | 19 | 1.69 | 9.7 |
| 5 | 4.99; $J^{5,6a} = 9.6$; $J^{5,6b} = 1.4$ | 84.6 | 20a | 4.32; $J^{20a,20b} = 8.4$ | 76.6 |
| 6a | 2.57; $J^{6a,6b} = 14.2$; $J^{6a,7} = 6.9$ | 35.8 | 20b | 4.17 | |
| 6b | 1.86; $J^{6b,7} = 11.0$ | | C=O OBz | | 167.3 |
| 7 | 4.48 | 72.5 | q-OBz | | 129.6 |
| 8 | | 59.0 | *o*-OBz | 8.12; $J^{o,m} = 7.6$ | 130.2 |
| 9 | | 204.3 | *m*-OBz | 7.50; $J^{m,p} = 7.4$ | 130.0 |
| 10 | 6.34 | 76.4 | *p*-OBz | 7.62 | 135.0 |
| 11 | | 132.2 | C=O OAc-4 | | 170.8 |
| 12 | | 146.5 | Me OAc-4 | 2.29 | 22.8 |
| 13 | 4.91 | 68.2 | OH-7 | 2.49 | |
| 14a | 2.31; $J^{14a,14b} = \sim 9^a$ | 38.8 | C=O OAc-10 | | 171.5 |
| 14b | 2.06 | | Me OAc-10 | 2.26 | 21.1 |
| 15 | | 21.1 | | | |

[a] Inaccurate due to the overlap between the signals of H-14a and H-18 as well as between the signals of H-14b and Me OAc-4.

