

# Nowicky Pharma

Dipl.-Ing. DDr. Dr.h.c. W. Nowicky



Margaretenstraße 7  
A-1040 Vienna, Austria  
tel.: + 43-1-586 12 24  
fax: + 43-1-586 89 94  
nowicky@ukrin.com

## Results

### From National Cancer Institute Human Tumor Cell Line Screen

## Ukrain

## Chelidonine Hydrochloride

## Thiotepa

Results  
From National Cancer Institute  
Human Tumor Cell Line Screen

Ukrain

Chelidonine Hydrochloride

Thiotepa



National Institutes of Health  
National Cancer Institute  
Bethesda, Maryland 20892  
Telex: 908111

September 29, 1990

Dr. J. Wassyl Nowicky  
Ukrain Anti-Cancer Inst.  
Laimgrubengasse 19/7  
A-1060 Wien  
Austria

Dear Dr. Nowicky,

The NCI's new investigational *in vitro* disease-oriented primary antitumor screen has now entered a fully operational phase. A recent description of the rationale, background, current status and other details are provided in the enclosed review from PPO Updates. Currently, the cell panel consists of 60 lines against which compounds are tested at a minimum of five concentrations at 10-fold dilutions. A 48 hour continuous drug exposure protocol is used, and a sulforhodamine B (SRB) protein assay is used to estimate cell viability or growth.

The development and applications of numerical and statistical methods of screening data analysis and refinement of selection criteria is continuing. However, we have adopted, on an interim basis, a set of data display and analysis procedures described briefly in the attachments provided. Pilot studies performed during the past 1 1/2 years have provided the basis for calibration of the screen with respect to known standards and the development of methods to facilitate the detection of patterns of interest.

The checklist includes observations of the mean response parameters, differential cellular sensitivity, and any subpanel-specific patterns of sensitivity. Enclosed are data obtained in the current operational screening protocol on compound(s) submitted by you recently and/or during the earlier development or pilot phases of the new *in vitro* screen. Screening program staff have reviewed the data and have provided the attached summary checklist for your information. Also indicated is any further pending action by NCI.

Our staff welcomes your comments, criticisms, and advice concerning the information provided and the display formats employed. If you have specific questions concerning the enclosed data, or the status of your compounds, please contact Dr. K. Paull, Chief, Information Technology Branch (Fax 301-496-8333) or Dr. Ven Narayanan, Chief, Drug Synthesis and Chemistry Branch (Phone 301-496-8795, Fax 301-496-8333).

Michael R. Boyd MD PhD

National Cancer Institute Developmental Therapeutics Program  
In-Vitro Screening Data Review Checklist

NSC: 631570-W/ 1	Experiment ID: 9006NS33	Source:
Test Date: June 27, 1990	Review Date: 10/5/90 A✓	

Review Summary

1. MG\_MID Response Parameters

a. Log <sub>10</sub> GI50	<u>-5.55</u>
b. Log <sub>10</sub> TGI	<u>-4.80</u>
c. Log <sub>10</sub> LC50	<u>-4.17</u>

2. Selectivity Analysis

a. Differential Cellular Sensitivity

(1) GI50	<u>0.52</u>
(2) TGI	<u>0.78</u>
(3) LC50	<u>0.79</u>

(Low [ $\Delta < 1$ ]=L, Mod [ $1 < \Delta < 3$ ]=M, High [ $\Delta > 3$ ]=H)

b. Differential Subpanel Sensitivity

Subpanels\* showing a statistical measure of differential sensitivity (see definition and derivation attached) with respect to the indicated response parameter (GI50, TGI, or LC50)

Response Parameter	Subpanel Specificity		
(1) GI50	LEU	MEL	SCL
(2) TGI	SCL	_____	_____
(3) LC50	SCL	_____	_____

\* Tumor cell line subpanels are identified as follows:

CNS = Brain;	COL = Colon;	LEU = Leukemia/Lymphoma;
LNS = Non-small Cell Lung;	MEL = Melanoma;	MIS = Miscellaneous
OVA = Ovary;	REN = Kidney;	SCL = Small Cell Lung;

Pending Action by NCI

1. \_\_\_\_\_ None
2. \_\_\_\_\_ Repeat testing in Primary Screen  
5 Log<sub>10</sub> \_\_\_\_\_; 10 Log<sub>10</sub> \_\_\_\_\_; Other \_\_\_\_\_.
3. \_\_\_\_\_ Refer to Biological Evaluation Committee
4. \_\_\_\_\_ Refer to Natural Products Program Committee

