A Strategy for Primary High Throughput Cytotoxicity Screening in Pharmaceutical Toxicology

Peter J Bugelski,[1] Uzma Atif,[1,4] Sarah Molton,[3] Ilana Toeg,[2] Peter G Lord,[1] and D. Gwyn Morgan[1]

Received June 13, 2000; accepted July 7, 2000

Purpose. Recent advances in combinatorial chemistry and high throughput screens for pharmacologic activity have created an increasing demand for *in vitro* high throughput screens for toxicological evaluation in the early phases of drug discovery.

Methods. To develop a strategy for such a screen, we have conducted a data mining study of the National Cancer Institute's Developmental Therapeutics Program (DTP) cytotoxicity database.

Results. Using hierarchical cluster analysis, we confirmed that the different tissues of origin and individual cell lines showed differential sensitivity to compounds in the DTP Standard Agents database. Surprisingly, however, approaching the data globally, linear regression analysis showed that the differences were relatively minor. Comparison with the literature on acute toxicity in mice showed that the predictive power of growth inhibition was marginally superior to that of cell death.

Conclusions. This datamining study suggests that in designing a strategy for high throughput cytotoxicity screening: a single cell line, the choice of which may not be critical, can be used as a primary screen; a single end point may be an adequate measure and a cut off value for 50% growth inhibition between 10^{-6} and 10^{-8} M may be a reasonable starting point for accepting a cytotoxic compound for scale up and further study.

KEY WORDS: cultured cells; data mining; correlation; National Cancer Institute Developmental Therapeutics Program.

INTRODUCTION

Traditionally, selection of lead compounds for development has been based on screening for desired pharmacologic activity in animal models and tissue preparations (1). In the recent era of rational drug design, specific molecular targets were defined and primary screening conducted in *in vitro* systems (2), followed by more extensive secondary testing in animal models for evidence of efficacy (3). Safety concerns typically received little attention. Presently, advances in combinatorial chemistry and high throughput, molecular mechanism based screens for pharmacologic activity are producing a plethora of potential drug candidates (2). Moreover, while the numbers of potential leads is high, typically they are synthesized in very small amounts; typically ≤ 100 mg (4).

The pharmaceutical industry is thus faced with the challenge of how to prioritize scale up of potential leads from a few milligrams available from combinatorial chemistry to the grams required for secondary testing and safety assessment. In addition to current strategies for prioritisation such as biotransformation (5) and SAR for mutagenicity (6), we believe there is the potential to conduct a simple *in vitro* screen for cytotoxicity at an early stage of discovery.

It is well established that *in vivo,* there are species differences in the susceptibility to toxicity and marked differences in the susceptibility of different tissues to toxicity. These differences receive great attention by many scientists who seek to design screens for toxicity as a replacement for toxicity testing in animals. Great emphasis is often placed on the species and tissue of origin of the test system and on the absolute value for potency against a particular end point (7–9). Despite the efforts of numerous laboratories over the past two decades to define *in vitro* systems which can replace testing in animals, a validated, predictive, species and tissue specific *in vitro* screen, which requires small amounts of compound remains largely unrealized (10–11).

A practical way of addressing this challenge is to create a simple screen which would have some predictive value for those potential leads which will ultimately cause unacceptable side effects during either preclinical or clinical development. In this paper, we propose a strategy for high throughput *in vitro* screening. To support these arguments, we have conducted a data mining study drawing on the National Cancer Institute (NCI) Developmental Therapeutics (DTP) database for cytotoxicity data (12). The NCI database contains cytotoxicity and cytostasis data for 60 well-characterized cell lines and data for over 30,000 compounds are currently available on-line (http://dtp.nci.nih.gov/). Our study has been conducted on the "Standard Agents" subset of the DTP database. The Standard Agents data are on 170 compounds selected to include anticancer drugs used in clinical practice and novel anticancer compounds under development. Serendipitously, the Standard Agents database also contains data on other therapeutic classes of compounds which are cytotoxic, so the relevance of the data may be of importance for drugs in general. To complete our analysis, we have supplemented the NCI data with data from the literature on murine lethality for 57 of the Standard Agents.

MATERIALS AND METHODS

Cytotoxicity Data. The technical details of the NCI DTP screening process and the origins of the cell lines are described (12) in the web-site at, http://dtp.nci.gov/docs/cancer/cancer_screen.html. The Standard Agents database contains cytotoxicity data for 170 compounds (listed in: http://dtp.nci.nih.gov/docs/cancer/searches/standard_agent_table.html).

Mean log growth inhibitory (GI) and log cytolytic (LC) concentration data for the 170 Standard Agents for the individual cell lines from each tissue type, [melanoma, non-small cell lung carcinoma (NSCL), small cell lung (SCL), renal,

[1] Safety Assessment, SmithKline Beecham Pharmaceuticals, The Frythe, Welwyn, Hertfordshire AL6 9AR, United Kingdom.
[2] Information Mangement, SmithKline Beecham Pharmaceuticals, The Frythe, Welwyn, Hertfordshire AL6 9AR, United Kingdom.
[3] Present address: Babraham Institute, Cambridge, United Kingdom.
[4] To whom correspondence should be addressed. (e-mail: Uzma_Atif-1@sbphrd.com)

ABBREVIATIONS: GI 50, 50% growth inhibitory concentration; LC 50, 50% lethal concentration; LD10, 10% murine lethal concentration; NCI, National Cancer Institutes; DTP, development therapeutics database; SAR, structure activity relationship; HCA, hierarchical cluster analysis; NSCLC, non-small cell lung cancer.

0724-8741/00/1000-1265$18.00/0

breast, colon, prostate, ovarian, central nervous system (CNS), leukemia/lymphoma and murine] as well as individual cell line data for A549 (NSCL), M14 (melanoma), MOLT4 and CCRF-CEM (human lymphoma) and P388 (murine lymphoma) were downloaded into Excel 97 SR1 spreadsheet format (Microsoft, Inc. Redmond WA). For a few of the compounds, there are no data available for all the tissues or cell lines but data for a minimum of 162 compounds were available for this analysis.

