Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 7, No. 18, pp. 2387-2392, 1997

0960-894X/97 $17.00 + 0.00

PII: S0960-894X(97)00439-3

3D QSAR ANALYSIS OF TAXOIDS FROM *TAXUS CUSPIDATA* VAR. *NANA* BY COMPARATIVE MOLECULAR FIELD APPROACH

Hiroshi Morita, Akira Gonda, Lan Wei, Koichi Takeya, and Hideji Itokawa*

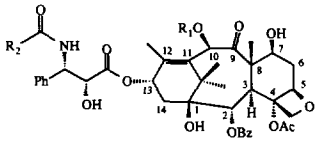
Department of Pharmacognosy, School of Pharmacy, Tokyo University of Pharmacy & Life Science, 1432-1 Horinouchi, Hachioji, Tokyo 192-03, Japan

Abstract: A series of taxoids, isolated from the stems of *Taxus cuspidata* var. *nana*, showing a cell growth inhibitory activity, were investigated for their 3D QSAR by using the comparative molecular field (CoMFA) analysis. The results indicated a strong correlation between the P388 cell growth inhibitory activity of these taxoids and the steric and electrostatic factors which modulate their biological activity, and accounted for the potent activities of the taxoids with the N-acylphenylisoserine derivatives at C-13 and the weak activities of those without this kind of ester group. It was also suggested that the substituent at C-2 is not a requisite for the biological activity: smaller substituent seemed to favor the biological activity.

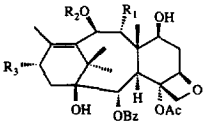
Paclitaxel (Taxol®), isolated from various species of the genus *Taxus* (Taxaceae), and docetaxel (Taxotere®), a semisynthetic analog, are highly promising anticancer agents especially for the treatment of advanced ovarian cancer.[1,2] They are also considered as the leading compounds in cancer chemotherapy. The mechanism of the action of these diterpenes is unique involving the facilitated assembly and stabilization of microtubules.[3]

Extensive studies on syntheses of taxol analogs and their structure-activity relationships (SAR) have been reported,[1] but SAR of natural taxoids without N-acylphenylisoserine moiety has not. Recently, the structures of many taxoids from *Taxus cuspidata* and varieties of their biological activities, such as cytotoxicity, reduction of $CaCl_2$-induced depolymerization of microtubules, and the enhanced cellular accumulation of vincristine in multidrug-resistant tumor cells as potent as verapamil, have been reported.[4] Among them, there are some taxoids, showing cytotoxicity, but having neither an oxetane ring nor N-acylphenylisoserine moiety, which have been considered to be requisite for cytotoxicity.[4]

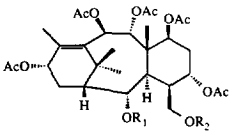
Previously, we conducted a thorough analysis of taxoids contained in the stem of *Taxus cuspidata* Sieb. et. Zucc.var. *nana* Rehder. (Taxaceae), a popular garden tree in Japan,[5] and by chromatographic purification of its taxoid fraction isolated of 35 cytotoxic taxoid compounds.[6] These taxoids belong to either of the two skeletal taxoids including 6-8-6 and 6-10-6 ring systems. Some of the taxoids with an N-acylphenylisoserine ester group at C-13 showed a potent cytotoxicity and some did nor. In the present study, we investigated these taxoids by using comparative molecular field analysis (CoMFA). Here we report relationships between the structure of these taxoids and their growth inhibitory activity on P388 cells by CoMFA.



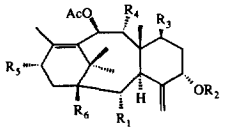
1 R_1=Ac, R_2=Ph (taxol)
2 R_1=H, R_2=Ph
3 R_1=Ac, R_2=C_4H_7
4 R_1=H, R_2=C_4H_7
5 R_1=Ac, R_2=*n*-C_5H_{11}
6 R_1=H, R_2=*n*-C_5H_{11} C-7 epimer
7 R_1=H, R_2=*n*-C_5H_{11}
8 R_1=Ac, R_2=C_6H_{13}
9 R_1=Ac, R_2=*n*-C_3H_7



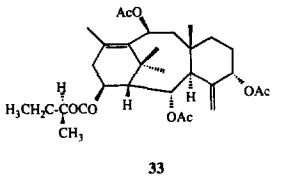
10 R_1=O, R_2=Ac, R_3=OH
11 R_1=OH, R_2=Ac, R_3=OAc
12 R_1=OH, R_2=H, R_3=O
13 R_1=OH, R_2=H, R_3=OAc
14 R_1=OH, R_2=Ac, R_3=O