# National Cancer Institute Developmental Therapeutics Program

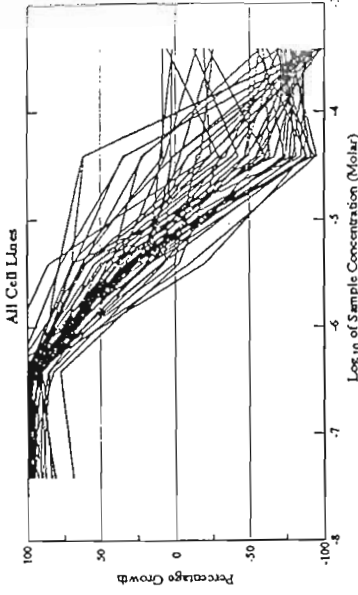
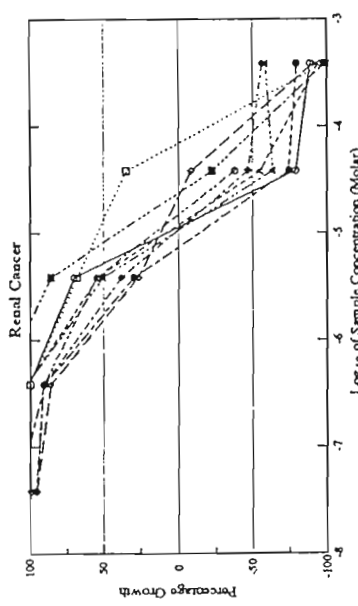
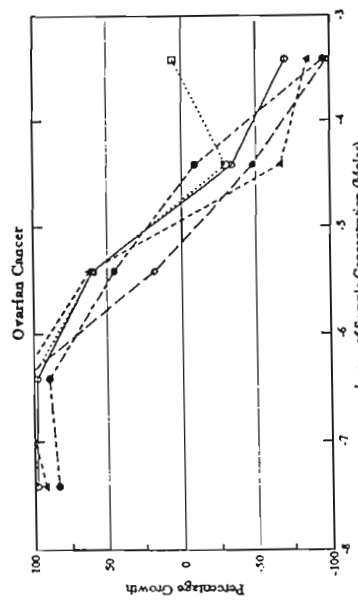
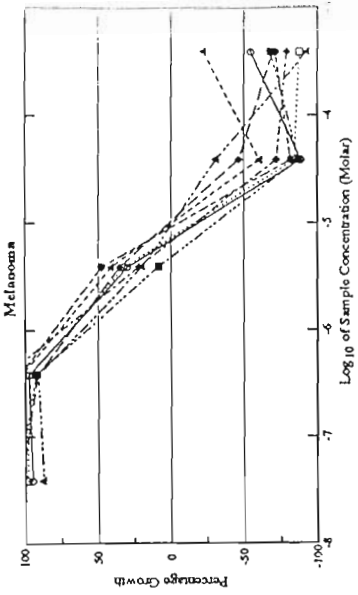
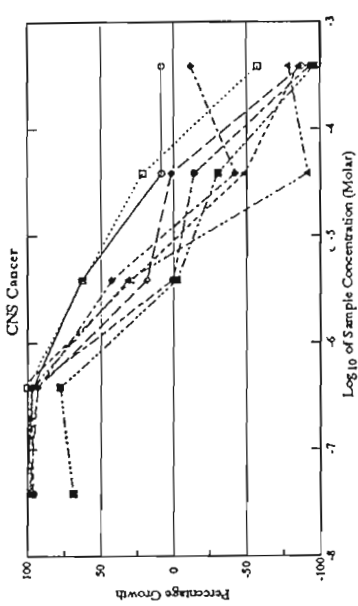
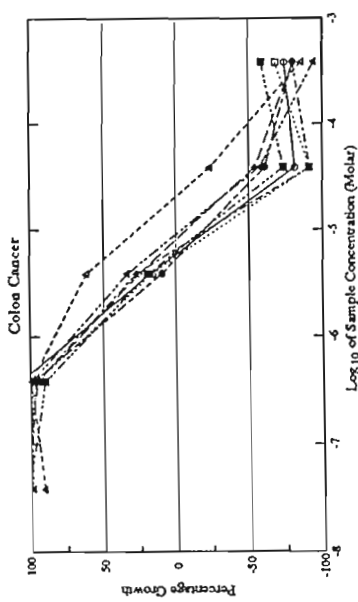
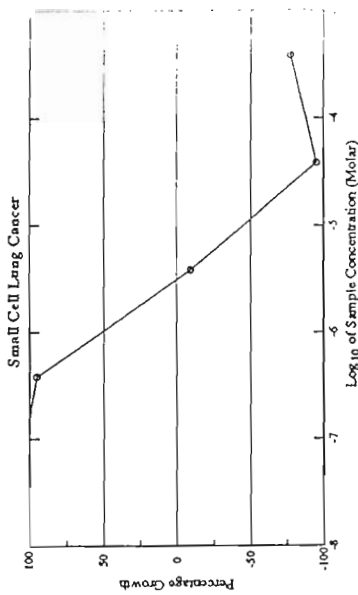
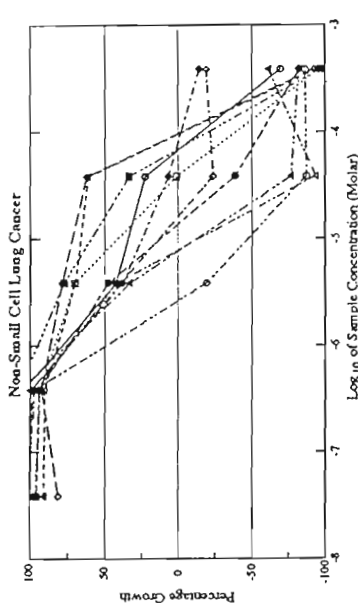
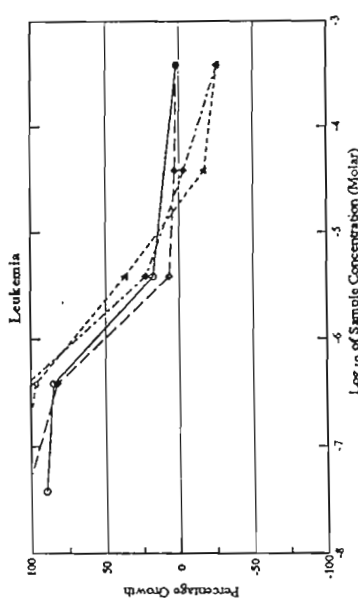
## Dose Response Curves

