

Hansch analysis of anticancer activities of C₂-modified paclitaxel analogs against human ovarian carcinoma 1A9, human colon carcinoma HCT116 and human burkitt lymphoma CA46 cell lines

Kunal Roy^{*†}, Dipak Kumar Pal[‡], A U De[‡] & Chandana Sengupta[†]

Division of Pharmaceutical Chemistry, Seemanta Institute of Pharmaceutical Sciences, Jharpokharia, Mayurbhanj 757 086, Orissa, India

[†]QSAR Lab, Division of Medicinal & Pharmaceutical Chemistry, Jadavpur University, Calcutta 700 032, India

Received 15 April 1999; accepted 20 August 1999

Anticancer activities, recently reported by Kingston *et al.*, of paclitaxel and its thirty analogs, having substitution(s) on the phenyl ring of 2-benzoyloxy moiety, against human ovarian carcinoma 1A9, human colon carcinoma HCT116 and human burkitt lymphoma CA46 cell lines have been subjected to QSAR analyses by extrathermodynamic approach of Hansch. The aims of the study are to explore the effects of the phenyl ring substituents on the activities and to unravel the quantitative relations of the activities with physico-chemical parameters of the substituents.

For the activity against ovarian carcinoma, presence of *meta* and absence of *para* substituents on the phenyl ring (with respect to the ester linkage on the 2-position of taxane skeleton) has been found to be conducive to the activity. Molar refractivity and electronic parameter (Hammett σ) of the *meta* substituents show parabolic relationships with the activity. In case of the activity against colon carcinoma, *para* substitution is not preferred. However, lipophilicity of the *para* substituent, if present, has been found to favour to the activity which was further negatively contributed by square term of molar refractivity of the *meta* substituents. In the relation best obtained, the activity against burkitt lymphoma is found to be negatively contributed by square term of molar refractivity of the *meta* substituents and those of electronic parameters (Hammett σ) of the *para* and *meta* substituents. The differences among the equations relating activity against three cell lines with physico-chemical parameters suggest that there might be some differences among the receptor structures.

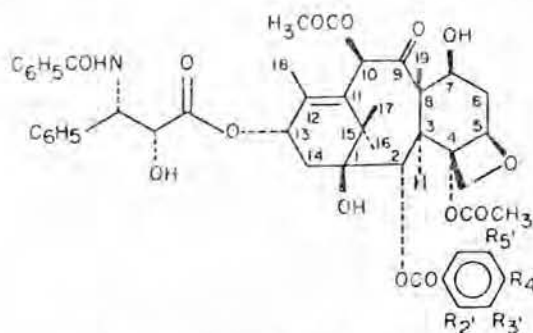
In the ongoing process of exploring potential of natural products for the cure of the deadly diseases like cancer and AIDS, taxoids have gained significant attraction in recent years for their promising anticancer activity. Paclitaxel (Taxol), a naturally occurring diterpenoid, is now established as a broad-spectrum antileukemic and antitumor agent¹.

In 1960s, during the National Cancer Institute's screening programme for antineoplastic agents from plant source, extracts of the bark of Pacific Yew, *Taxus brevifolia*, were found to possess activity against a number of cell lines². The active principle was isolated in 1967 and named Taxol. Its structure was characterised in 1971³. Paclitaxel (**1**) is a highly functionalised diterpenoid having taxane skeleton with four ester groups and one oxetane ring. It is 5 β -20-epoxy-1,2 α ,4,7 β ,10 β ,13 α -hexahydroxytax-11-en-9-one-4,10-diacetate-2-benzoate-13-ester with (2*R*,3*S*)-*N*-benzoyl-3-phenylisoserine.

Initially paclitaxel was not considered as a

promising clinical candidate because of its low yield, limited aqueous solubility and only modest *in vivo* activity in various leukemia assays⁴. Later in 1970s, with the development of new assay methods, encouraging positive results were obtained in new *in vivo* models and paclitaxel entered into preclinical studies in the late '70s. It reached the clinics more than two decades after its discovery.

Apart from *T. brevifolia*, paclitaxel has also been isolated from *T. baccata*, *T. cuspidata*, *T.*



1. Paclitaxel; R₂' = R₃' = R₄' = R₅' = H

2. C₂-modified analogs of paclitaxel

^{*}Address for correspondence: C/o Dr. A G Saha, Flat No. 2E, 8, Dr Ashutosh Sastry Road, Calcutta 700 010 (India)

wallichiana, *T. chinensis*, *T. canadensis*, *T. media*, *T. yunnanensis* and *Taxomyces adreanae*⁵. Paclitaxel has a unique mode of action. It inhibits cell division by promoting assembly of stable microtubule bundles which leads to cell death². Now paclitaxel is approved for treatment of ovarian and breast cancer⁶. It also showed encouraging activity against P-388, P-1534 and L-1210 mouse leukemia, B-16 melanocarcinoma, Walker 256 carcinoma, Lewis lung carcinoma, sarcoma 180 and CX-1 colon, LX-1 lung xenographs^{4,5}. It also has potential for treatment of skin, head and neck cancer⁶. The slow development of paclitaxel as a drug is due to the difficulty of its isolation, high toxicity and its lipophilicity which makes formulation difficult⁵.