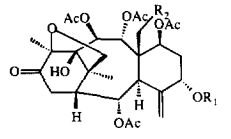
29 R_1=Ac, R_2=H
30 R_1=H, R_2=Ac



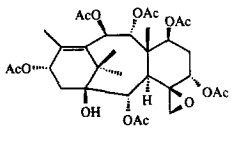
15 R_1=H, R_2=H, R_3=OBz, R_4=OAc, R_5=OH, R_6=OH
16 R_1=H, R_2=Ac, R_3=H, R_4=OAc, R_5=OAc, R_6=H
17 R_1=H, R_2=COCH=CHPh, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H
18 R_1=OAc, R_2=COCH=CHPh, R_3=H, R_4=OAc, R_5=O, R_6=H
19 R_1=OAc, R_2=Ac, R_3=H, R_4=OAc, R_5=OAc, R_6=H
20 R_1=H, R_2=Ac, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H
21 R_1=H, R_2=H, R_3=OAc, R_4=OAc, R_5=OH, R_6=H
22 R_1=OCOC$_4$H$_9$, R_2=Ac, R_3=OAc, R_4=H, R_5=H, R_6=H
23 R_1=H, R_2=H, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H
24 R_1=OAc, R_2=COCH=CHPh, R_3=OAc, R_4=OAc, R_5=O, R_6=H
25 R_1=OAc, R_2=H, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H
26 R_1=H, R_2=H, R_3=OAc, R_4=OAc, R_5=O, R_6=H
27 R_1=OAc, R_2=H, R_3=OAc, R_4=OAc, R_5=O, R_6=H
28 R_1=OH, R_2=COCH=CHPh, R_3=OAc, R_4=OAc, R_5=OAc, R_6=H



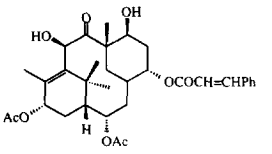
33



31 R_1=H, R_2=OBz
32 R_1=COCH=CHPh, R_2=H



34



35

Taxoids

The taxoids used in this experiment were prepared as described previously[5] or identified with the authentic samples.[6]

Computational Methods

Sybyl molecular modeling software was used for the CoMFA approach. Structure optimization was performed by using the systematic pseudo Monte Carlo (MC) search,[7] in the MacroModel (version 4.5) on an IRIS 4D computer *in vacuo*. Each conformation generated by each MC calculation (3000 step) was minimized by the use of molecular mechanics calculation of MM2* force field[8] to reduce the gradient rms to less than 0.001 kcal/Å·mol with a distance-dependent dielectric, $\varepsilon=R_{ij}$. To eliminate possible duplicate conformations, comparison was performed on the heavy atoms only. The extended cutoff distances employed were 8 Å for van der Waals, 20 Å for charge/electrostatics and 10 Å for charge/multipole electrostatics. The stable conformation of the N-acylphenylisoserine side chain at C-13 of taxol-type compounds was selected by the clustering analysis according to their hydrophobic collapse conformation as a consequence of hydrophobic interaction of nonpolar groups at 2-benzoyl, 3'-phenyl, and 4-acetyl groups.[9] The taxoids with the abeotaxane skeleton such as taxuspinanane B and taxuspinanane F from the same source,[5] were not included in the present CoMFA analysis, because of their conformational flexibility.[10]

CoMFA was then initiated by using the minimum-energy conformations obtained as described above. The default SYBYL settings were used unless otherwise noted. These compounds were aligned *via* root mean square (rms) fitting of backbone atoms, C1 to C14, to the corresponding atoms of taxol chosen as the template

molecule. In choosing these particular atoms for the alignments, care was taken so that optimal fit among the compounds and also the minimal loss of conformational flexibility are to be provided. Partial atomic charges required for calculation of the electrostatic interaction energies were calculated by using the PM3 method.[11]

PLS and CoMFA Analysis.

For each of the alignment sets, the steric and Coulombic potential energy fields were individually calculated at each lattice intersection of a regularly spaced grid of 2.0Å units in all x, y, and z directions. The steric terms represent the van der Waals interactions, whereas the Coulombic terms represent the electrostatic interactions for which a distance-dependent dielectric expression $\varepsilon=R_{ij}$ was adopted. The grid pattern was generated automatically by the SYBYL/CoMFA routine. An sp^3 carbon atom with a +1.0 charge was selected as a probe for the calculation of the steric and electrostatic field. Values of the steric and electrostatic energies were truncated at 30 kcal/mol.