NSC: 631570-W/ 1

SSPL: V27D Exp. ID: 9006NS3

Report Date: September 29, 1990

Test Date: June 25, 1990

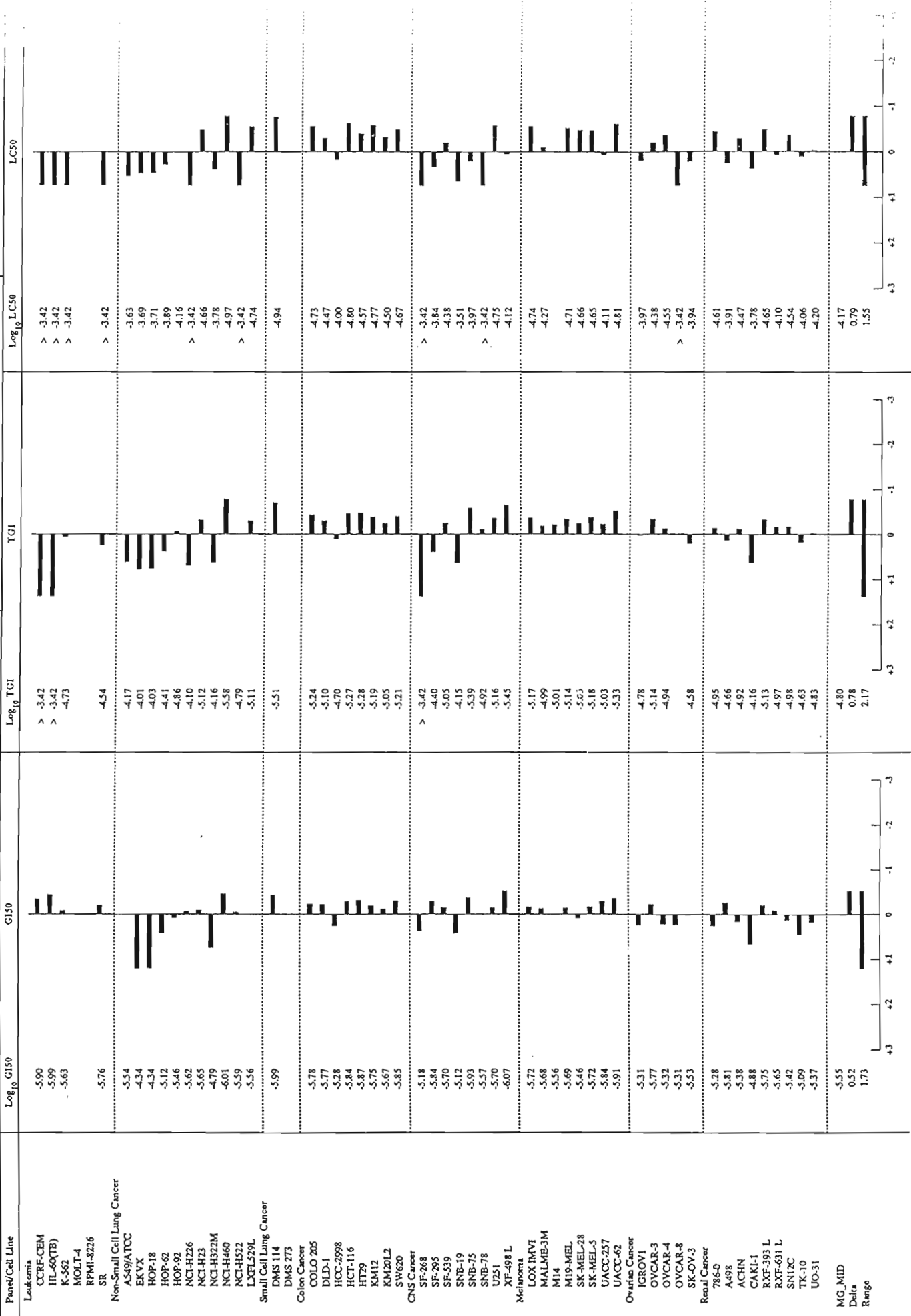




# National Cancer Institute Developmental Therapeutics Program Mean Graphs

NSC: 631570-W/ 1      Units: Molar  
Report Date: September 29, 1990

SSPL: V27D      Exp. ID: 9306(N33)  
Test Date: June 25, 1990



Dose Response Matrix

Test Date: June 25, 1990

Stain: PRO

Report Date: September 29, 1990

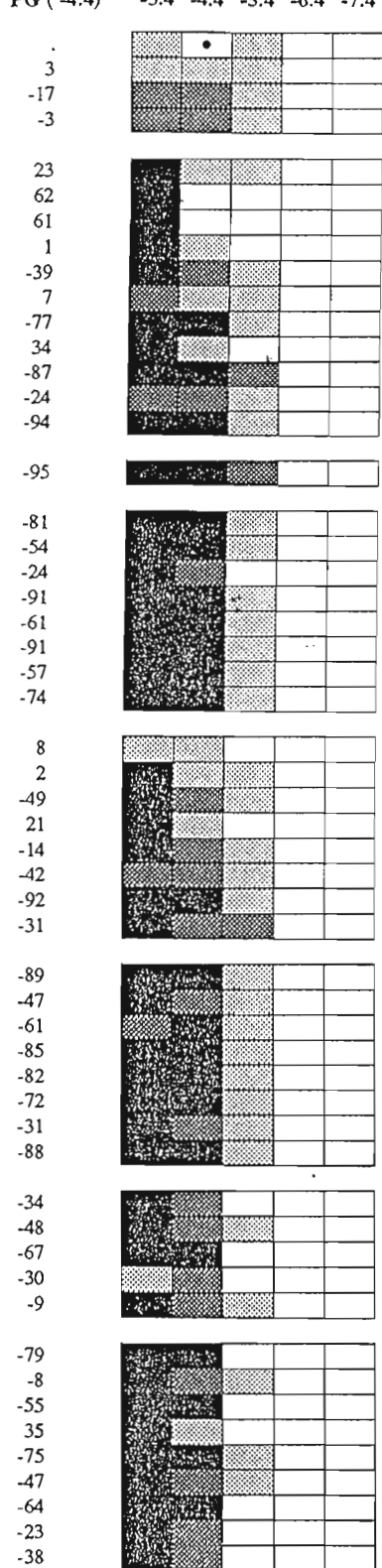
SSPL: V27D

Log<sub>10</sub> Concentration (Molar)  
PG (-4.4) -3.4 -4.4 -5.4 -6.4 -7.4

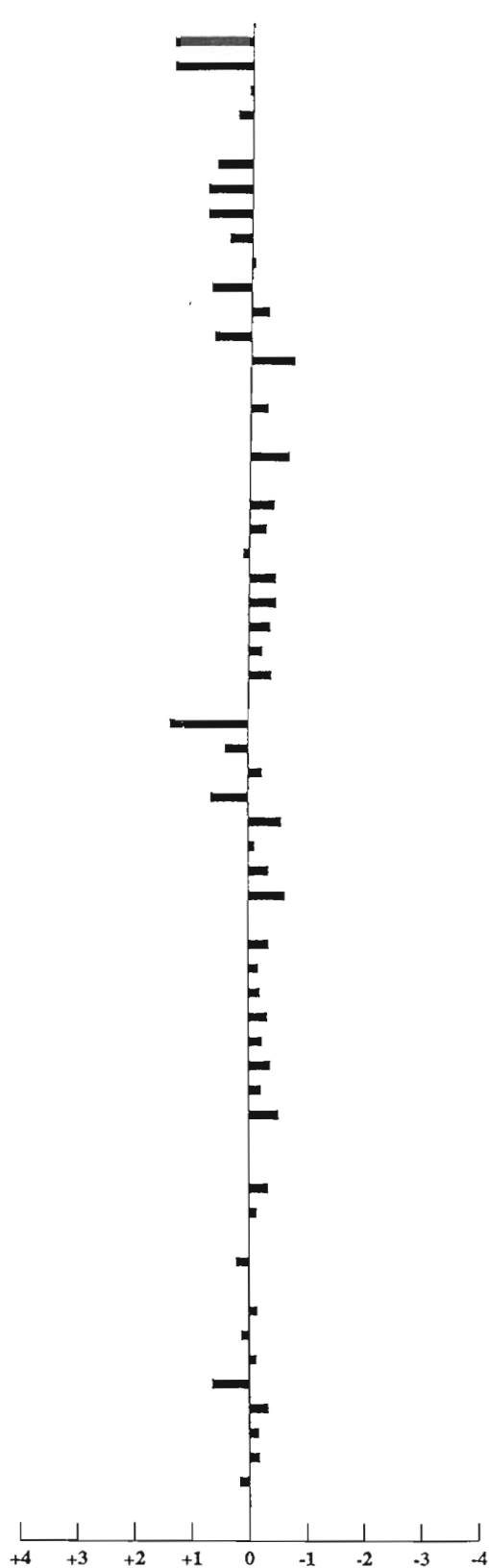
Panel/Cell Line

Log<sub>10</sub> TGI

TGI



Panel/Cell Line	Log <sub>10</sub> TGI
<b>Leukemia</b>	
CCRF-CEM	> -3.42
HL-60(TB)	> -3.42
K-562	-4.73
SR	-4.54
<b>Non-Small Cell Lung Cancer</b>	
A549/ATCC	-4.17
EKVX	-4.01
HOP-18	-4.03
HOP-62	-4.41
HOP-92	-4.86
NCI-H226	-4.10
NCI-H23	-5.12
NCI-H322M	-4.16
NCI-H460	-5.58
NCI-H522	-4.79
LXFL529L	-5.11
<b>• Small Cell Lung Cancer •</b>	
DMS 114	-5.51
<b>Colon Cancer</b>	
COLO 205	-5.24
DLD-1	-5.10
HCC-2998	-4.70
HCT-116	-5.27
HT29	-5.28
KM12	-5.19
KM20L2	-5.05
SW620	-5.21