To overcome the shortcomings of low aqueous solubility and to reduce undesirable side effects, chemical modifications of paclitaxel have been attempted by the scientists. There are many reports on structure activity relation studies of paclitaxel derivatives^{7,8}. It has been concluded that the C₁₃ ester side chain and the oxetane ring are essential for the activity⁶. The six membered ring fused with the oxetane ring⁹ and C₁₁-C₁₂ double bond¹⁰ are also required for the activity. Expansion of the six membered ring to seven membered ring or opening of the oxetane ring results in loss of activity. The α -benzoyloxy functionality at C₂ present in paclitaxel appears to play an essential role in binding of paclitaxel to its receptor¹¹. The acetyloxy group on C₄ also appears to play significant role for anchoring hydrophobically clustered conformation in its proper orientation¹². The hydroxy group at C₇ position¹³ and acetate group at C₁₀ position¹⁴ contribute very little in the receptor binding. Addition of oxygenated substituent at C₆ does not significantly improve the activity¹⁵. The removal of C₉-carbonyl group results in equally active 9-deoxypaclitaxel. However, simultaneous removal of C₁₀-acetoxy, C₉-carbonyl and C₇ hydroxy groups results in less active compound¹⁶. Water soluble esters of 2'-hydroxyl group in the ester side chain with impressive biological action have been reported. Taxotere, which is approximately twice as potent as paclitaxel and more water soluble, possesses a *tert*-butyloxycarbonyl moiety in place of the benzoyl group as in paclitaxel and acetyloxy group at C₁₀ is replaced by a hydroxy group¹⁷.

Recently, the benzoyloxy moiety at C₂ of paclitaxel has been the target of modifications for Quantitative

Structure Activity Relationship (QSAR) studies^{18,19}. Syntheses and evaluation of antineoplastic activities of 2-debenzoyl-2-aryloxy-paclitaxels have been very recently reported by Kingston *et al.*⁶. The present communication is an attempt to give a quantitative basis of the relations between changes in the activities of such analogs of paclitaxel and changes in their physicochemical or structural parameters.

Materials and Methods

Kingston *et al.*⁶ reported cytotoxicity of C₂-modified paclitaxel analogs against three human carcinoma cell lines, viz., human ovarian carcinoma 1A9, human colon carcinoma HCT116 and human burkitt lymphoma CA46 lines. Cytotoxicity data of a total of thirty four compounds were published in the communication⁶, though quantitative data of some of the compounds were not available for one or two cell line(s). Four compounds containing heterocyclic ring containing substituents at C₂ position were excluded in the present theoretical study. The rest thirty compounds differing from paclitaxel only in having substitution(s) on the phenyl ring of the 2-benzoyloxy moiety were subjected to QSAR study, along with paclitaxel itself, by multiple regression analyses with different physico-chemical and structural parameters, following linear free energy - related approach of Hansch²⁰, to explore the effects of the substituents on the activity.

Table I shows different substituents present on C₂-benzoyloxy moiety of the paclitaxel analogs (2) considered in the present study and their relative activity in comparison to that of paclitaxel. The relative activity value of each compound is expressed as the ratio of IC₅₀ value of the compound to that of paclitaxel for a particular cell line. Hence, any compound having relative activity less than unity is more cytotoxic than paclitaxel against a particular cell line. The average IC₅₀ values of paclitaxel were 2.1 nM, 2.0-2.5 nM and 30 nM against ovarian carcinoma, colon carcinoma and burkitt lymphoma cell lines respectively. For the present QSAR study, relative activity values were converted to the negative logarithm of IC₅₀ values (pC) which were then subjected to multiple regression analyses (MRA) with different physico-chemical parameters including electronic (Hammett σ), lipophilicity (π) and steric (MR, molar refractivity) terms. Some dummy parameters and indicator variables encoding some structural features which could not be described by

Table I—Relative activity of C₂ modified analogs (2) of paclitaxel in comparison to paclitaxel (1) against different human carcinoma cell lines

Compd	Substitutions				Relative activity* in different cell lines		
	R ₂	R ₃	R ₄	R ₅	OC	CC	BL
1	OCH ₃	H	H	H	1.4	1.9	0.67
2	N ₃	H	H	H	0.77	0.36	—
3	H	OC ₂ H ₅	H	H	1.7	1.6	1.0
4	H	OPr ⁱ	H	H	15	17	1.3
5	H	OCH ₃	H	H	0.095	0.35	0.27
6	H	SCH ₃	H	H	3.3	—	1.0
7	H	N ₃	H	H	0.18	1.4	0.17
8	H	F	H	H	0.40	0.45	—
9	H	I	H	H	1.6	0.85	2.3
10	H	Cl	H	H	0.05	1.7	0.67
11	H	Br	H	H	0.32	0.32	—
12	H	CF ₃	H	H	1.4	3.0	0.67
13	H	CN	H	H	0.32	1.1	—
14	H	NO ₂	H	H	1.3	0.88	1.0
15	H	H	OH	H	31	52	3.0
16	H	H	Cl	H	7.6	—	1.3
17	F	F	H	H	14	0.55	—
18	F	H	H	F	0.67	0.44	0.67
19	F	H	F	F	3.1	1.9	1.0
20	OCH ₃	OCH ₃	H	H	>48	59	>3.3
21	OCH ₃	H	H	OCH ₃	0.12	0.14	0.67
22	H	OCH ₃	OCH ₃	OCH ₃	>48	110	>3.3
23	H	OCH ₃	H	OCH ₃	3.3	5.4	1.0
24	H	F	F	H	3.4	2.7	0.33
25	H	N ₃	H	N ₃	1.3	1.6	1.0
26	H	Cl	Cl	H	1.0	0.77	—
27	H	F	H	F	0.12	0.36	—
28	H	F	F	F	3.2	3.4	1.3
29	H	Cl	H	Cl	0.38	0.51	0.67
30	H	NO ₂	H	NO ₂	20	3.1	2.3