To obtain a 3D-QSAR, partial least squares (PLS) fitting was used. The PLS method has been applied to rationalize those structural features affecting the biological activity. The PLS algorithm was initially used with the cross-validation option to obtain the optimal number of components needed for the subsequent analysis of the data. In the leave-one-out cross-validation, each compound was systematically excluded from the set and its activity predicted by a model derived from the rest of the compounds. The optimal number of components (n=5) was then chosen so that it would yield either the smallest rms error or the largest cross-validated r^2 value. A final PLS analysis was then performed by using the reported optimum number of components, with no cross-validation. This generated a fitted correlation of the entire training set with conventional r^2 values. The steric and electrostatic fields were scaled according to the CoMFA standard deviations in order to give them the same potential weights on the resulting QSAR. The 3D-QSAR calibration model so derived was then employed to predict the inhibitory values of some taxoids in the test set, after conformational analysis and alignments were performed by the method as indicated above.

Results and Discussion

The results of the CoMFA analysis are summarized in Table 1, and their actual and predicted activities and residuals are in Table 2. It is evident that the CoMFA-derived QSAR gave good cross-validated r^2, indicating thereby a considerably reliable capacity of this method for predicting their growth inhibition of the P388 leukemia cells. The relative contributions of the steric and electrostatic fields to this model were almost equal. As already known, the cytotoxicities of the taxoids are influenced by the presence of N-acylphenylisoserine ester at C-13. The N-acylphenylisoserine ester at C-13 is considered to enhance the drug-receptor interactions: in the case of Taxol and Taxotere, the taxane ring system is recognized first by the receptor, followed by additional interactions of the binding site with the side chain.[12] Our present analysis, which well correlate for both the activities of taxoids with N-acylphenylisoserine ester at C-13 and inactivities of those without it, is one of the strong pieces of evidence to support the above assumption. Evidence for the reliability of estimate by the CoMFA-derived model is provided in Fig. 1, which shows plots of actual versus predicted growth inhibition [log ($1/IC_{50}$)].

The major steric and electrostatic features of the QSAR are represented in Fig. 2 in the form of three-dimensional contour maps, displayed as transparent surfaces. The surfaces at the top of Fig. 2 indicate the areas in space around the template molecule (**1**) whose increases (green region) and/or decreases (yellow region) in the

steric bulk could enhance the growth inhibition of P388 cells. Contour maps of the electrostatic field contributions provided lower in Fig. 2.

Table 1. Summary of CoMFA-PLS results

opt. no. of components	5
probe atom	c (sp^3, +1)
cross-validated r^2	0.818
std error of estimate	0.196
r^2	0.979
F values (n1=5, n2=29)	267.621
contributions	
steric	0.520
electrostatic	0.480

Table 2. Actual and predicted activities and residuals by CoMFA analysis.

	log 1/IC_{50}		
compounds	actual	predicted	residual
1	1.66	1.48	0.18
2	1.37	1.41	-0.04
3	0.77	0.84	-0.09
4	1.18	1.26	-0.08
5	1.09	1.27	-0.19
6	1.05	1.14	-0.09
7	1.39	1.38	0.01
8	1.74	1.54	0.20
9	0.77	0.66	0.11
10	-1.20	-1.02	-0.18
11	-1.28	-1.38	0.10
12	-1.00	-1.27	0.27
13	-1.54	-1.39	-0.15
14	-1.32	-1.29	-0.04
15	-1.60	-1.63	0.03
16	-0.34	-0.80	0.46
17	-0.64	-0.49	-0.15
18	-2.00	-2.05	0.05
19	-1.78	-1.42	-0.36
20	-0.62	-0.64	0.01
21	-1.20	-1.31	0.11
22	-0.48	-0.46	-0.02
23	-1.36	-1.56	0.20
24	-2.00	-1.96	-0.04
25	-1.90	-2.09	0.19
26	-1.91	-1.42	-0.49
27	-1.18	-1.46	0.28
28	-0.59	-0.53	-0.06
29	-1.85	-1.79	-0.06
30	-1.91	-1.94	0.03
31	-1.57	-1.39	-0.18
32	-2.00	-1.95	-0.05
33	-0.64	-0.65	0.01
34	-2.00	-1.97	-0.03
35	-1.32	-1.40	0.08