<b>CNS Cancer</b>	
SF-268	> -3.42
SF-295	-4.40
SF-539	-5.05
SNB-19	-4.15
SNB-75	-5.39
SNB-78	-4.92
U251	-5.16
XF-498 L	-5.45
<b>Melanoma</b>	
LOX IMVI	-5.17
MALME-3M	-4.99
M14	-5.01
M19-MEL	-5.14
SK-MEL-28	-5.05
SK-MEL-5	-5.18
UACC-257	-5.03
UACC-62	-5.33
<b>Ovarian Cancer</b>	
IGROV1	-4.78
OVCAR-3	-5.14
OVCAR-4	-4.94
OVCAR-8	-4.94
SK-OV-3	-4.58
<b>Renal Cancer</b>	
786-0	-4.95
A498	-4.66
ACHN	-4.92
CAKI-1	-4.16
RXF-393 L	-5.13
RXF-631 L	-4.97
SN12C	-4.98
TK-10	-4.63
UO-31	-4.83



PG > GI50  
PG ≤ GI50  
PG ≤ TGI  
PG ≤ LCS0

D<sub>GI50</sub> = 39.0 (-5.4)  
D<sub>TGI</sub> = 96.2 (-5.4)  
D<sub>LCS0</sub> = 57.7 (-4.4)  
D<sub>H</sub> = 100 (-4.4)

MG\_MID TGI = -4.80  
Delta = 0.78  
Range = 2.17  
MGD<sub>H</sub> = 96.15





National Institutes of Health  
National Cancer Institute  
Bethesda, Maryland 20892

December 2, 1991

J. W. Nowicky  
Laimgrubengasse 19/7  
A-1060 Vienna  
Austria

Dear Dr. Nowicky:

In response to your letters of November 10 and 26, 1991, to Dr. Broder, I am enclosing the data you requested for doxorubicin (NSC 123127), daunomycin (NSC 82151), 5-fluorouracil (NSC 19893), and your material Ukrain (NSC 631570).