* Relative activity values are expressed as ratios of IC₅₀ values of the compounds to that of paclitaxel.

OC = Ovarian Carcinoma, CC = Colon Carcinoma, BL = Burkitt Lymphoma

Prⁱ = isopropyl

continuous variables were also included in the MRA as predictor variables. Considering the average IC₅₀ value of paclitaxel against colon carcinoma to be 2.25 nM, the corresponding biological activity (pC₂) was calculated. The cytotoxicity data of the compounds are listed in Table II. The values of the physico-chemical parameters (Table III) were taken from literature²¹.

The following general pattern was followed while deriving an equation best describing the changes in a particular activity with changes in physicochemical property and structural features. On examination of

the dataset, it was revealed that majority of the compounds possessed *meta* - substituents. There were only few compounds having only *para* or only *ortho* substituents. The relative positions of the substituents (*ortho*, *meta* and *para*; or, 2', 3' and 5' and 4') were determined with reference to the ester linkage joining the phenyl ring to the taxane skeleton. At the beginning of MRA with each activity dataset, only mono-*meta* substituted compounds were considered to identify the parameters best describing the changes in activity. Next were included the compounds with di-*meta* substituents. After the relation of the

Table II—Observed and calculated biological activity (cytotoxicity) against three human carcinoma cell lines of paclitaxel and its C₂-modified analogs

Compd	Biological activity (-log IC ₅₀) against different cell lines								
	pC ₁ ^a			pC ₂ ^b			pC ₃ ^c		
	Obs.	Calc.	Res.	Obs.	Calc.	Res.	Obs.	Calc.	Res.
1	-0.468	-0.387	-0.082	-0.631	-0.015	-0.616	-1.303	-1.350	0.047
2	-0.209	-0.387	0.178	0.092	-0.015	0.106	—	—	—
3	-0.553	-0.205	-0.348	-0.556	-0.545	0.082	-1.477	-1.418	*-0.059
4	-1.498	-1.571	0.072	-1.583	-1.158	-0.425	-1.591	-1.472	-0.119
5	0.700	0.483	0.217	0.104	-0.280	0.384	-0.909	—	—
6	-0.841	-0.548	-0.293	—	—	—	1.477	-1.435	-0.042
7	0.423	0.161	0.261	-0.498	-0.446	-0.052	-0.708	—	—
8	0.076	0.143	-0.068	-0.005	-0.013	0.009	—	—	—
9	-0.526	-0.677	0.150	-0.282	-0.793	0.511	-1.839	—	—
10	0.979	0.452	0.527	-0.583	-0.176	-0.406	-1.302	-1.392	0.089
11	0.173	0.257	-0.084	0.143	-0.347	0.490	—	—	—
12	-0.468	—	—	-0.829	—	—	-1.302	-1.396	0.093
13	0.173	0.275	-0.103	-0.394	-0.192	-0.202	—	—	—
14	-0.436	0.048	-0.484	-0.297	-0.249	-0.048	-1.477	-1.467	-0.010
15	-1.814	-1.356	-0.457	-2.068	-0.270	0.101	-1.954	-1.962	0.008
16	-1.203	-1.356	0.153	—	—	—	-1.591	-1.587	-0.004
17	-1.468	—	—	-0.093	—	—	—	—	—
18	-0.148	0.143	-0.292	0.004	-0.013	0.018	-1.303	-1.371	0.068
19	-0.814	-0.826	0.013	-0.631	-0.873	0.242	-1.477	-1.387	-0.090
20	—	—	—	-2.123	—	—	—	—	—
21	0.599	0.483	0.116	0.502	—	—	-1.303	-1.381	0.078
22	—	—	—	-2.394	-1.992	-0.402	—	—	—
23	-0.841	—	—	-1.085	-0.876	-0.208	-1.477	-1.451	-0.023
24	-0.854	-1.259	0.405	-0.784	-0.968	0.185	-0.996	—	—
25	-0.436	—	—	-0.556	-0.736	0.180	-1.477	-1.557	0.080
26	-0.322	-0.518	0.196	-0.239	-0.125	-0.114	—	—	—
27	0.599	0.451	0.147	0.092	-0.012	0.103	—	—	—
28	-0.827	-0.518	-0.309	-0.884	-0.872	-0.012	-1.591	-1.450	-0.141
29	0.098	0.226	-0.128	-0.060	-0.514	0.455	-1.303	-1.502	0.199
30	-1.623	-1.770	0.146	-0.844	-0.766	-0.077	-1.839	-1.796	-0.042
31*	-0.322	-0.387	0.064	-0.352	-0.015	-0.337	-1.477	-1.350	-0.127