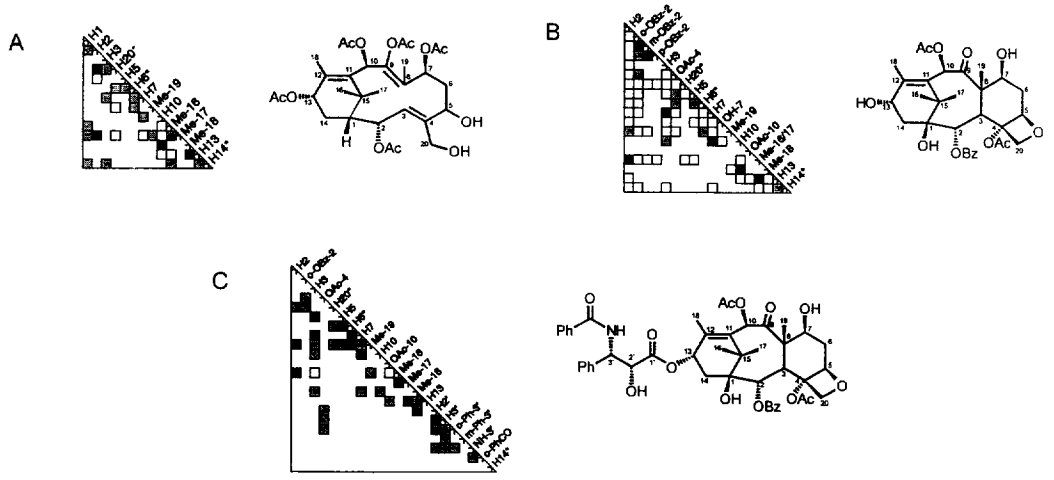


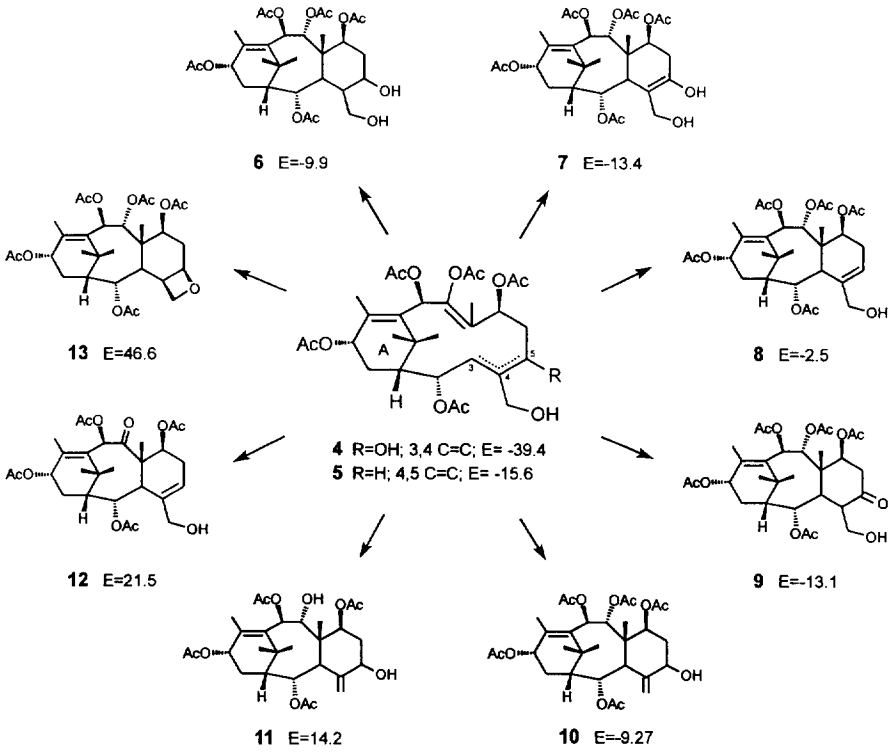

**Fig. 2.** Diagonal plot of NOESY connectivities observed in $CDCl_3$ for (A) canadensene **4**[7], (B) baccatin III **2** and (C) paclitaxel **1**[10]. An asterisk indicates the presence of more than one proton on a same carbon atom. The color of individual squares indicates the intensity of the connectivity: white, weak or < 5 Å; gray, medium or < 4 Å and black, strong or < 3 Å.

The NOESY connectivities obtained for canadensene **4** and baccatin III **2** are summarized in Fig. 2A and

2B, respectively. No significant differences in NOESY connectivities were observed between the three solvents. For both molecules, the connectivities are consistent with a close proximity between all the protons. In the case canadensene **4**, 26 NOE distance constraints and two dihedral angle constraints calculated from the vicinal coupling constants between single neighbouring protons were used for molecular modeling. In the case of baccatin III **2**, 59 NOE distance constraints and one dihedral angle constraint were used. For paclitaxel **1**, 39 NOE constraints from the literature were used[10] (Fig. 2C).

*Molecular Modeling: Energy Calculations*

Our first goal was to obtain the energy requirements of the bicyclic and tricyclic structures related to canadensene **4** and taxadiene **3**, respectively. Could these values clarify the following question: if taxadiene **3**, a tricyclic hydrocarbon with no oxygenated substituents is a precursor to paclitaxel **1**, why is the oxygenated bicyclic taxane canadensene **4** a co-metabolite of paclitaxel **1**? Could canadensene derive from the opening of a tricyclic taxane? Is there any conformation difference between the bicyclic canadensene **4** and the putative bicyclic

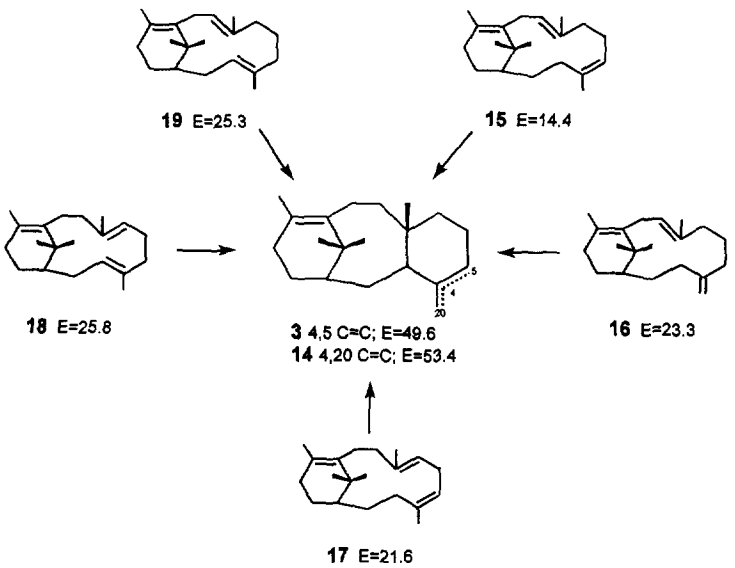



**Scheme 1**. Chemical structures of bicyclic canadensene **4** and analogue **5** as well as possible tricyclic derivatives used for molecular modeling calculations. Minimal energies in kcal/mol are indicated for each structure.

precursors of taxadiene **3**. Molecular modeling calculations were therefore performed for canadensene **4**, a putative canadensene like-compound **5** (Scheme 1) with the double bond at C4/C5, and their respective tricyclic analogues. In addition, similar calculations were done for taxadienes **3** and **14** with their bicyclic putative precursors (Scheme 2). Each structure was energy minimized in the presence of a box of water molecules.