The information supplied is in the form of the standard supplier reporting package for the three most recent experiments on each of the standard agents and two for your material. The supplier report packages include: In Vitro Testing Results, Dose Response Curves, Mean Graphs, and the Dose Response Matrix.

The original review summary page of the supplier report package has been phased out, since it was felt to cause some confusion in interpretation. The items you requested in your outline marked A. 2. a. 1, 2, 3, and A. 2. b. are those that were primarily responsible for the confusion and were therefore phased out of the summary page. However, the items in your outline marked A. 1. a, b, c (the MG\_MID and the LOG10 values for every cell line at the GI50, TGI, and LC50 levels of effect) continue to be available in the Mean Graphs.

If I can be of any further assistance, please let me know.

Sincerely,

Saul A. Schepartz, Ph.D.  
Deputy Associate Director  
Developmental Therapeutics Program  
Division of Cancer Treatment  
National Cancer Institute

cc: Dr. Broder  
Dr. Chabner

# National Cancer Institute Developmental Therapeutics Program

## Dose Response Curves

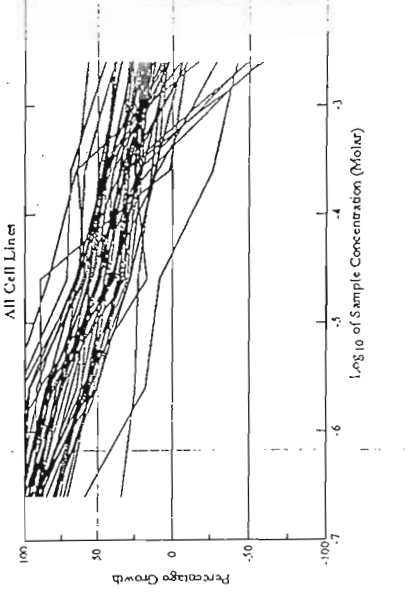
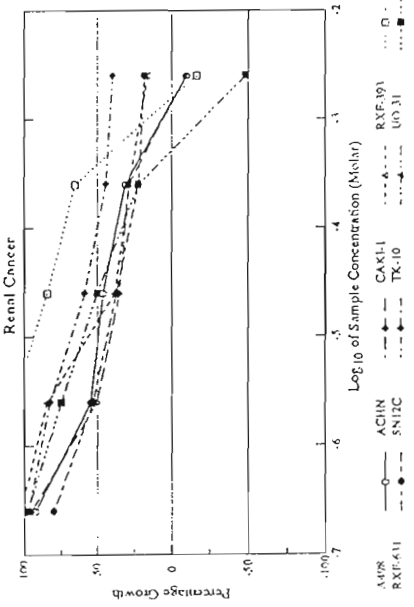
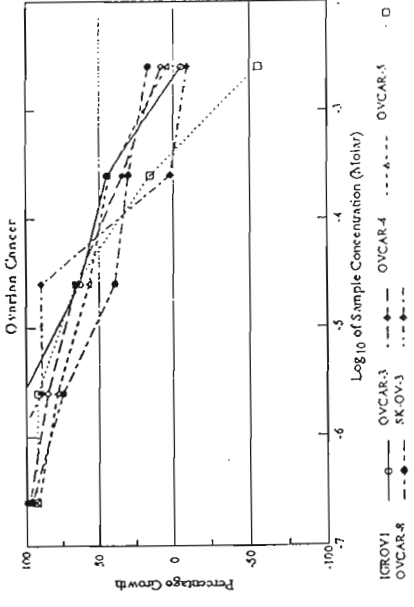
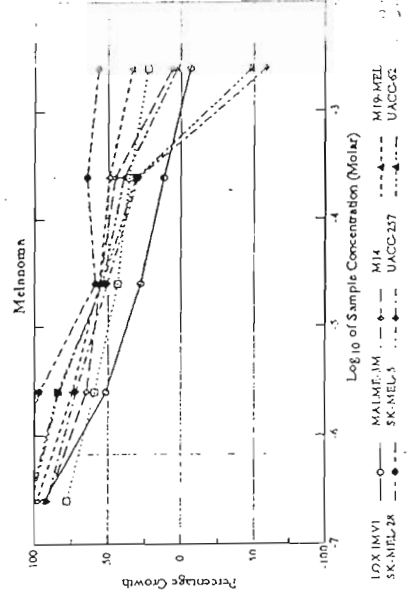
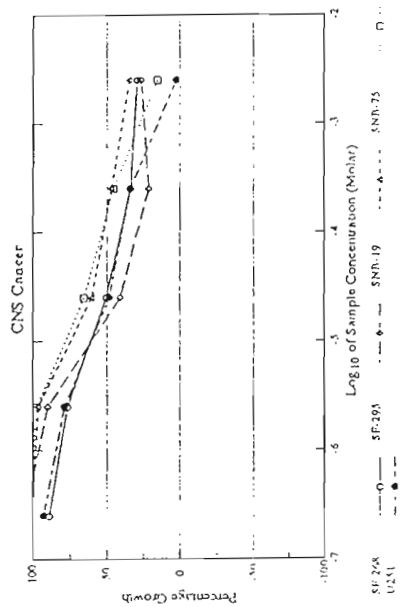
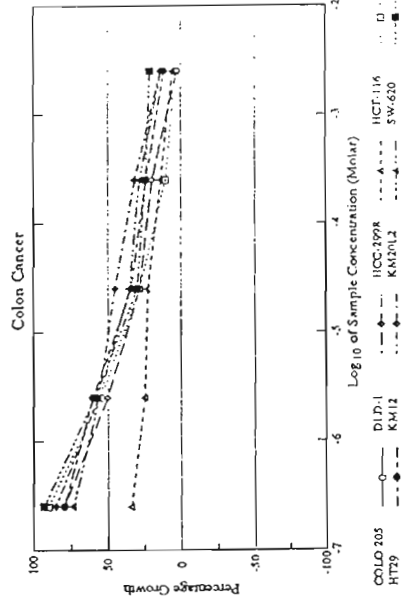
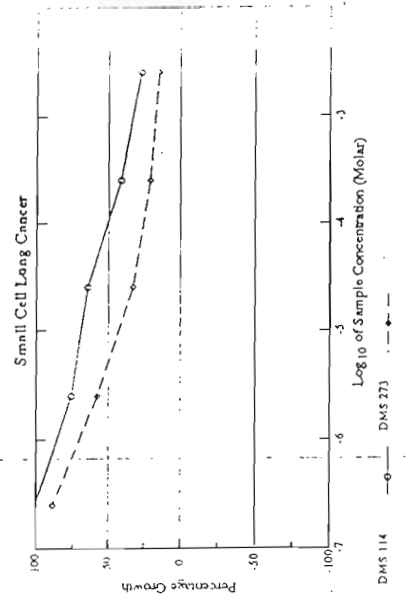
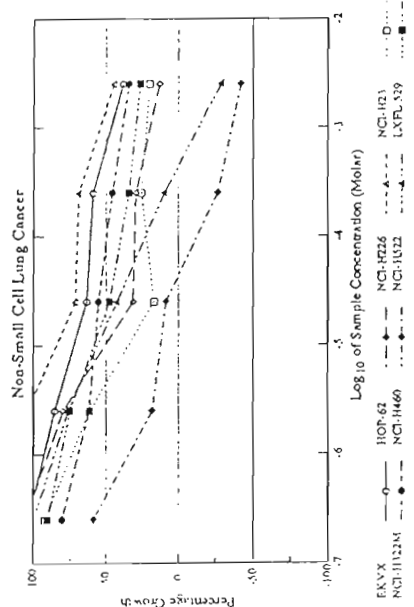
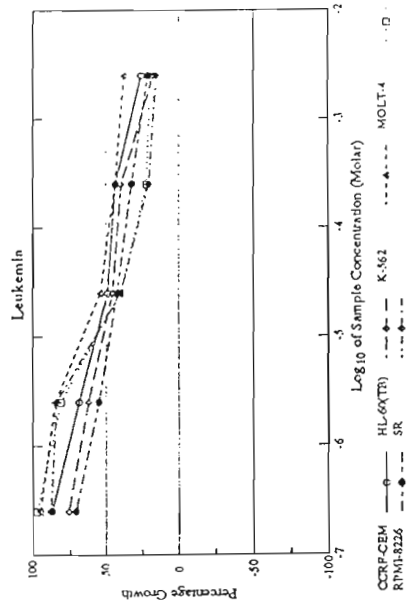
NSC: 19893 -G/10