Obs. = Observed, Calc. = Calculated, Res. = Residual = Obs. - Calc.

a. Calculated according to Eq. OC7; Compd. Nos. 12, 17, 23 & 25 were not included in MRA while deriving Eq. OC7

b. Calculated according to Eq. CC6; Compd. Nos. 12, 17, 20 & 21 were not included in MRA while deriving Eq. CC6

c. Calculated according to Eq. BL6; Compd. Nos. 5, 7, 9 & 24 were not included in MRA while deriving Eq. BL6

* Compound 31 is paclitaxel

biological activity of the *meta* substituted compounds with the selected parameters best fitted as the predictor variables were obtained, the compounds having *para* substituents (singly, or in addition to *meta*,) were also included in the dataset to explore the contribution pattern of the *para* substitution to the response variable (pC). Finally, all compounds, quantitative activity data of which were known, were

considered to derive the final equations. The compounds having structural features with skewed distribution pattern were excluded as a sound statistical practice. Finally analysis of residuals was applied to exclude the outliers or the compounds showing aberrant behaviour to improve statistical quality of the equations. The quality of the equations was judged by examining the parameters like EV%

Table III—Different physico-chemical parameters of the aromatic substituents

Substituents	Physico-chemical parameters ^a			
	MR*	σ		π
		<i>meta</i>	<i>para</i>	
OCH ₃	0.787	0.12	-0.27	-0.02
N ₃	1.020	0.27	—	0.46
OC ₂ H ₅	1.247	0.10	—	0.38
OPr ⁱ	1.706	0.10	—	0.85
SCH ₃	1.382	0.15	—	0.61
N ₃	1.020	0.27	—	0.46
F	0.092	0.34	0.06	0.14
I	1.394	0.35	—	1.12
Cl	0.603	0.37	0.23	0.71
Br	0.888	0.39	—	0.86
CF ₃	0.502	0.43	—	0.88
CN	0.633	0.56	—	-0.57
NO ₂	0.736	0.71	—	-0.28
OH	0.285	—	-0.37	-0.67

a. Taken from Ref. 21.

* MR values are scaled by a factor of 0.1 as usual.

(explained variance), r or R (correlation coefficient), SEE (standard error of estimate) and F (variance ratio). The significance of the regression coefficients and constants (intercepts) were judged by 't' test. In case an intercept was statistically insignificant and omission of this did not cause harm to statistical quality of the equation, the same was done to obtain a more acceptable equation. Unless otherwise stated, a compound was considered as an outlier in case of a particular equation when the residual value was more than twice as the SEE of the equation. The use of the predictor variables in a multivariate equation was also justified by autocorrelation study.

The multiple regression analyses²² were done using a software program *RRR98* written in GW-BASIC and developed by one of the authors²³.

Results and Discussion

The equations relating the cytotoxicity of the compounds against human ovarian carcinoma cell lines are listed in **Table IV**. While relating the activity data of mono-*meta* substituted compounds ($n = 12$) with different physico-chemical parameters, electronic (MR_m) and steric parameters (σ_m) were found to be important, and both were found to follow parabolic relationships with the activity though linear (1st degree) term was absent in case of the electronic parameter (Eq. OC1). The parameters for the *meta* position were derived by adding the values for the

substituents at 3' and 5' positions of the phenyl ring. The intercept, being insignificant, was dropped without hampering statistical quality of the equation. When both mono- and di- *meta* substituted compounds ($n = 17$) were considered, a dummy parameter N_m denoting number of *meta* substitutions was found to be a significant term in the corresponding equation (no. OC2). With the inclusion of the *para* substituted compounds ($n = 22$) a new indicator variable I_p denoting presence or absence (value 1 or 0 respectively) of *para* substituents showed significant contribution to the response variable. When all compounds ($n = 29$) were considered, another indicator variable $I_{2,3}'$, denoting presence (value 1) or absence (value 0) of substituents on both positions 2' and 3', had to be considered (Eqn OC4). Though regression coefficient of $I_{2,3}'$ was significant, the compound **17**, which was having a value of unity for $I_{2,3}$, was excluded on the ground that rest of the compounds were devoid of such structural feature (Eq. OC5). The equation OC5 could explain 61.5% variation in the activity. Analysis of residuals showed that the residuals of the compounds **23** and **25** were more than 1.9 times as the SEE of the equation OC5. Further exclusion of these two compounds ($n = 26$) gave equation OC6 explaining 77.9% of the variance and the F value rose by 9 units. The compound **12**, an outlier in case of Eq. OC6, was excluded to get a final equation (OC7) explaining 84.4% of the variation of the activity and without having any outlier. The predictor variables used in the equation OC7 were not much autocorrelated [autocorrelation (r) matrix: MR_m vs $\sigma_m^2 / N_m / I_p = 0.25 / 0.35 / 0.46$; MR_m^2 vs $\sigma_m^2 / N_m / I_p = 0.24 / 0.32 / 0.40$; σ_m^2 vs $N_m / I_p = 0.61 / 0.54$; N_m vs $I_p = 0.12$]. The calculated biological activity values according to Eq. OC7 are shown in **Table II**.