In Scheme 1 minimal energy values are shown for some representative putative tricyclic compounds which could originate from canadensene **4** or the hypothetical analogue **5**. The calculated energy for **5** which has a double bond at C4/C5 as in taxadiene **3**, was -15.6 kcal/mol, a little less stable than canadensene. The trend is very obvious: all the tricyclic compounds are less stable than canadensene or compound **5**. The positioning of the double bonds or the oxidation states at C5 in the tricyclic structures (**6-11**) influences its inherent energies but still remain higher than their bicyclic counterparts **4** and **5**. Compound **12** with a keto group at C9 is a probable product of cyclization of canadensene but is in a higher energy state. The oxetane ring at C4/C5 (**13**) as in paclitaxel shows the highest energy.

Using the structure of taxadiene **3** (with the double bond at C4/C5, which converts to paclitaxel[4]) and taxadiene **14** (with the double bond at C4/C20, the presumed precursor of paclitaxel) which are devoid of polar substituents, the energies of different plausible bicyclic precursors were compared. Representative bicyclic structures with their respective energies are shown in Scheme 2. Similarly to the previous case, all the bicyclic structures are more stable than the tricyclic **3** and **14**. Comparing Schemes 1 and 2 leads to the following trend: the oxygenated bicyclic structures **4** and **5** seem more amenable to cyclization than the non-oxygenated ones **15-19**. Indeed, the energy differences between **4**, **5** and **6-11** are inferior than between **15-19** and **3**, **14**.




**Scheme 2.** Chemical structures of tricyclic taxadiene **3** and analogue **14** as well as possible bicyclic precursors used for molecular modeling calculations. Minimal energies in kcal/mol are indicated for each structure.

*Molecular Modeling: 3D Structures*

The three-dimensional structures of paclitaxel **1**, baccatin III **2** and canadensene **4** were calculated using the distance and dihedral angle constraints derived from NMR data (Fig. 2, Tables 1-3). For baccatin III **2**, the lowest energy structures originated from the U-shape family. Surprisingly, the same shape was obtained for the bicyclic canadensene **4** and for both compounds the energies were equivalent whether the extremities pointed up or down. The lowest energy structures for **4** and **2** are shown in Fig. 3A and 3B, respectively. Root-mean-square deviation (RMSD) values for all atoms were calculated as 2.88 Å and 1.74 Å for the ten minimized structures of canadensene **4** and baccatin III **2**, respectively. The lowest energy structure of paclitaxel is illustrated in Fig. 3C and is quite similar to the docetaxel structure obtained by X-ray diffraction.[16] Most of our structures are however of the inverse U shape compared to the reported X-ray structure of docetaxel. The RMSD value was 1.8 Å for the ten paclitaxel structures. For all three molecules, no distance constraint violations exceeding 0.5 Å were observed following molecular modeling calculations. The ring core structures of the three molecules are superimposed in Fig. 3D and are found to be very similar, a slightly more pronounced bending being observed for paclitaxel **1**. Interactions involving the sidechain in position 13 are likely responsible for this higher degree of curvature (Fig. 3C). In the case of paclitaxel **1**, two hydrogen bonds are possible in the minimized structure, one between the hydrogen atom of OH-7 and the carbonyl oxygen atom of OAc-4 and the other between the hydrogen atom of OH-7 and the ester oxygen of OAc-4. One hydrogen bond is forming in canadensene **4** between the hydrogen atom of OH-20 and the oxygen atom of OH-5.