Mouse LD10

Literature searches of the Medline database (1966–July 1999) and Chemical Abstracts database (1967–July 1999) were conducted to obtain references containing murine LD10 data, using appropriate keywords. Additional key articles were obtained from the cited references provided in these articles. The final set of 9 papers provided LD10 data for 57 compounds (Table 1). LD10s quoted in these papers varied depending on the strain, route of administration and dosing schedule used, but as far as possible, a figure corrected for single daily parenteral dose in mg/kg or mg/ m^2 (the latter being converted to mg/kg for this analysis) was obtained. When multiple values were found, the lowest values were brought forward. Data for all compounds was converted to moles/kg using molecular weights obtained from the Standard Agents database and finally, to allow direct comparison to the log LC50 and log GI50 data from the Standard Agents database, the LD10 data were log transformed. The compounds and the LD10 values included in our study are listed in Table 1.

Data Analysis

Hierarchical cluster analysis was performed using Statistica v5.1(StatSoft Inc, Tulsa OK). Clustering was determined using complete linkages and Euclidean distance (d). The results were plotted as normalized vertical clustering trees [d (link)/d (max)] ×100.

Scatter plots were made with Excel 97 SR1 (Microsoft) or TableCurve v1.0 (Jandel Scientific, San Rafael CA) and linear regression analysis was done with TableCurve. Pearson correlation analysis was done with SigmaStat v1.0 (Jandel) and t tests were done manually from values for standard error of the mean (sem) obtained from TableCurve. P values >0.05 were accepted as significant.

RESULTS

Differences among the Tissues of Origin and Among Individual Cell Lines

From the *in vitro* cytotoxicity data in the DTP Standard Agents database we have analyzed the 50% lethal concentration (LC50) and the 50% growth inhibitory concentration (GI50) only. The data cover 11 different tissues of origin and up to 60 individual cell lines.

To determine if there are differential responses among

Table 1. Compounds for Which Murine LD10 Data Were Analyzed

Compound	Log mouse LD10 (moles/kg)	Reference	Compound	Log mouse LD10 (moles/kg)	Reference
3-deazauridine	–3.1	21,22	hexamethylene melamine	–3.5	25
5-azacytidine	–4.3	21,23	hydroxyurea	–2.8	25
5-azadeoxycytidine	–3.9	24	ICRF-159	–2.7	21,23
5-FU	–3.3	22,25	amsacrine	–5.1	21,22,26
5-FUDR	–3.4	22,25	maytansine	–6.2	21,22
6 mercaptopurine	–3.7	24,25	melphelan	–4.9	26,28
actinomycin D	–6.8	21,22,25,26	methotrexate	–5.4	22,25
anguidine	–4.6	22	methyl-CCNU	–3.7	21,23
aphidicolin glycinate	–3.6	24	methyl-GAG	–3.8	22,28
acivicin	–5.0	24	mitomycin C	–5.2	22,26,27
BCNU	–4.3	22,26	mitramycin	–6.0	23
bleomycin	–4.9	21,23,24,26	PALA	–2.5	21,22
busulphan	–3.7	25	piperazinedione	–4.6	21
carboplatin	–3.4	24	porfiromycin	–3.7	21,25
CCNU	–3.6	21,23	rhizoxin	–5.2	24
chlorambucil	–4.5	25	R-methyl formamide	–2.4	25
chlorozotocin	–4.1	21,22	soluble Baker's antifol	–4.0	21
cisplatin	–4.9	21,24,26	teroxirone	–3.3	24
cyclophosphamide	–4.1	22,24,25	thalacarpine	–3.5	21,22
cytocine arabinoside	–2.7	24	thioguanine	–5.1	25
daunorubicin	–5.5	23	thio-tepa	–4.7	22,25
doxorubicin	–5.3	22,26	topotecan	–4.3	29
DTIC	–3.7	23	triethylenemelamine	–5.3	25
brequinar	–3.5	24	uracil nitrogen mustard	–5.1	25
flavoneacetic acid ester	–3.1	24	vinblastin sulphate	–6.7	26,28
fludarabine phosphate	–2.7	24	vincristine sulfate	–7.7	22,28
ftorafur	–2.9	27	etoposide	–5.2	24,26
galluim nitrate	–3.7	21,22	Yoshi-864	–3.9	23
guanazole	–2.5	23			

the different tissues of origin, the data for average cytotoxicity as measured by LC50 for each tissue of origin were analyzed by hierarchical cluster analysis (HCA). HCA (Figure 1a) showed that the tissues fell into three clusters; melanoma grouped with small cell lung carcinoma, murine lymphoid cells grouped with the human leukemia/lymphoma cells and the remaining tissues formed the third group.

To determine if these groupings were an artifact of analyzing average data for the different tissues of origin, LC50 data for individual cell lines selected from these three broad clusters were analyzed in a similar manner. The cell lines: P-388 - murine lymphoma; MOLT 4 and CCRF-CEM human lymphoma/leukemia; M14–melanoma and A549–non-small cell lung carcinoma (NSCLC) were chosen on the basis of their falling into the three "tissue of origin" clusters found by HCA for the tissue data. Two leukemia/lymphoma lines MOLT4 and CCRF-CEM were chosen to allow comparison within a cluster and P388 was selected because it is of murine, rather than human origin. A549 and M14 cells were chosen because they are the most widely studied cells of the non-small cell lung carcinoma and melanoma panels, respectively. HCA by cell lines revealed four clusters, shown in Figure 1b. P388 grouped with A549 non-small cell lung carcinoma cells rather than the human leukemia/lymphoma cells and M14 was a wide-outlier.

For the GI50 data, three clusters were found for the tissues and cells (Figure 1c and d). However, the groupings were different from those found using the LC50 data, indicating that the tissues and cells responded differentially in terms of growth inhibition and cytolysis.

The results of hierarchical cluster analysis, at face value, support the contention that data from a single tissue of origin or single cell line may be inadequate to describe the *in vitro* cytotoxicity for the Standard Agents. However, multivariate techniques are, by design, intended to emphasize the differences among the data. To address the relevance of the differences, the LC50 data for the mean values for 10 of the tissue types were plotted as a function of human NSCLC (Figure 2a). Although there was some scatter in the data, there was an obvious linear trend. These data were subjected to linear regression analysis on a tissue by tissue basis (Table 2). In each case, the data for one tissue type was found to be linearly related to all the other tissue types and there were only small, albeit statistically significant differences in the slopes and intercepts when compared to the expectation values (1 and 0, respectively) by one tailed t test.