When the term N_m in the equation OC5 was excluded, the first degree term of σ_m could find a place in the equation without much compromise in the quality as seen in case of equation OC8. Next, the compound (no. **12**) having residual more than 1.75 times the SEE of Eq. OC8 was excluded to obtain a statistically more acceptable equation (no. OC9) which was further improved by exclusion of the compound **14** (Eqn OC10).

From the equations of the **Table IV** it is evident that the substituents at *meta* position significantly contribute to the activity. The positive regression coefficient of N_m signifies that the *meta*-substitution is conducive to the activity. Moreover, a di-*meta*

Table IV—Relations of cytotoxicity of paclitaxel and its C₂ modified analogs against human ovarian carcinoma cell line with some physico-chemical and structural parameters

Eq. no.	Type(s) of Substitution(s)	n	Equations*	Statistics	
				EV% (r or R)	F [#] (SEE)
OC1	mono- <i>meta</i>	12	$pC_1 = 2.14 (\pm 1.29) MR_m - 1.67 (\pm 0.85) MR_m^2 - 2.05 (\pm 2.07)^b \sigma_m^2$	68.9 (0.86)	9.4 (0.39)
OC2	mono- <i>meta</i> + di- <i>meta</i>	17	$pC_1 = -0.47 (\pm 0.24) MR_m^2 - 0.79 (\pm 0.62) \sigma_m^2 + 0.53 (\pm 0.38) N_m$	55.9 (0.78)	8.8 (0.49)
OC3	mono- <i>meta</i> + di- <i>meta</i> + <i>para</i> + <i>para</i> & <i>meta</i>	22	$pC_1 = -0.46 (\pm 0.22) MR_m^2 - 0.82 (\pm 0.60) \sigma_m^2 + 0.54 (\pm 0.34) N_m - 1.23 (\pm 0.50) I_p$	59.2 (0.81)	12.3 (0.48)
OC4	All	29	$pC_1 = 1.04 (\pm 1.18)^b MR_m - 0.92 (\pm 0.59) MR_m^2 - 0.90 (\pm 0.53) \sigma_m^2 + 0.63 (\pm 0.37) N_m - 0.96 (\pm 0.49) I_p - 1.60 (\pm 0.97) I_2 - 0.56 (\pm 1.14)^d$	63.2 (0.84)	9.0 (0.44)
OC5	All excluding (<i>ex.</i>) Compd. 17	28	$pC_1 = 1.04 (\pm 1.18)^b MR_m - 0.92 (\pm 0.59) MR_m^2 - 0.90 (\pm 0.53) \sigma_m^2 + 0.63 (\pm 0.37) N_m - 0.96 (\pm 0.49) I_p - 0.56 (\pm 1.17)^d$	61.5 (0.83)	9.6 (0.44)
OC6	All <i>ex.</i> Compds. 17, 23 & 25	26	$pC_1 = 1.91 (\pm 1.08) MR_m - 1.48 (\pm 0.60) MR_m^2 - 0.95 (\pm 0.44) \sigma_m^2 + 0.65 (\pm 0.31) N_m - 0.92 (\pm 0.38) I_p - 0.76 (\pm 0.96)^c$	77.9 (0.91)	18.7 (0.34)
OC7	All <i>ex.</i> Compds. 12, 17, 23 & 25	25	$pC_1 = 2.09 (\pm 0.95) MR_m - 1.60 (\pm 0.53) MR_m^2 - 0.96 (\pm 0.38) \sigma_m^2 + 0.66 (\pm 0.27) N_m - 0.97 (\pm 0.33) I_p - 0.75 (\pm 0.86)^b$	84.4 (0.94)	27.0 (0.29)
OC8	All excluding Compd. 17	28	$pC_1 = 0.91 (\pm 1.22)^c MR_m - 0.72 (\pm 0.61) MR_m^2 + 2.17 (\pm 1.39) \sigma_m - 1.99 (\pm 1.11) \sigma_m^2 - 0.95 (\pm 0.50) I_p - 0.51 (\pm 1.27)^c$	58.9 (0.82)	8.75 (0.45)
OC9	All <i>ex.</i> Compds. 12 & 17	27	$pC_1 = 0.99 (\pm 1.15)^b MR_m - 0.79 (\pm 0.58) MR_m^2 + 2.37 (\pm 1.32) \sigma_m - 2.13 (\pm 1.05) \sigma_m^2 - 1.00 (\pm 0.47) I_p - 0.51 (\pm 1.21)^d$	65.2 (0.85)	10.7 (0.42)
OC10	All <i>ex.</i> Compds. 12, 14 & 17	26	$pC_1 = 1.18 (\pm 1.07) MR_m - 0.90 (\pm 0.54) MR_m^2 + 2.68 (\pm 1.25) \sigma_m - 2.30 (\pm 0.97) \sigma_m^2 - 1.04 (\pm 0.44) I_p - 0.57 (\pm 1.12)^d$	72.1 (0.88)	13.9 (0.39)

n = No. of data points.