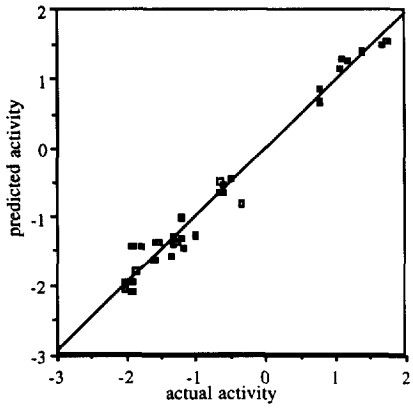


Fig. 1. Plot of actual vs predicted inhibitory activities on P388 cells (r^2=0.979). The activities are expressed by log 1/IC_{50}.

In the steric CoMFA map, as expected, the side chain at C-13 would enhance the growth inhibition. In addition, that at C-5, such as cinnamoyl moiety had a small contribution to the action. Surprisingly, the side chain at C-2 of a smaller bulk would enhance the activity. In spite of extensive SAR studies on paclitaxel and docetaxel, little has been reported on the modification at the C-2 benzoate group. The present result suggesting that the ester function at C-2 with a smaller bulk function is more effective for the growth inhibition of P388 cells agrees with recent study by Ojima et al. that the taxoids derivatives with a less bulk at C-2 than that of benzoyl group displayed more pronounced effects in the *in vitro* cytotoxicity assay.[13] Whereas, 2-deoxy taxol was known to possess much reduced activity with respect to taxol.[14] Some hydrophobic group at C-2 has considered to be requisite for hydrophobic collapse conformation.[9] This map also indicated that smaller bulks at C-7 and C-9 positions would also enhance the activity. The 6-8-6 skeletal taxoids (**15-28** and **31-33**) with exo-methylene moiety at C-4 showed a range of IC_{50} values, as shown in Table 2. This indication might explain why **24** and **28**, both including similar esteric moieties except for C-2 and C-13, exhibited different values of growth inhibition (**24**: >100 μg/ml; **28**: 3.9 μg/ml).

In the electrostatic CoMFA map, the regions where a more positive electrostatic interaction would improve the growth inhibition are encompassed around the N-acylphenylisoserine ester at C-13, whereas, those where negative electrostatic interactions are locating around each of the ester side chains and oxetane rings.