To determine if the lack of major differences among the tissue types was an artifact of analyzing mean data, LC50 data for the 4 other cell lines were plotted as a function of A549, Figure 2b. Although there was some scatter, M14 showing the

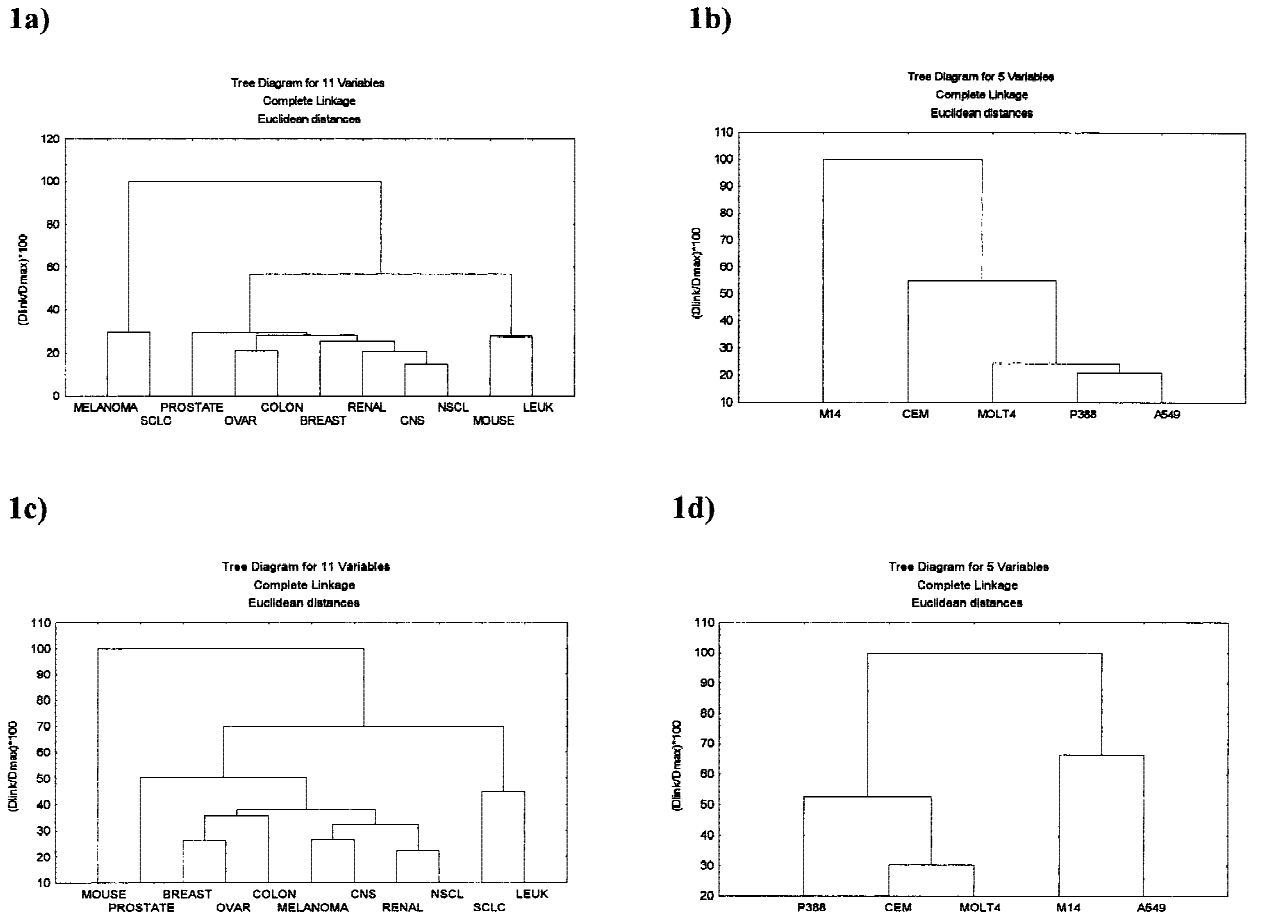

Fig. 1. a) Hierarcial cluster analysis (HCA) for mean cytotoxicity data as measured by LC_{50} for each tissue of origin. b) HCA to determine if the groupings in Fig 1a were an artifact of analysis of mean data. LC_{50} data for individual cell lines selected from the clusters shown in Figure 1a. c) HCA for mean cytotoxicity data as measured by GI_{50} for each tissue of origin. d) HCA to determine if the groupings in Fig 1a were an artifact of analysis of mean data. GI_{50} data for individual cell lines selected from the clusters shown in Figure 1c.