* Regression coefficients and constants are significant at 95% level if not superscripted. 95% confidence intervals are shown within parentheses. Degree of freedom = n - np - 1; np = no. of predictor variables; I = 1 if intercept is present; I = 0 otherwise.

Degrees of freedom of F ratio = (np, n - np - 1); F values are significant at 99% level if not superscripted.

a, b, c, d and e correspond to significance levels of 95-99%, 90-95%, 80-90%, 60-80% and 50-60% respectively.

 $MR_m = MR_{R3'} + MR_{R5'}$; $\sigma_m = \sigma_{R3'} + \sigma_{R5'}$

substitution is preferred over mono-*meta* substitution provided the combined MR value does not exceed the optimum limit. The parabolic relation of the biological activity is found with both MR and σ values of the meta positions. The optimum MR value for the meta position (sum of values for positions 3' and 5') was calculated from Eq. OC7 and found to be

0.653. The optimum σ_m obtained from Eq. OC10 was 0.583. The negative regression coefficient of I_p signifies that the presence of *para* substituent lowers the activity. No term relating to the *ortho* position was found to significantly improve the quality of the equations.

Table V shows the equations derived from MRA

Table V—Relations of cytotoxicity of paclitaxel and its C₂ modified analogs against human colon carcinoma cell line with some physico-chemical and structural parameters

Eq. No.	Type(s) of Substitution(s)	n	Equations*	Statistics	
				EV% (r or R)	F [#] (SEE)
CC1	mono- <i>meta</i>	11	$pC_2 = -0.35 (\pm 0.18) MR_m^2$	35.0 (0.59)	18.9 (0.39)
CC2	mono- <i>meta</i> + di- <i>meta</i>	16	$pC_2 = -0.28 (\pm 0.11) MR_m^2$	29.9 (0.54)	27.1 (0.40)
CC3	mono- <i>meta</i> + di- <i>meta</i> + <i>para</i> + <i>para</i> & <i>meta</i>	21	$pC_2 = -0.35 (\pm 0.10) MR_m^2 - 1.13 (\pm 0.32) I_p$ $+ 1.58 (\pm 0.66) \pi_p$	75.6 (0.88)	47.2 (0.34)
CC4	All	29	$pC_2 = -0.36 (\pm 0.12) MR_m^2 - 1.08 (\pm 0.37) I_p$ $+ 1.61 (\pm 0.84) \pi_p - 0.96 (\pm 0.63) I_{2'-3'}$	61.8 (0.81)	24.8 (0.44)
CC5	All <i>ex.</i> Compds. 17 & 20	27	$pC_2 = -0.35 (\pm 0.10) MR_m^2 - 1.08 (\pm 0.31) I_p$ $+ 1.59 (\pm 0.71) \pi_p$	69.3 (0.85)	40.0 (0.37)
CC6	All <i>ex.</i> Compds 12, 17, 20 & 21	25	$pC_2 = -0.35 (\pm 0.09) MR_m^2 - 1.08 (\pm 0.27) I_p$ $+ 1.60 (\pm 0.61) \pi_p$	77.2 (0.89)	55.9 (0.31)

*See foot notes of **Table IV**

of cytotoxicity data against human colon carcinoma. In case of the mono-*meta* substituted compounds ($n = 11$, Eq. CC1), only molar refractivity term of the *meta* position was found to significantly contribute to the activity through a parabolic relationship. Both the linear term of MR_m and the intercept were absent. On inclusion of di-*meta* substituted compounds ($n = 16$, Eq. CC2) similar pattern was retained with some fall in quality (EV% and r) of the equation though F value rose by 8 units. When the *para* substituted compounds were included ($n = 21$, Eq. CC3), two terms relating to the *para* position appeared in the equations. The new terms were lipophilicity (π_p) and an indicator variable I_p denoting presence or absence (value 1 or 0) of the *para* substituent. When all equations were considered, another term $I_{2'-3'}$, defined previously, appeared. Only two compounds (compounds **17** and **20**) were having substitutions on both 2' and 3' positions and considering skewness of distribution of the structural feature, these were excluded to get Eq. CC5 explaining 69.3% variance. For further improvement of the quality, the compounds **12** and **21** having residuals more than twice as the SEE of Eq. CC5 were excluded to obtain Eq. CC6 having 77.9% EV and a F ratio of 55.9 (degrees of freedom = 3, 22). The predictor variables

were mutually orthogonal as evidenced by autocorrelation study [autocorrelation (r) matrix of Eq. CC6 : MR_m^2 vs I_p / $\pi_p = 0.23$ / 0.21 ; I_p vs $\pi_p = 0.10$]. The calculated biological activity values according to Eq. CC6 are shown in **Table II**.