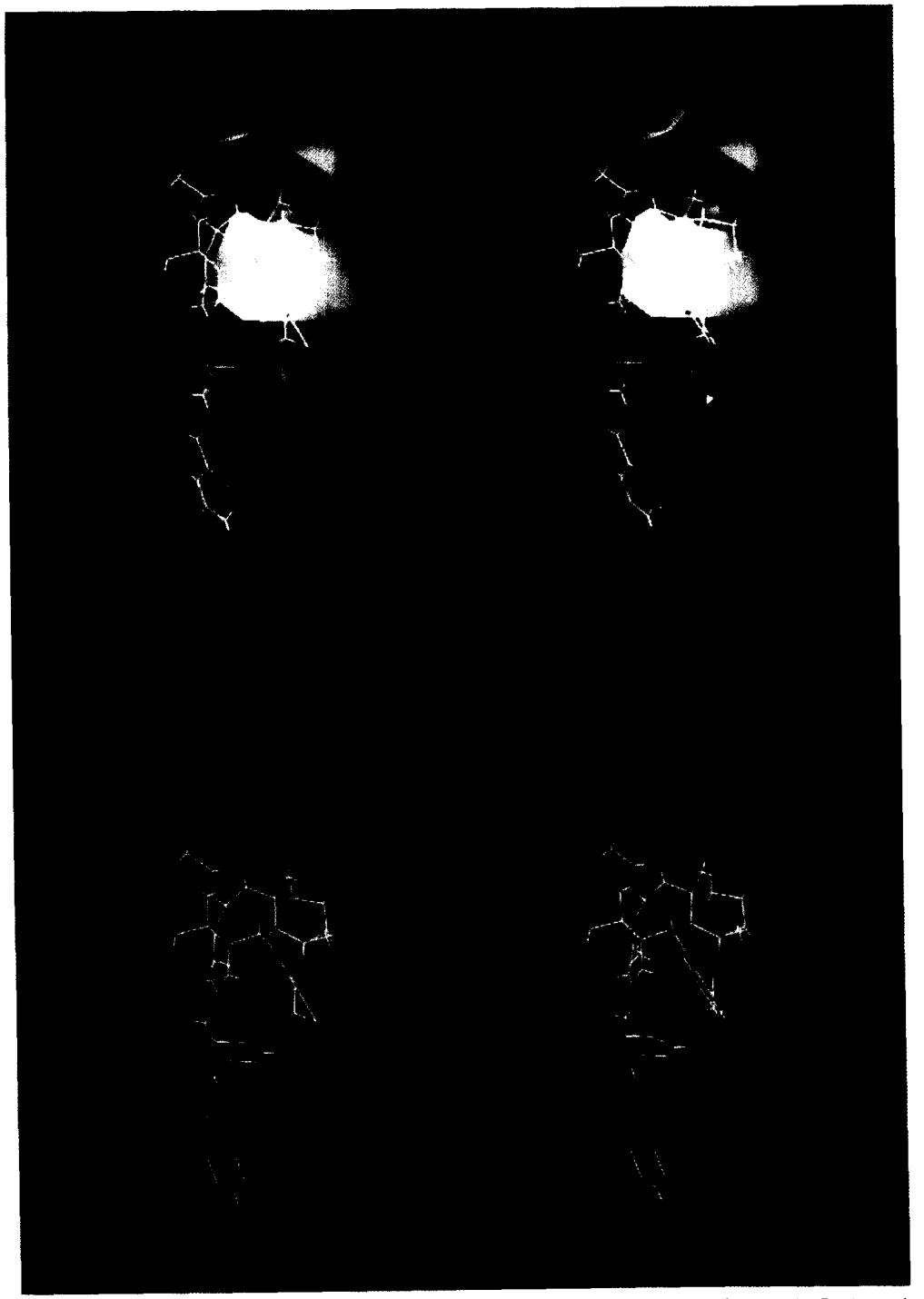


Fig. 2 Stereoview of the CoMFA contour plot from the PLS analysis. Above: steric, below: electrostatic. Regions where increased steric bulk is associated with enhanced activity are indicated in green, whereas regions where increased steric bulk is associated with diminished activity are indicated in yellow. On the other hand, regions where increased positive charge is favorable for activity are indicated in blue, whereas regions where increased negative charge is favorable are indicated in red.

The CoMFA model derived from the above taxoids was used to calculate the IC_{50} values of the test taxoids having a cyclohexyl substitution at C-3' and/or C-2 of Taxotere, and the obtained data were compared with the reported values.[15] The observed and corresponding CoMFA-predicted IC_{50} values for these test taxoids were as follows (observed:predicted): docetaxel: (2.10:1.16), 3'-dephenyl-3'-cyclohexyldocetaxel (1.20:0.94), 2-(hexahydro) docetaxel (1.12:0.94), 3'-dephenyl-3'cyclohexyl-2-(hexahydro)docetaxel (1.12:0.94), and 2-(hexahydro)paclitaxel (0.35:0.84). The result shows that the present model can be generally used for estimation of the activities of active derivatives (predictive r^2 0.908), though inactive derivatives should also be included in the calculation.

Thus, the CoMFA results have been applied successfully to rationalize the cell growth inhibitory activities of the taxoids from *T. cuspidata* var. *nana* in terms of their steric and electrostatic properties, and it became clear that similar correlation may be observed in both the taxoids having an N-acylphenylisoserine group at C-13 and those having no ester group at C-13. The predictive utility of the respective CoMFA-derived models is evident in the good results. It is distinctly possible that a combination of various substituents and conformation of the whole molecule may occur potency of inhibition in growth of leukemia cells. Estimation of biological action by the use of CoMFA derived models may further be improved if more substituents are taken into calculation.

Acknowledgments: We thank the Ministry of Education, Science and Culture, Japan, for financial support through Grant-in-Aid for General Scientific Research, and M. Katsuyama and K. Saitou, Sumisho Electronics Co., Ltd., for support of the molecular simulation study.