SSPL: Q87M

Exp. ID: 91115N43

Report Date: December 2, 1991

Test Date: November 4, 1991



NCI Developmental Therapeutics Program  
Dose Response Matrix

NSC: 19893 -G/10

Exp. ID: 9111SR43

Test Date: November 4, 1991

Stain: PROTEIN-51

Report Date: December 2, 1991

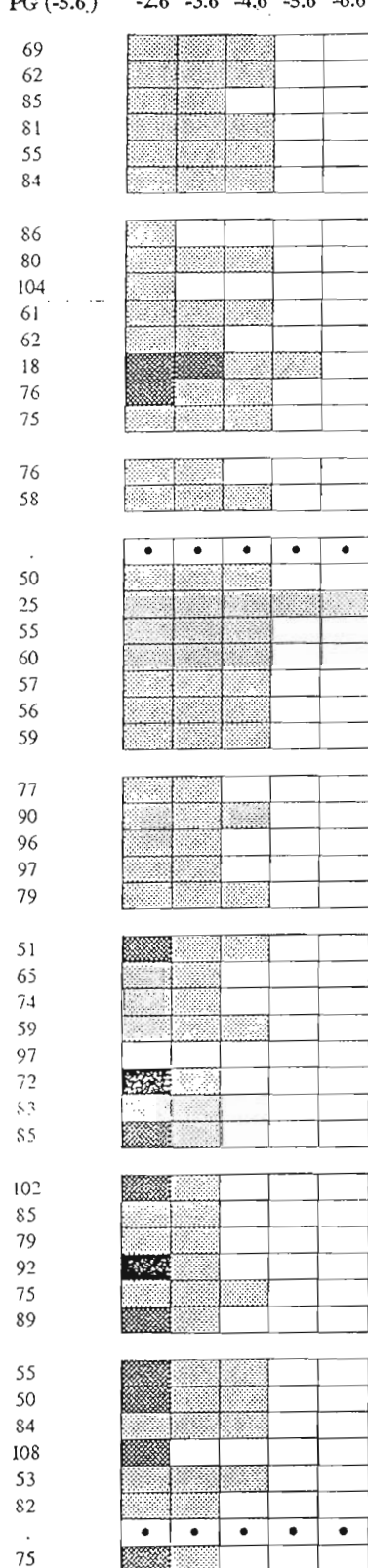
SSPL: Q87M

Log<sub>10</sub> Concentration (Molar)  
PG (-5.6) -2.6 -3.6 -4.6 -5.6 -6.6

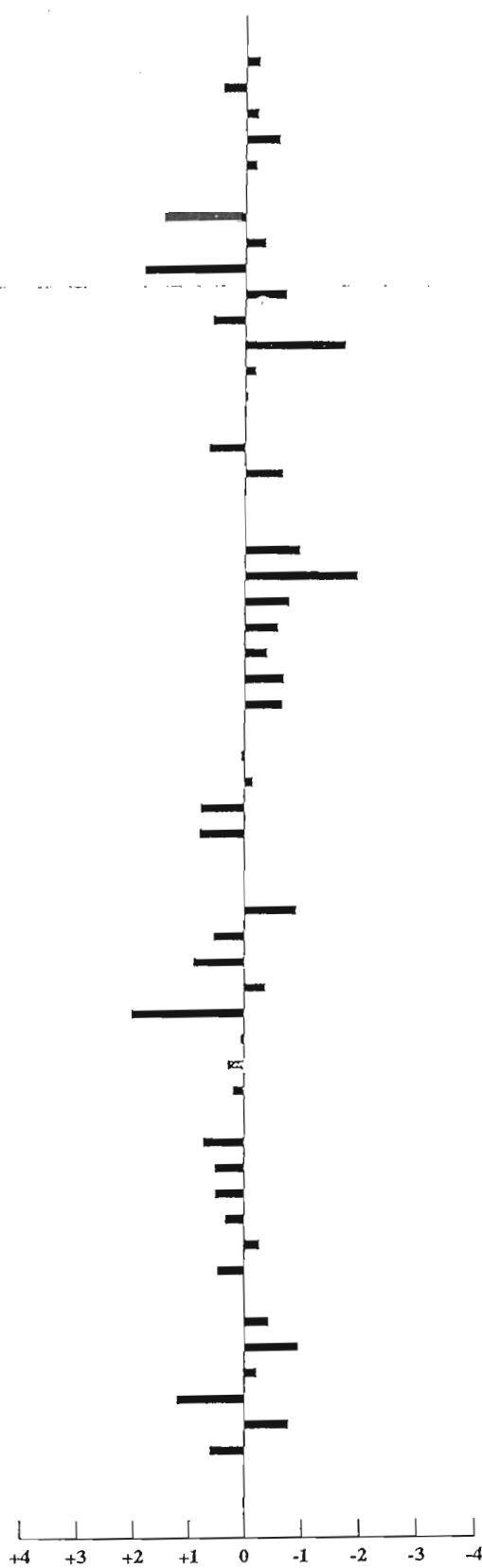
Panel/Cell Line

Log<sub>10</sub> GI50

GI50



Panel/Cell Line	Log <sub>10</sub> GI50
<b>Leukemia</b>	
CCRF-CEM	-4.66
HL-60(TB)	-4.87
K-562	-4.21
MOLT-4	-4.84
RPMI-8226	-5.22
SR	-4.83
<b>Non-Small Cell Lung Cancer</b>	
EKVX	-3.17
HOP-62	-4.99
NCI-H226	-2.82
NCI-H23	-5.35
NCI-H322M	-4.05
NCI-H460	-6.39
NCI-H522	-4.82
LXFL 529	-4.69
<b>Small Cell Lung Cancer</b>	
DMS 114	-3.98
DMS 273	-5.29
<b>Colon Cancer</b>	
COLO 205	
DLD-1	-5.60
HCC-2998	< -6.60
HCT-116	-5.41
HT29	-5.22
KM12	-5.03
KM20L2	-5.32
SW-620	-5.30
<b>CNS Cancer</b>	
SF-268	-4.57
SF-295	-4.79
SNB-19	-3.84
SNB-75	-3.83
U251	-4.66
<b>Melanoma</b>	
LOX IMVI	-5.54
MALME-3M	-4.07
M14	-3.71
M19-MEL	-5.02
SK-MEL-28	> -2.60
SK-MEL-5	-4.57
UACC-257	-4.34
UACC-62	-4.43
<b>Ovarian Cancer</b>	
IGROV1	-3.89
OVCAR-3	-4.11
OVCAR-4	-4.11
OVCAR-5	-4.29
OVCAR-8	-4.91
SK-OV-3	-4.15
<b>Renal Cancer</b>	
A498	-5.07
ACHN	-5.58
CAKI-1	-4.86
RXF-393	-3.42
RXF-631	-5.42
SN12C	-4.01
TK-10	
UO-31	-4.60



PG > GI50  
PG ≤ GI50  
PG ≤ TGI  
PG ≤ LC50

D<sub>GI50</sub> = 52.0 (-4.6)  
D<sub>TGI</sub> = 40.0 (-2.6)  
D<sub>LC50</sub> = 14.0 (-2.6)  
D<sub>H</sub> = 76.19 (-5.6)

MG<sub>MID GI50</sub> = -4.63  
Delta = 1.97  
Range = 4.00  
MG<sub>DH</sub> = 79.59



## DIFFERENT ACTION OF UKRAIN, CELANDINE ALKALOIDS AND THIOTEPA ON CANCER CELL LINES

Ukrain is produced from alkaloids from greater celandine and Thiotepa.

The National Cancer Institute (Bethesda, Maryland, USA) has proved that Ukrain (NSC 631570) has a completely different effect on malignant cells to Thiotepa (NSC 6396) and chelidonine hydrochloride (NSC 406034; chelidonine is the main alkaloid of greater celandine).