On examination of the equations of the **Table V** it appears the higher MR values of the *meta* substituents are not favourable for the activity against colon carcinoma cell line because the regression coefficients of the second order terms of MR_m bear negative values in all cases. The negative coefficient of I_p signifies that the *para* substitution should be avoided for having good activity. Again, if any *para* substitution is present, it should have higher lipophilicity to favour the activity, as evinced by the positive coefficient of π_p . No physico-chemical or indicator term relating *ortho* position could find a place in the equations.

The equations relating cytotoxicity of the compounds against human burkitt lymphoma cell lines are shown in **Table VI**. MRA on cytotoxicity data of mono-*meta* substituted compounds revealed important contribution of molar refractivity and electronic (Hammett σ) terms ($n = 9$, Eq. BL1). Both the parameters showed parabolic relationships with the response variable though the first degree terms

Table VI—Relations of cytotoxicity of paclitaxel and its C₂ modified analogs against human burkitt lymphoma cell line with some physico-chemical parameters

Eq. no.	Type(s) of Substitution(s)	n	Equations*	Statistics	
				EV% (r or R)	F ^a (SEE)
BL1	mono- <i>meta</i>	09	$pC_3 = -0.27 (\pm 0.27)^b MR_m^2 - 1.18 (\pm 1.68)^c \sigma_m^2 - 0.81 (\pm 0.57)$	30.3 (0.69)	2.7 ^c (0.29)
BL2	mono- <i>meta</i> + di- <i>meta</i>	13	$pC_3 = -0.12 (\pm 0.15)^c MR_m^2 - 0.23 (\pm 0.31)^c \sigma_m^2 - 1.10 (\pm 0.33)$	25.7 (0.62)	3.1 ^b (0.27)
BL3	mono- <i>meta</i> + di- <i>meta</i> + <i>para</i> + <i>para</i> & <i>meta</i>	17	$pC_3 = -0.10 (\pm 0.12)^b MR_m^2 - 0.26 (\pm 0.29)^b \sigma_m^2 - 6.25 (\pm 4.44) \sigma_p^2 - 1.14 (\pm 0.25)$	34.9 (0.69)	3.9 ^a (0.27)
BL4	All	22	$pC_3 = -0.07 (\pm 0.10)^c MR_m^2 - 0.23 (\pm 0.26)^b \sigma_m^2 - 5.38 (\pm 3.75) \sigma_p^2 - 1.24 (\pm 0.17)$	30.6 (0.64)	4.1 ^a (0.24)
BL5	All <i>ex. Compds.</i> 5, 7 & 9	19	$pC_3 = -0.05 (\pm 0.06)^b MR_m^2 - 0.19 (\pm 0.15) \sigma_m^2 - 4.77 (\pm 2.18) \sigma_p^2 - 1.31 (\pm 0.10)$	56.9 (0.80)	8.9 (0.14)
BL6	All <i>ex. Compds.</i> 5, 7, 9 & 24	18	$pC_3 = -0.04 (\pm 0.04)^b MR_m^2 - 0.18 (\pm 0.11) \sigma_m^2 - 4.47 (\pm 1.59) \sigma_p^2 - 1.35 (\pm 0.08)$	70.2 (0.87)	14.4 (0.10)

vide foot notes of **Table IV**

were absent. On addition of the di-*meta* substituted compounds ($n = 13$, Eq. BL2), similar pattern of relations was obtained. The quality of the equations in both the cases was not good ($EV < 31\%$, regression coefficient and F values were significant below 95% level). When *para* - substituted compounds were included ($n = 17$, Eq. BL3), second degree term of σ_p appeared in the equation. The quality of the equation in this case slightly improved ($EV = 34.9\%$, F value significant at 95% level). When all compounds were considered ($n = 22$, Eq. BL4), an equation explaining 30.6% of variance was obtained. No other term improving the quality of the equation was available. Analysis of residuals showed that the compounds **5**, **7** and **9** had residuals more than 1.5 times as the SEE of the Eq. BL4. To obtain a statistically acceptable Eq., those three compounds were excluded leading to Eq. BL5 with 56.9% variance and one outlier (compound **24**). Omission of the outlier gave the final Eq. (no. BL6) explaining 70.2% variance and having F value significant at 99% level and regression coefficients significant at 95% level (except for MR_m which was significant at 90% level). The independent variables were not much autocorrelated [autocorrelation (r) matrix of Eq. BL6 : MR_m^2 vs σ_m^2 / $\sigma_p^2 = 0.20/0.29$;

σ_m^2 vs $\sigma_p^2 = 0.17$]. The calculated biological activity values according to Eq. BL6 are listed in **Table II**.

After careful examination of the equations of **Table VI** it appears that the *meta* substituents significantly contribute to the variation of the activity. Negative coefficient of the square term of MR_m imply that higher value of MR of the *meta* substituents is not preferred for better activity. The significance of the negative coefficients of the square terms of electronic parameters (σ_m and σ_p) is that the electronic influence (withdrawal or donation) of the substituents through inductive (mainly for *meta*) and resonance (for *para*) effects should be minimum. The *ortho* - substituents did not significantly contribute to the activity.