References and Notes

1. Kingston, D.G.I.; Molinero, A.A.; Rimoldi, J.M. *Progress in the Chemistry of Organic Natural Products*, **1993**, *61*, 1-206; Suffness, M., Ed. *In Taxol®: Science and Applications*; CRC Press: Boca Raton, Fl, 1995.
2. Swenerton, K.; Eisenhauer, E.; ten Bokkel Huinink, W. *Proc. Am. Soc. Clin. Oncol.*, **1993**, *12*, 256.
3. Schiff, P. B.; Fant, J.; Horwitz, S. B. *Nature*, **1979**, *277*, 665-667.
4. Kobayashi, J.; Hosoyama, H.; Wang, X.-x.; Shigemori, H.; Koiso, Y.; Iwasaki, S.; Sasaki, T.; Naito, M.; Tsuruo, T. *Bioorg. Med. Chem. Lett.*, **1997**, *7*, 393-398.
5. Morita, H.; Gonda, A.; Wei, L.; Yamamura, Y.; Takeya, K.; Itokawa, H. *J. Nat. Prod.*, **1997**, *60*, 390-392; Morita, H.; Wei, L.; Gonda, A.; Takeya, K.; Itokawa, H. *Phytochemistry*, **1997**, in press; Morita, H.; Gonda, A.; Wei, L.; Yamamura, Y.; Wakabayashi, H.; Takeya, K.; Itokawa, H. *Planta Medica*, accepted.
6. Taxoids isolated are as follows: taxol (**1**), 10-deacetyl taxol (**2**), taxol B (**3**), 10-deacetyl taxol B (**4**), taxol C (**5**), 7-epi-10-deacetyl taxol C (**6**), 10-deacetyl taxol C (**7**), taxuspinanane A (**8**), taxol D (**9**), baccatin III (**10**), 9-dihydro-14-acetyl baccatin III (**11**), taxuspinanane C (**12**), 7, 9, 10-deacetyl baccatin VI (**13**), taxuspinanane D (**14**), brevifoliol (**15**), taxusin (**16**), 2α-deacetoxy taxinine J (**17**), taxinin (**18**), taxa-4(20), 11-diene-2α, 5α, 9α, 10β, 13α-pentaol pentaacetate (**19**), taxa-4 (20), 11-diene-5α, 7β, 9α, 10β, 13α-pentaol pentaacetate (**20**), taxa-4 (20), 11-diene-5α, 7β, 9α, 10β, 13α-pentaol 7β, 9α, 10β-triacetate (**21**), 2α-α-methyl butyryloxy-5α-7β, 10β-triacetyl-(4), 20, 11-taxadine (**22**), taxa-4(20), 11-diene-5α, 7β, 10β, 13α-pentaol 7β, 9α, 10β, 13α tetra-acetate (**23**), taxinin B (**24**), decinnamoyl taxinine J (**25**), taxuspinanane K (**26**), taxuspine F (**27**), taxuspinanane G (**28**), taxuspine L (**29**), taxchin A (**30**), taxinine M (**31**), taxagifine (**32**), taxa-4 (20), 11-taxadiene-2α, 5α, 10β, 14β-tetraol 2α, 5α, 10β-triacetate-14β-(s)-2'-methyl butyrate (**33**), 1β-hydroxy-baccatin I (**34**), and taxuspinanane H (**35**).
7. Goodmann, J.M.; Still, W.C. *J. Comput. Chem.*, **1991**, *12*, 1110-1117.
8. Mohamadi, F.; Richards, N.G.J.; Guida, W.C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W.C. *J. Comput. Chem.*, **1990**, *11*, 440-467.
9. Velde, D.G.V.; Georg, G.I.; Grunewald, G.L.; Gunn, C.W.; Mitscher, L.A. *J. Am. Chem. Soc.* **1993**, *115*, 11650-11651.
10. Fuji, K.; Tanaka, K.; Li, B.; Shingu, T.; Yokoi, T.; Sun, H.; Taga, T. *Tetrahedron* **1995**, *51*, 10175-10188.
11. Stewart, J.J.P. *J. Computational Chem.*, **1989**, *10*, 221-264.
12. Dubois, J.; Guenard, D.; Guerrite-Voegelein, F.; Guedira, N.; Potier, P.; Gillet, B.; Beloeil, J.-C. *Tetrahedron* **1993**, *49*, 6533-6544.
13. Ojima, I.; Kuduk, S. D.; Pera, P.; Veith, J. M.; Bernacki, R. J. *J. Med. Chem.* **1997**, *40*, 279-285.
14. Chen, S.-H.; Wei, J.-M.; Farina, V. *Tetrahedron Lett.* **1993**, *34*, 3205-3206.
15. Ojima, I.; Duclos, O.; Zucco, M.; Bissery, M.-C.; Combeau, C.; Vrignaud, P.; Riou, J. F.; Lavelle, F. *J. Med. Chem.* **1994**, *37*, 2602-2608.

(*Received in Japan* 7 *July* 1997; *accepted* 21 *August* 1997)