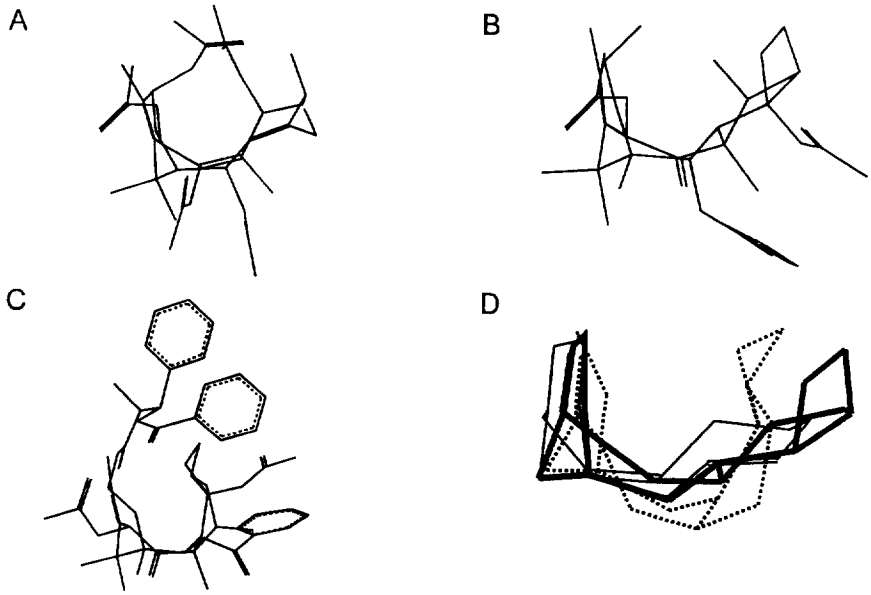



**Fig. 3.** Molecular models of (A) canadensene **4**, (B) baccatin III **2** and (C) paclitaxel **1** calculated using NMR distance (Fig. 1) and dihedral angle constraints. Superimposition of the ring cores of the three molecules are illustrated in (D): canadensene **4** (narrow line), baccatin III **2** (thick line) and paclitaxel **1** (dotted line).

Having obtained a U-shape structure for canadensene **4** similar to that of baccatin III and paclitaxel, we decided to investigate whether the bicyclic precursors of taxadiene **3** would have the same shape. The 3D structure of bicyclic canadensene **4** with the bicyclic precursors of taxadiene **3** (which was shown[4,5] in *T. brevifolia* to be a paclitaxel biosynthetic precursor) were therefore derived from molecular modeling. Interestingly, there was a marked difference in curvature for the overall conformation of canadensene **4** or canadensene-like **5** versus the pre-taxadiene bicyclic conformers **15-19**. Indeed, we can see in Fig. 4 that the structures of the bicyclic canadensene **4** or **5** are much more curved in the U-shape than the structures of the bicyclic taxadiene precursors **15-19**. These results illustrate the importance of the polar substituents for the U-shape conformation of the bicyclic molecule.

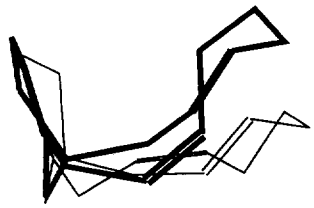
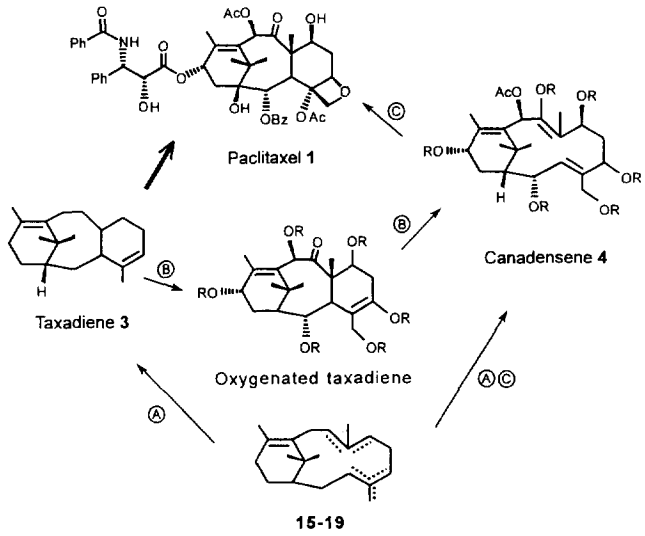


**Fig. 4.** Superimposition of the bicyclic cores of canadensene **4** (thick line) and taxadiene analogue **19** (narrow line) illustrating the difference in curvature.