In summary, molar refractivity term of the *meta* substituents was found to be important in all three biological activity datasets. In all the cases MR_m showed a parabolic relationship with the activities. It appears that the volume of the *meta* substituents has an influence on the preferred orientation of the C₂ substituent of the taxane ring which may take part in the interaction with other part of the molecule (like C₁₃ ester side chain⁶) or the receptor site. No term relating to *ortho* substitution showed significance in

any dataset. In ovarian carcinoma and colon carcinoma datasets, absence of *para* substitution was preferred. In case of both ovarian carcinoma and burkitt lymphoma, electronic parameter of the *meta* substituents showed parabolic relationships. The activity against colon carcinoma is further positively contributed by lipophilicity of the *para* substituent and that against burkitt lymphoma is negatively by square term of electronic parameter of *para* substitution. The differences among the equations relating activity against three cell lines with physico-chemical parameters imply that there may be some differences among the receptor structures.

Substitutional requirements for optimum receptor fit are reflected in the equations discussed above. Judicious selection of the substituents based on the rigorous examination of the equations may lead to more active compounds of therapeutic interest.

Acknowledgement

The authors thank Prof K D Chowdhury, Principal, SIPS, Orissa for providing necessary computational facility to one of them (KR).

References

- 1 Suffness M in *Annual Reports in Medicinal Chemistry*, Vol 28, edited by J A Bristol (Academic Press, New York) **1993**, 305.
- 2 Buss A D & Waigh R D in *Burger's Medicinal Chemistry and Drug Discovery*, Vol 1, edited by M E Wolff (John Wiley & Sons, New York) **1995**, 996.
- 3 Wani M C, Taylor H L, Wall M E, Coggon P & McPhail A T, *J Am Chem Soc*, **93**, **1971**, 2325.
- 4 Kapoor V K & Mahindroo N, *Indian J Chem*, **36B**, **1997**, 639.
- 5 Das B & Rao S P, *Indian J Chem*, **35B**, **1996**, 883.
- 6 Kingston D G I, Chaudhury A G, Chordia M D, Gharpure M, Gunatilaka A A L, Higgs P I, Rimoldi J M, Samala L, Jagtap P G, Giannakakou P, Jiang Y Q, Lin C M, Hamel E, Long B H, Fairchild C R & Johnston K A, *J Med Chem*, **41**, **1998**, 3715.
- 7 Wall M E & Wani M C, In *Alkaloids: Chemical & Biological Perspectives*, Vol 9, edited by S W Pelletier (Pergamon, Oxford) **1995**, 1.
- 8 Chen S H & Farina V in *The Chemistry and Pharmacology of Taxol and its Derivatives*, Vol 22, edited by V Farina (Elsevier, Amsterdam), **1995**, 165.
- 9 Bouchard H, Pulicani J P, Vuilhorgne M, Bourzat J D & Commercon A, *Tetrahedron Lett*, **35**, **1994**, 9713.
- 10 Marder R, Dubois J, Gu'enard D, Vogelein F G & Potier P, *Tetrahedron*, **51**, **1995**, 1985.
- 11 Chen S H, Wei J M & Farina V, *Tetrahedron Lett*, **34**, **1993**, 3205.
- 12 Neidigh K A, Gharpure M M, Rimoldi J M & Kingston D G I, *Tetrahedron Lett*, **35**, **1994**, 6839.
- 13 Chen S H, Huang S & Farina V, *Tetrahedron Lett*, **35**, **1994**, 41.
- 14 Chaudhury A G & Kingston D G I, *Tetrahedron Lett*, **34**, **1993**, 4921.
- 15 Liang X, Kingston D G I, Lin C M & Hamel E, *Tetrahedron Lett*, **36**, **1995**, 2901.
- 16 Klein L L, Yeung C M, Li L & Plattner J J, *Tetrahedron Lett*, **35**, **1994**, 4707.
- 17 Mangatal L, Adeline M T, Guenard D, Gueritta-Voegelein F & Potier P, *Tetrahedron*, **45**, **1989**, 4177.
- 18 Chen S, Farina V, Wei J, Long B, Fairchild C, Mamber S W, Kadow J F, Vyas D & Doyle T W, *Bioorg Med Chem Lett*, **4**, **1994**, 479.
- 19 Grover S, Rimoldi J M, Molinero A A, Chaudhary A A, Kingston D G I & Hamel E, *Biochemistry*, **34**, **1995**, 3927.
- 20 Hansch C, in *Drug Design*, Vol 1, edited by E J Ariens (Academic Press, New York) **1971**, 271.
- 21 Kubinyi H, in *Burger's Medicinal Chemistry and Drug Discovery*, Vol 1, edited by M E Wolff (John Wiley & Sons, New York), **1995**, 507.
- 22 Snedecor G W & Cochran W G, *Statistical Methods* (Oxford & IBH Publishing Co Pvt Ltd, New Delhi) **1967**, 381.
- 23 The GW-BASIC program *RRR98* was developed by Kunal Roy (1998) and standardized on known datasets.