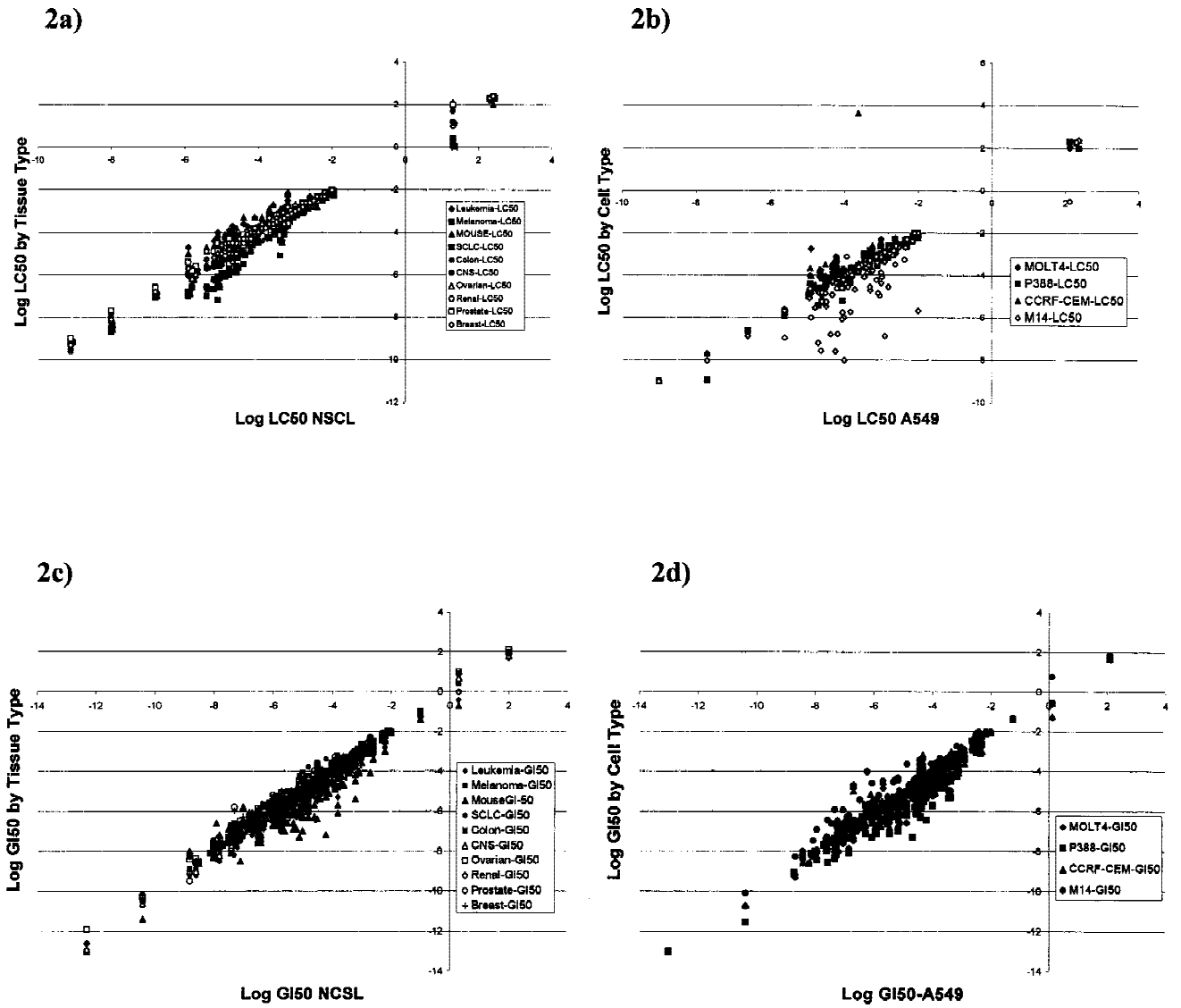

Fig. 2. a) Mean LC_{50} data for the 10 tissue types plotted as a function of non-small cell lung carcinoma. b) LC_{50} data for MOLT4, P388, CCRF-CEM and M14 plotted as a function of LC_{50} for A549. c) Mean GI_{50} data for the 10 tissue types plotted as a function of non-small cell lung carcinoma. b) GI_{50} data for MOLT4, P388, CCRF-CEM and M14 plotted as a function of LC_{50} for A549.

greatest, there was a strong linear relationship between A549 and the other cell lines. Linear regression analysis for all the individual cell lines confirmed the linear relationship (Table 3). In all cases, a statistically significant linear relationship was found between the data for each of the different cell lines and again, there were small but significant differences in slope or intercept.

Differences Between GI50 and LC50

To determine if the choice of end point affected the results of the linear regression analysis, the analyses shown in Figures 2a and 2b and Tables 2 and 3 were repeated for the GI50 data (Figures 2c and 2d and Table 3). The analyses with GI50 data gave results similar to LC50, with linear relationships being found as well as small but statistically significant differences between the measured values and expectation values for slope and intercept. The GI50 values tended to be lower than LC50 values.

Predictive Power

To determine the predictive power of GI50 and LC50 data for *in vivo* data, correlation analysis between cytotoxicity and murine LD10 was performed for data from the 5 cell lines. As shown in Table 4, there was a statistically significant correlation between LC50 and LD10 and also between GI50 and LD10 (Pearson correlation) for each of the tissues and cell lines. Regression analysis revealed that the data could be modeled by a linear function, but the r^2 values were not very high. However, as reflected by the larger Pearson correlation coefficients and r^2 values, GI50 was a marginally better predictor of LD10 than was LC50. Using the linear fit to estimate the rates of false positives (i.e., *in vitro* toxic potency overestimating the *in vivo* toxic potency) and false negatives, between 5 and 18% of compounds were divergent greater than ± one standard error of the mean from the predicted values. However, only 0–4% of compounds were divergent greater than ± 2 times the sem, indicating a reasonable predictive power.

Table 2. Linear Regression Analysis for Mean LC_{50} for Different Tissues as a Function of Mean LC_{50} for Non-Small Cell Lung Carcinoma, Melanoma, Lymphoid and Murine Cells

	NSCL			Melanoma			Human leukemia/lymphoma			Murine lymphoma		
Tissue type	r^2	Slope ± sem	Intercept ± sem	r^2	Slope ± sem	Intercept ± sem	r^2	Slope ± sem	Intercept ± sem	r^2	Slope ± sem	Intercept ± sem
Leukemia	0.95	0.92 ± 0.02*	−0.09 ± 0.06	0.85	0.77 ± 0.03*	−0.49 ± 0.10*	—	—	—	0.95	1.00 ± 0.02	−0.07 ± 0.04
CNS	0.99	1.01 ± 0.01	0.03 ± 0.02	0.95	0.88 ± 0.02*	−0.29 ± 0.06*	0.95	1.05 ± 0.02*	−0.04 ± 0.07	0.94	1.02 ± 0.02	−0.08 ± 0.07
Melanoma	0.94	1.10 ± 0.02	0.14 ± 0.08	—	—	—	0.86	1.10 ± 0.04*	−0.04 ± 0.13	0.84	1.07 ± 0.04	−0.09 ± 0.13
Ovarian	0.99	0.99 ± 0.01	0.01 ± 0.03	0.90	0.84 ± 0.02*	−0.37 ± 0.08*	0.96	1.03 ± 0.02	−0.02 ± 0.06	0.97	1.01 ± 0.02	−0.06 ± 0.06
Renal	0.99	1.02 ± 0.01*	0.02 ± 0.03	0.95	0.88 ± 0.02*	−0.30 ± 0.07*	0.94	1.05 ± 0.02*	−0.05 ± 0.07	0.93	1.02 ± 0.02	−0.10 ± 0.08
Breast	0.99	1.00 ± 0.01	0.02 ± 0.00	0.93	0.86 ± 0.02*	−0.31 ± 0.07*	0.95	1.04 ± 0.02*	−0.03 ± 0.06	0.95	1.01 ± 0.02	−0.07 ± 0.07
Colon	0.99	1.01 ± 0.01	0.02 ± 0.03	0.92	0.87 ± 0.02*	−0.33 ± 0.08*	0.95	1.05 ± 0.02*	−0.02 ± 0.06	0.94	1.02 ± 0.02	−0.06 ± 0.07
Prostate	0.99	0.99 ± 0.01	−0.01 ± 0.03	0.92	0.85 ± 0.02*	−0.33 ± 0.08*	0.95	1.02 ± 0.02	−0.06 ± 0.07	0.94	0.99 ± 0.02	−0.11 ± 0.07
SCL	0.96	1.06 ± 0.02	0.08 ± 0.06	0.98	0.95 ± 0.01*	−0.13 ± 0.04*	0.88	1.07 ± 0.03*	−0.07 ± 0.11	0.87	1.04 ± 0.03	−0.11 ± 0.11
NSCL	—	—	—	0.94	0.92 ± 0.02*	−0.03 ± 0.00*	0.95	1.03 ± 0.02	−0.06 ± 0.06	0.95	1.00 ± 0.02	−0.11 ± 0.07
Murine	0.95	0.94 ± 0.02	−0.06 ± 0.06	0.84	0.78 ± 0.03*	−0.48 ± 0.11*	0.98	1.01 ± 0.01	0.04 ± 0.04	—	—	—

* Statistically different from expectation value for $y = mx + b$ where $m = 1$ and $b = 0$ ($P < 0.05$, One tailed t test).