From the energy calculations we understand that canadensene could not spontaneously cyclize to a tricyclic taxane. This cyclization might require activation (for example with a phosphate group at C20 or C5). On the basis of our present work, we can try to formulate some biogenetic schemes to explain the biosynthesis of canadensene **4**, a co-metabolite of paclitaxel in *T. canadensis*. If we assume that there is only *one* pathway for paclitaxel production in all the yews, and since taxadiene **3** has been shown to be a precursor[4], then canadensene **4** would be a shunt metabolite of the taxadiene pathway. One possibility for its formation could be the one shown in Scheme 3A. It would derive from extensive oxygenations of some bicyclic hydrocarbon (a variant of **15-19**) which is the probable precursor to taxadiene **3**. In this case, canadensene **4** could be a dead end metabolite derived by a side reaction driven by the stability of canadensene (this work) and the ubiquitous presence of oxygenases. In fact, since the oxygenated substituents on canadensene have the same absolute configuration as in paclitaxel, these two natural products might involve the exact same oxygenases. Another biogenetic hypothesis of canadensene **4** is shown in Scheme 3B. It would derive from ring opening of an oxygenated tricyclic precursor probably originating from taxadiene **3**. This hypothesis is less plausible since the conversion of tricyclic taxanes to open-bicyclic structures has never been reported in any yew species. On the other hand, the U-shape of the *bicyclic* canadensene versus the planar *bicyclic* pre-taxadiene structures might imply some biosynthetic correlations between paclitaxel and canadensene. A third biogenetic possibility is that along the taxadiene pathway, there is a second pathway involving canadensene (which might be more prevalent in *T. canadensis*). This mechanism could be described by Scheme 3C with a different sequence of reactions. In one case (the taxadiene pathway), *cyclization* of the bicyclic hydrocarbon (**15-19**) led to taxadiene **3** (Scheme 3A), and in order to convert to paclitaxel **1** *oxygenations* are needed as well as *addition of the side chain.* The canadensene pathway (Scheme 3C: where the key intermediate would be canadensene **4** or some deacetylated derivative) would differ in that the

*cyclization* and *addition of the side chain* would occur only *after* the *oxygenations* had occurred. As opposed to the taxadiene pathway, we have seen that the cyclization process of the canadensene pathway would involve substrate and product (paclitaxel) of similar U-shape structures. Furthermore, based on our energy calculations, this substrate would be more easily amenable to cyclizations than that of the taxadiene pathway. The sequence




**Scheme 3**. Biosynthetic theories related to canadensene. The bold arrow represents a proven conversion.[4]

of events differs: it was first formed by *oxygenations* of a bicyclic hydrocarbon and in order to make paclitaxel *cyclization* and *addition of the side chain* should be the subsequent steps. This could be a very plausible mechanism for the biosynthesis of paclitaxel. It is possible that the true biosynthetic precursor is a hydroxylated derivative of canadensene **4**. This possibility can be supported by the difference in taxane composition between *T. canadensis* and other yews.[17,18] There may be two pathways for the formation of paclitaxel: the canadensene and the taxadiene pathways. Depending on the *Taxus* species, one of them could predominate. This situation would be analogous to the arogenate and the prephenate pathways in the biosynthesis of aromatic amino acids.[20,21] Labelling experiments with different yew species are in progress.

## ACKNOWLEDGMENTS

We thank the Natural Sciences and Engineering Research Council of Canada and the Canadian Breast Cancer Research Initiative for support of this work via operating grants to L. O. Z. The Fondation Armand-Frappier is acknowledged for a fellowship to G. C. We wish to acknowledge Dr. René Gagnon for the preparation of baccatin III according to published procedures, the technical assistance of Mr. Alain Larocque for the acquisition and analysis of NMR spectra as well as the assistance of Mr. Louis Senécal for computer maintenance and software development.

## REFERENCES


1. Schiff, P. B.; Fant, J.; Horwitz, S. B. *Nature* **1979**, *277*, 665-667.
2. Potier, P.; Guéritte-Voegelein, F.; Guénard, D. *J. Am. Chem. Soc.* **1992**, *203*, 1-3.
3. Nicolaou, K. C.; Yang, Z.; Liu, J. J.; Ueno, H.; Nantermet, P. G.; Guy, R. K.; Claiborne, C. F.; Renaud, J.; Couladouros, E. A.; Paulvannan, K.; Sorensen, E. J. *Letters to Nature* **1994**, *367*, 630-634. Holton, R. A.; Somoza, C.; Kim, H.-B.; Liang, F.; Biediger, R. J.; Boatman, P. D.; Shindo, M.; Smith, C. C.; Kim, S.; Nadizadeh, H.; Suzuki, Y.; Tao, C.; Vu, P.; Tang, S.; Zhang, P.; Murthi, K. K.; Gentile, L. N.; Liu, J. H. *J. Am. Chem. Soc.* **1994**, *116*, 1597-1598. Holton, R. A.; Kim, H. B.; Somoza, C.; Liang, F.; Biediger, R. J.; Boatman, P. D.; Shindo, M.; Smith, C. C.; Kim, S.; Nadizadeh, H.; Suzuki, Y.; Tao, C.; Vu, P.; Tang, S.; Zhang, P.; Murthi, K. K.; Gentile, L. N.; Liu, J. H. *J. Am. Chem. Soc.* **1994**, *116*, 1599-1600.
4. Koepp, A. E.; Hezari, M.; Zajicek, J.; Stofer Vogel, B.; LaFever, R. E.; Lewis, N. G.; Croteau, R. *J. Biol. Chem.* **1995**, *270*, 8686-8690.
5. Hezari, M.; Lewis, N. G.; Croteau, R. *Arch. Biochem. Biophys.* **1995**, *322*, 437-434.
6. Kingston, D.G.I.; Molinero, A.A.; Rimoldi, J.M. In *Progress in the Chemistry of Organic Natural Products* **1993**, *61*, pp. 154-155.
7. Zamir, L.O.; Zhou, Z.-H.; Caron, G.; Nedea, M. E.; Sauriol, F.; Mamer, O. *J. Chem. Soc., Chem. Commun.* **1995**, 529-530.
8. Zamir, L. O.; Caron, G. unpublished results. Begley, M. J.; Jackson, C. B.; Pattenden, G. *Tetrahedron Lett.* **1985**, *26*, 3397-3400.
9. Fang, W.-S.; Fang, Q.-C.; Liang, X.-T.; Lu, Y.; Zheng, Q.-T. *Tetrahedron* **1995**, *51*, 8483-8490.
10. Hilton, B. D.; Chmurny, G. N.; Muschik, G. M. *J. Nat. Prod.* **1992**, *55*, 1157-1161.
11. Williams, H. J.; Scott, A. I.; Dieden, R. A.; Swindell, C. S.; Chirlian, L. E.; Francl, M. M.; Heerding, J. M.; Krauss, N. E. *Tetrahedron* **1993**, *49*, 6545-6559. Williams, H. J.; Scott, A. I.; Swindell, R. A. C. S.; Chirlian, L. E.; Francl, M. M.; Heerding, J. M.; Krauss, N. E. *Can. J. Chem.* **1994**, *72*, 252-260.
12. Chmurny, G. N.; Hilton, B. D.; Brobst, S.; Look, S. A.; Witherup, K.; Beutler, J. A. *J. Nat. Prod.* **1992**, *55*, 414-423.
13. Falzone, C. J.; Benesi, A. J.; Lecomte, J. T. J. *Tetrahedron Lett.* **1992**, *33*, 1169-1172.
14. Baker, J. K. *Spectrosc. Lett.* **1992**, *25*, 31-48.
15. Dubois, J.; Guénard, D.; Guéritte-Voegelein, F.; Guedira, N.; Potier, P.; Gillet, B.; Beloeil, J.-C. *Tetrahedron* **1993**, *49*, 6533-6544.
16. Guéritte-Voegelein, F.; Mangatal, L.; Guénard, D.; Potier, P.; Guilhem, J.; Cesario, M.; Pascard, C. *Acta Crystallogr.* **1990**, *C46*, 781-784.
17. Zamir, L. O.; Nedea, M. E.; Bélair, S.; Sauriol, F.; Mamer, O.; Jacqmain, E.; Jean, F. I.; Garneau, F-X. *Tetrahedron Lett.* **1992**, *33*, 5173-5176.
18. Zamir, L. O.; Nedea, M. E.; Zhou, Z.-H.; Bélair, S.; Caron, G.; Sauriol, F.; Jacqmain, E.; Jean, F. I.; Garneau, F.-X.; Mamer, O. *Can. J. Chem.* **1995**, *73*, 655-665.
19. Zhang, S.; Chen, W. M.; Chen, Y. H.; *Yaoxue Xuebao* **1992**, *27*, 268-272.
20. Zamir, L. O.; Jensen, R. A.; Arison, B. H.; Douglas, A. W.; Albers-Schongerg, G.; Bowen, J. R. *J. Am. Chem. Soc.* **1980**, *102*, 4499-4504.
21. Jensen, R. A. In *The Evolution of Metabolic Function*; Mortlock R. P. Ed. C. R. C. Press, Inc.: Boca Raton, FL, **1992**, pp. 205-236.


(*Received in USA* 29 *February* 1996; *revised* 1 *May* 1996; *accepted* 6 *May* 1996)