As seen in Figure 3, GI50 has an approximately 5 log wider dynamic range than LC50 for A549 cells. There was no remarkable threshold, but interestingly only 2 of 35 clinically useful antineoplastics had log LC50 values below –5 M and only 5 of 34 had log GI50 values below –8 M.

DISCUSSION AND CONCLUSIONS

We have conducted a data mining study of the DTP Standard Agents database in an attempt to address some of the issues in designing a primary screen for cytotoxicity such as the choice of tissue of origin, specific cell line and measure of cytotoxicity. We compared data for GI50 and LC50 for 170 compounds, across cell lines derived from 11 tissue types and have studied the data for 5 cell lines in detail. Although the DTP screen is primarily directed toward aiding discovery of antineoplastics, the database contains data for compounds such as rapamycin and colchicine, which are not used as antineoplastics. The compounds also cover a wide range of the recognized cytotoxic agents, including:, alkylating agents (melphelan); anti-DNA or anti-DNA/RNA metabolites (thiopurine or methotrexate); inhibitors of topoisomerase I (topotecan) and topoisomerase II (amsacrine); anti-mitotics (colchicine) and inducers of apoptosis (cytarabine). Thus, although the results of our analysis are most applicable to drugs intended to be cytotoxic, they may have some relevance to pharmaceuticals in general.

Choice of Tissue of Origin and Cell Line

As shown by hierarchical cluster analysis (Figure 1), there are differences among the responses of the different tissues and cell lines to the Standard Agents, in agreement with the literature. The expression of the multidrug resistance gene MRP in the 60 DTP cell lines was studied and it was found that the CNS and lung carcinoma lines express a 14 fold variation in MRP expression (13). Metallothionine (MT) expression was studied in 53 of the cell lines and a 400 fold range in basal MT expression was determined (14).The latter authors also showed that basal MT expression correlated with sensitivity of the different cell lines to metal containing compounds.

The conclusion that the different tissues of origin and individual cell lines show differential responses to cytotoxic compounds seems acceptable but one is left with the question of whether these differences are important when designing a primary screen. As shown in Tables 2 and 3, linear regression analysis whether conducted on a tissue by tissue or individual cell line basis always revealed a strong, statistically significant linear relationship between tissues and between cell lines. These analyses indicate that the differences revealed by HCA and reported previously are subtle and may not be important when designing a primary screen.

Choice of End Point and Predictive Power

To determine if the choice of cytotoxic end point influenced the conclusions of the screen, differences between the LC50 and GI50 data were examined. Although GI50 correlated with LC50, the correlation coefficients were relatively low and the linear fits were relatively poor (Table 3). Use of the analytical tool COMPARE (15) indicated that there are greater differences among the cell lines for GI50 data than for

Table 3. Linear Regression Analysis for LC_{50} for 5 Different Cell Lines as a Function of LC50 for A549, M14, CCRF-CEM and P388 Cells

	A549 (ordinate)			M14 (ordinate)			CCRF-CEM (ordinate)			P388 (ordinate)		
Cell type (abscissa)	r^2	Slope ± sem	Intercept ± sem	r^2	Slope ± sem	Intercept ± sem	r^2	Slope ± sem	Intercept ± sem	r^2	Slope ± sem	Intercept ± sem
A549	—	—	—	0.72	0.69 ± 0.03*	−0.74 ± 0.14*	0.80	0.84 ± 0.03*	−0.66 ± 0.11*	0.96	0.95 ± 0.15*	−0.14 ± 0.05*
M14	0.72	1.04 ± 0.05	−0.30 ± 0.18	—	—	—	0.56	0.86 ± 0.06*	−1.01 ± 0.21*	0.71	1.00 ± 0.05	−0.40 ± 0.19*
MOLT4	0.96	0.96 ± 0.02*	−0.08 ± 0.06	0.67	0.66 ± 0.04*	−0.83 ± 0.15*	0.81	0.83 ± 0.03*	−0.63 ± 0.11*	0.95	0.93 ± 0.02*	−0.16 ± 0.06*
CCRF-CEM	0.96	0.96 ± 0.01*	−0.05 ± 0.05	0.56	0.65 ± 0.05*	−0.76 ± 0.19*	—	—	—	0.97	0.94 ± 0.04*	−0.11 ± 0.13
P388	0.96	1.01 ± 0.02	0.01 ± 0.06	0.71	0.71 ± 0.04*	−0.71 ± 0.15*	0.79	0.86 ± 0.03*	−0.62 ± 0.12*	—	—	—

* Statistically different from expectation value for $y = mx + b$ where $m = 1$ and $b = 0$ ($P < 0.05$, One tailed t test).

Table 4. Pearson Correlation and Linear Regression for Log LD10 (moles/kg) as a Function of Cell Type

	Log LD10 (abscissa)							
Cell type (LC50 (ordinate)	# of cpnds	Pearson correlation coefficient	P >	Linear regression r^2	Slope	Intercept	% Compounds > + 1 sem from predicted value	% Compound < – 1 sem from predicted value
A549	57	0.54	0.001	0.29	0.54 ± 0.11	–2.34 ± 0.42	14	12
M14	56	0.60	0.001	0.36	0.51 ± 0.09	–2.28 ± 0.38	11	14
MOLT4	57	0.53	0.001	0.28	0.54 ± 0.12	–2.35 ± 0.43	14	11
CCRF-CEM	57	0.34	0.011	0.27	0.67 ± 0.15	–0.50 ± 0.64	5	5
P388	54	0.53	0.001	0.28	0.53 ± 0.12	–2.33 ± 0.44	9	13
Cell type GI50 (ordinate)	# of cpnds	Pearson correlation coefficient	P >	Linear regression r^2	Slope	Intercept	% Compounds > + 1 sem from predicted value	% Compound < – 1 sem from predicted value
A549	57	0.65	0.001	0.42	0.39 ± 0.06	–2.13 ± 0.35	16	16
M14	56	0.67	0.001	0.45	0.40 ± 0.06	–2.09 ± 0.034	13	16
MOLT4	57	0.70	0.001	0.49	0.42 ± 0.06	–1.83 ± 0.35	18	14
CCRF-CEM	57	0.68	0.001	0.46	0.41 ± 0.06	–1.86 ± 0.36	16	16
P388	53	0.63	0.001	0.40	0.38 ± 0.07	–1.91 ± 0.40	11	13

LC50 data. A "fingerprint" analysis using the growth inhibitory data was conducted in 60 cell lines for 60,000 compounds and the data was found to contain patterns that "were remarkably rich" (16) and the data was found to accurately predict the mechanism of action for 141 known compounds with approximately 90% accuracy (17).

As seen in Table 4, the absolute predictive power of the *in vitro* data for *in vivo* toxicity was also rather weak. There

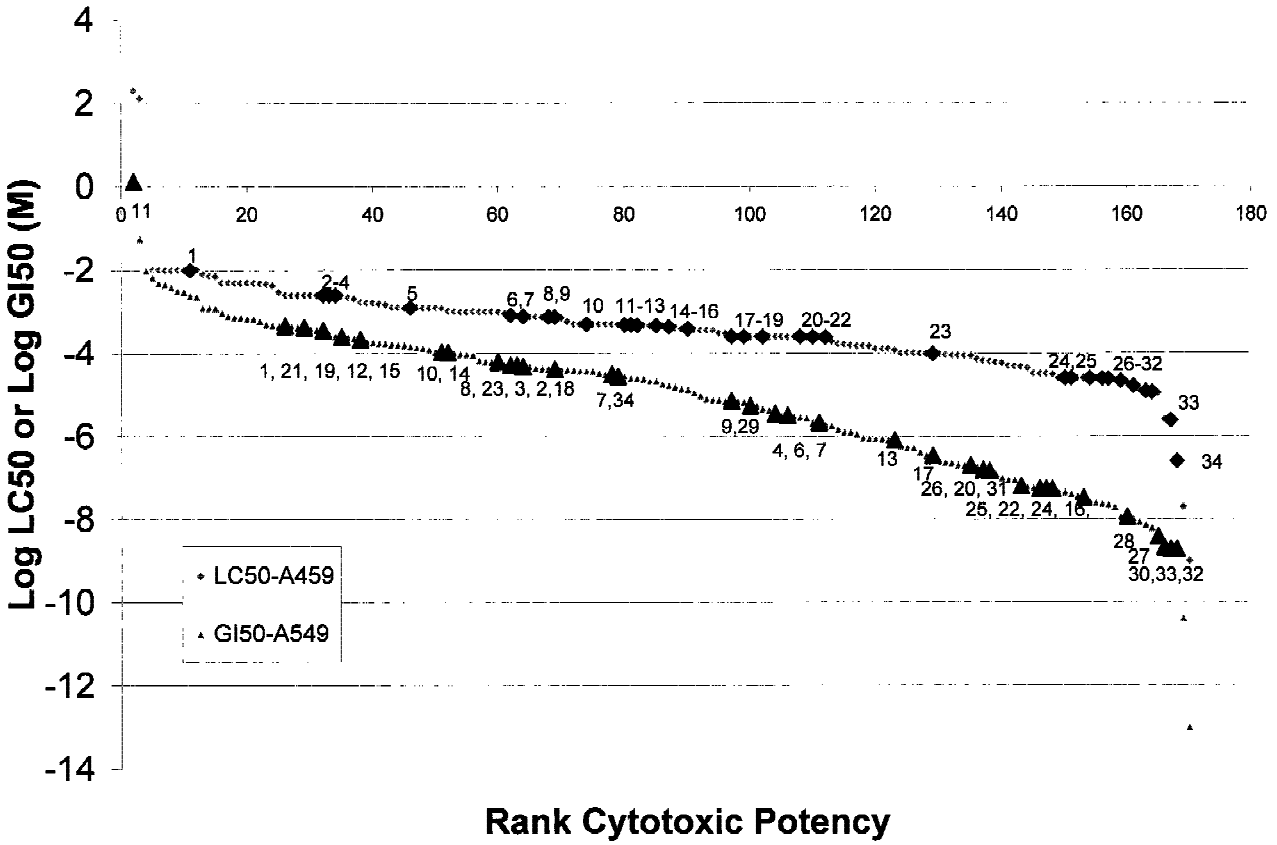

Fig. 3. Data for A549, LC_{50} (diamonds) and GI_{50} (triangles) plotted as a function of rank cytotoxic potency. The resultant curves are annotated for the position of approved antineoplastic drugs: PALA, 1; hydroxyurea, 2; 5-FUDR, 3; 5-FU, 4; fludarabine phosphate, 5; etoposide, 6; thio-tepa, 7; 6 mercaptopurine, 8; chlorambucil, 9; cytocine arabinoside, 10; procarbazine, 11; iphosphamide, 12; cisplatin, 13; BCNU, 14; vincristine sulfate, 15; CCNU, 16; methotrexate, 17; cyclophosphamide, 18; carboplatin, 19; busulphan, 20; melphelan, 21; thioguanine, 22; topotecan, 23; bleomycin, 24; paclitaxel, 25; doxorubicin , 26; mitomycin C, 27; teniposide, 28; tamoxifen, 29; daunorubicin , 30; amsacrine, 31; mitoxantrone, 32 vinblastin sulphate, 33; actinomycin D, 34.

are many possible reasons for the lack of correlation, foremost of which are the limitations inherent to continuous cell lines in culture. As described by Robert (18) in his critique of the NCI screening process, these include: limited drug-metabolic activity and the lack of influence of absorption; tissue disposition and elimination; lack of representation of the full range of the *in vivo* phenotype and limitations of the relevance of the *in vitro* endpoint to the *in vivo* pharmacology or toxicology.

Recommendation for a Primary Screen

Based on the results of our data mining study and the assumption that a primary screen for cytotoxicity will be deployed early in the drug discovery process, we feel some recommendations are warranted. In simplest terms, the definition of screening is straightforward, "a system for examining and separating into different groups" (19). In our experience in drug discovery, primary screening generally means selecting *for* an activity that is considered relevant to the desired therapeutic outcome, e.g., inhibiting a specific enzyme, and then choosing the most potent compounds of a series which express the selected activity. The stringency or *fineness of the mesh* of the screen is usually adjusted so as to provide no more than a manageable number of *hits*. In primary screening for desired activity, seemingly little concern is given to missing important leads that may be less potent. Such missed opportunities can be considered *false negatives*. Secondary or tertiary screens which address selectivity (often mislabeled specificity) are conducted to choose those hits which express a narrow range of pharmacologic activity. In this manner *false positives* are removed. It is only after a lead compound has passed a number of relatively low stringency screens that it will be scaled up for more extensive testing.

We propose that a similar approach can be taken for primary screening for toxicity and that *in vitro* cytotoxicity can serve this purpose in drug discovery and development. Applying this approach is a departure from widely held thinking about toxicity. The reason for this rests not in the relevance of the screen, but in the objectives of screening and when the screen is deployed in the Drug Discovery Cascade. A typical Drug Discovery Cascade is shown in Table 5. The cascade shown in the table, is intended to provide a framework for discussion and is not intended to be a definitive example representative of the strategy followed by any specific company, although the numbers and amounts are within an order of magnitude for most large pharmaceutical companies.

Based on the assumption that a primary screen for toxicity will be applied early in the Discovery Cascade, we can define some of the characteristics that a screen should have. High throughput (>100 compounds/yr); low compound requirement; and relatively low expense are self evident. Reproducibility over time is also essential so that data obtained from one set of compounds, particularly a validation or training set of compounds, can be used for classification of unknowns. This criterion can most easily be met by using banked cell lines. Prediction of potential for undesired *in vivo* effect is essential. If the screen cannot provide a forecast of some defined adverse effect it has no rationale and hence cannot be used to define rate of false positives and negatives.

Finally, defined rates of false positives and negatives are necessary so that a proper "weighting" of the screen data can be made when combining the results of the screen for toxicity with all the other data available for prioritization.

Although there were differences among the responses of the various tissues of origin and cell lines, when viewed globally the differences were not of great importance. We thus conclude that a single cell line is adequate for a primary screen and that the choice of the particular cell line is unlikely to be important.

The recommendation for choosing a single simple end point for a primary screen in drug discovery and development is based both on economic and scientific grounds. Regarding economics, as defined here, a primary screen must be high throughput, have a low compound requirement and be inexpensive to run. A similar conclusion has been reached to advocate using cell proliferation in a single "rapidly dividing transformed, undifferentiated cell line" as the first stage in acute toxicity testing, followed by more definitive models (20).

We feel that interpretation of GI50 data should depend on what other information is available about an individual compound. If nothing is known about the pharmacologic potency of the compound one can only make comparisons to compounds which have already been studied in depth, a *reference set* (or more loosely a *validation set*). The reference set could be studied entirely *de novo* or, preferably, be built up as an extension of an existing data set. This is the principle reason why we would advocate using one of the cell lines in the DTP screen. In this way the benefits of huge investment made to date in over 60,000 compounds by the DTP can be reaped by other laboratories.

An example of how such a reference set can be used is illustrated by Figure 3 which highlights data for 35 marketed

Table 5. The Drug Discovery Cascade

Stage	Research activity	Objectives	Number of compounds to be dealt with	Number of compounds emerging	Amount of compound available for study
I	Combinatorial Chemistry or Natural Products	Generate Candiate Pharmacophores	100K–1M	100K–1M	1–10 mg
II	Primary Screening for Pharmacology	Identify Candidate Pharmacophores Hits	100K–1M	1–100	
III	Primary Medicinal Chemistry	Expand of Number of Pharmacophores	1–100	100–10K	100 mg
IV	Secondary Screening for Pharmacology	Pharmacologic Selectivity of Hits	100–10K	1–100	
V	Secondary Medicinal Chemistry	Expand Supply of Hits	1–100	1–100	1–10 g
VI	Tertiary Pharmacology Testing	Lead Compound Selection	1–100	1–10	
VII	Synthetic Chemistry	Supply of Lead Compound	1–10	1–10	10–100 g

antineoplastic drugs. Only 5 of 35 marketed antineoplastics have log GI50 values less than -8. Thus, if no other information is available, setting a GI50 of 10^{-8} M as an initial cut off will probably avoid *throwing away the baby with the bath water* for most therapeutic targets. This cut off value would, of course be adjusted so as to keep the numbers of compounds "passing" the screen manageable and could be adjusted as additional information became available about the chemical or therapeutic class of compounds under study.

In conclusion, our datamining study of the DTP Standard Agents database has shown that when viewed in the context of a primary screen for pharmaceutical toxicity, the choice of cell line for the system may not be of critical importance and that a simple system for measuring growth inhibition can be of value in ranking compounds on the basis of toxic potential. Moreover, in the absence of efficacy data a log GI50 $\cong$ –8 M is a good starting point for a cut off value for accepting a compound for scale up and further study.

ACKNOWLEDGMENTS

The authors of this paper would like to thank the NCI for producing a valuable database (DTP) and making it publicly available. This work was supported by SmithKline Beecham Pharmaceuticals.

REFERENCES

1. D. Lednicer. Exploration of a chemical series as a drug discovery strategy. *Drug Design Del.* **78**:2–69 (1987).
2. K. S. Lam. Application of combinatorial library methods in cancer research and drug discovery. *Anti-Cancer Drug Design* **12**: 145–167 (1997).
3. D. R. Algate. Application of the pharmacological screening process. *Drug Metab. Rev.* **22**:809–820 (1990).
4. R. A. Fecik, K. E. Frank, E. J. Gentry, S. R. Menon, L. A. Mitscher, and H.Telikepalli. The search for orally active medications through combinatorial chemistry. *Med. Res. Rev.* **18**:149–185 (1998).
5. A. D. Rodrigues. Preclinical drug metabolism in the age of high-throughput screening: An industrial perspective. *Pharm. Res.* **14**: 1504–1510 (1997).
6. G. Klopman and H. S. Rosenkranz. International Commission for Protection Against Environmental Mutagens and Carcinogens. Approaches to SAR in carcinogenesis and mutagenesis. Prediction of carcinogenicity/mutagenicity using MULTI-CASE. *Mutation Res.* **305**:33–46 (1994).
7. A. Guillouzo. Liver cell models in *in vitro* toxicology. *Environ Health Pers.* **106**(supp 2): 511–532 (1998).
8. Z. Djuric, T. H. Corbett, F. A. Valeriote, L. K. Heilbrun, and L. H. Baker. Detoxification ability and toxicity of quinones in mouse and human tumor cell lines used for anticancer drug screening. *Cancer Chemother. Pharmaco.* **36**:20–26 (1995).
9. R. G. Ulrich, J. A. Bacon, C. T. Cramer, G. W. Peng, D. K. Petrella, R. P. Stryd, and E. L. Sun. Cultured hepatocytes as investigational models for hepatic toxicity; practical applications in drug discovery and development. *Toxicol. Lett.* **82**:107–115 (1995).
10. F. Ballet. Hepatotoxicity in drug development: detection, significance and solutions. *J. Hepatol.* **26**(supp 2):26–36 (1997).
11. J. W. Harbell, S. W. Koontz, R. W. Lewis, D. Lovell, and D. Acosta. IRAG working group 4. Cell cytotoxicity assays. Interagency regulatory alternatives Group. *Food Chem. Toxicol.* **35**: 79–126 (1997).
12. A. Monks, D. A. Scudiero, G. S. Johnson, K. D. Paull, and E. A. Sausville. The NCI anti-cancer drug screen: a smart screen to identify effectors of novel targets. *Anti-Cancer Drug Design* **12**: 533–541 (1997).
13. M. Alvarez, R. Robey, V.Sandor, K.Nishiyama, Y.Matsumoto, K. Paull, S. Bates, and T. Fojo. Using the National Cancer Institute anticancer drug screen to assess the effects of MRP expression on drug sensitivity profiles. *Mol. Pharm.* **54**:802–814 (1998).
14. E. S. Woo, A. Monks, S. C. Watikins, A. S. Wang, and J. S. Lazo. Diversity of metallothionine content and subcellular localization in the National Cancer Institute tumor panel. *Cancer Chemother. Pharmacol.* **41**:61–68 (1997).
15. K. D. Paull, R. H. Shoemaker, L. Hodes, A. Monks, D. A. Scudiero, L. Rubinstein, J. Plowman, and M. R. Boyd. Display and analysis of patterns of differential activity of drugs against human tumor cell lines: development of Mean Graph and COMPARE algorithm. *J Nat. Cancer Inst.* **81**:1088–1092 (1989).
16. J. N. Weinstein, T. G. Myers, P. M. O'Connor, S. H. Friend, A. J. Fornace Jr., K. W. Kohn, T. Fojo, S. E. Bates, L. V. Rubinstein, N. L. Anderson, J. K. Buolamwini, W. W. van Osdol, A. P. Monks, D. A. Scudiero, D. A. Sausville, D. W. Zaharvitz, B. Bunow, V. N. Viswanadhan, G. S. Johnson, R. E. Wittes, and K. D. Paul. An information-intensive approach to the molecular pharmacology of cancer. *Science* **275**:343–349 (1997).
17. A. D. Koutsoukos, L. V. Rubinstein, D. Faraggi, R. M. Simon, S. Kalyandrug, J. N. Weinstein, K. W. Kohn, and K. D. Paull. Discrimination techniques applied to the NCI *in vitro* anti-tumor drug screen: predicting biochemical mechanisms of action. *Statistics Med.* **13**:719–730 (1994).
18. J. Robert. Rôle des modèles *in vitro* dans l'évaluation préclinique das médicaments anticancéreux. *Bull. Cancer* **83**:801–808 (1996).
19. H. B. Woolf, editor. *Webster's New Collegiate Dictionary,* G&C Merriam Co., Springfield, MA 1981.
20. H .Siebert, M. Balls, J. H. Fentem, V. Bianchi, R. H. Clothier, P. J. Dierckx, B. Ekwall, M. J. Garle, M. J. Gómez-Lechón, Gribaldo, M. Gülden, M. Liebsch, E. Rasmussen, R. Roguet, R. Shrivatava, and E. Walum. Acute toxicity testing *in vitro* and the classification and labeling of chemicals. *ATLA Alternative Lab. Animal J.* **24**:499–510 (1996).
21. M. Rozencweig, D. D. Von Hoff, M. J. Staquet, P. S. Schein, J. S. Penta, A. Goldin, F. M. Muggia, E. J. Freireich, and V. T. DeVita Jr. Animal toxicology for early clinical trials with anticancer agents. *Cancer Clin. Trials* **4**:21–28 (1981).
22. J. S. Penta, M. Rozencweig, A. M. Guarino, and F. M. Muggia. Mouse and large-animal toxicology studies of twelve antitumor agents: relevance to starting dose for Phase I clinical trials. *Cancer Chemother. Pharmacol.* **3**:97–101 (1979).
23. M. A. Goldsmith, M. Slavik, S. K. Carter. Quantitative prediction of drug toxicity in humans from toxicology in small and large animals. *Cancer Res.* **35**:1354–1364 (1975).
24. E. Fuse, S. Kobayashi, M. Inaba, and H. Suzuki. Application of pharmacokinetically guided dose escalation with respect to cell cycle phase specificity. *J. Nat. Cancer Inst.* **86**:989-996 (1994).
25. D. P. Griswold, W. R. Laster Jr., M. Y. Snow, F. M. Schabel Jr., and H.E. Skipper. Experimental evaluation of potential anticancer agents. XII. Quantitative drug response of the SA180, CA755 and leukemia L1210 systems to a "standard list" of active and "inactive" agents. *Cancer Res.* **23**:271-298 (1963).
26. R .Taetle, F. Rosen, I. Abramson, J. Venditti, and S. Howell. Use of nude mouse xenograft as preclinical drug screens: *in vivo* activity of established chemotherapeutic agents against melanoma and ovarian carcinoma xenografts. *Cancer Treat. Reports* **71**:297–304 (1987).
27. I. Fichtner, W. Arnold, and H. Naundorf. Toxicologic and therapeutic examinations with ftorafur in different mouse strains. *Neoplasma.* **28**:133–140 (1981).
28. E. J. Freireich, E. A. Gehan, D. P. Rall, L. H. Schmidt, and H. E. Skipper. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey and man. *Cancer Chemother. Reports* **50**:219–245 (1966).
29. L. B.Grochow, E. K. Rowinsky, R. Johnson, S. Ludeman, S. H. Kaufmann, F. L. McCabe, B. R. Smith, L. Hurwitz, A. Delisa, R. C. Donehower, and D. A. Noe. Pharmacokinetics and pharmacodynamics of topotecan in patients with advanced cancer. *Drug Metab. and Dispos.* **20**:706–713 (1992).