For example:

- UKRAIN is at its least effective [ $\log(\text{TGI}) = -3.4$ ], with leukaemia-HL-60(TB) in contrast to chelidonine hydrochloride, which is very effective [ $\log(\text{TGI}) = -5.4$ ] and to Thiotepa, which is only moderately effective [ $\log(\text{TGI}) = -4.4$ ].
- UKRAIN is extremely effective [ $\log(\text{TGI}) = -5.6$ ] with Non-Small Lung-NCI-H460, chelidonine hydrochloride less effective [ $\log(\text{TGI}) = -4.0$ ] and Thiotepa shows very little effect [ $\log(\text{TGI}) = -4.5$ ].
- With Colon-SW-620 UKRAIN is very effective [ $\log(\text{TGI}) = -5.2$ ], chelidonine hydrochloride is not effective [ $\log(\text{TGI}) = -4.0$ ], and Thiotepa is also not effective [ $\log(\text{TGI}) = -4.2$ ].

The enclosed profiles of these three different substances show very clearly that their effects on the same cell lines are very different.

Results of the National Cancer Institute, Bethesda, USA, Human Cell Line Screen can be seen on the website of the Developmental Therapeutics Program NCI/NIH (National Cancer Institute, National Institute of Health) <http://dtp.nci.nih.gov/>.

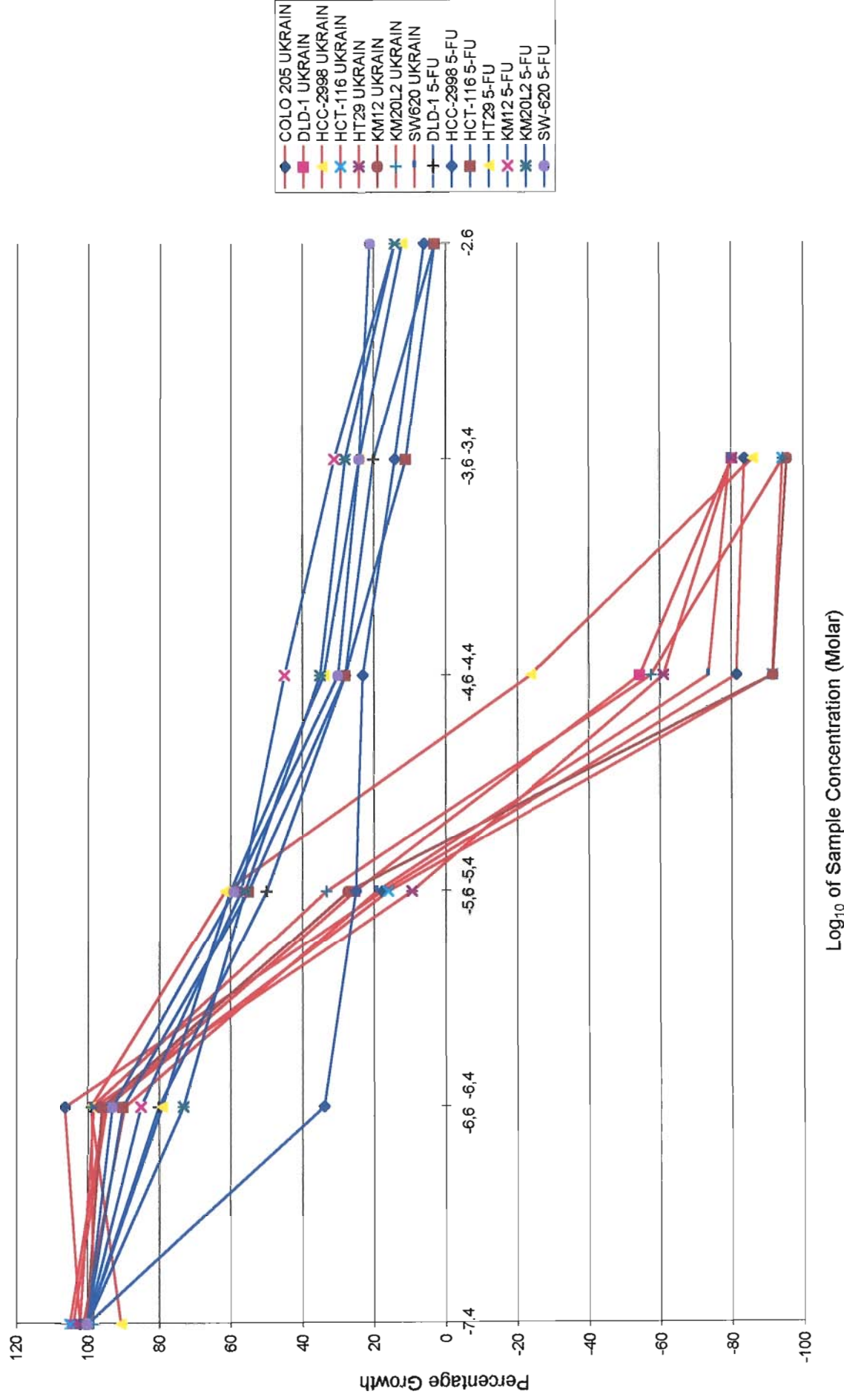
However, we believe that normal cell lines should also be used in screening tests. It is remarkable that neither the NCI nor EORTC (European Organization for Research and Treatment of Cancer) use healthy cell lines in their screening tests. The reason may be that within the framework of this programme tests are made almost exclusively on very toxic substances which have no selective effect and which are known in advance to be damaging to normal cells.

Prof. A. Liepins was the first also to examine normal cells. After observing very promising results he suggested to the NCI that normal cells should also be used in their screening programme. Unfortunately this suggestion has met with no response.

# NATIONAL CANCER INSTITUTE, BETHESDA, MARYLAND, USA

UKRAIN, NSC 631570      5-FU, NSC 19893

## Cytotoxicity of Ukrain and 5-FU on different human colorectal cancer cell lines



# TGI Mean Graph for Compound S631570

## Results from the National Cancer Institute Human Tumor Cell Line Screen

These results provided by the Developmental Therapeutics Program, DCTD, NCI

Highest concentration tested is 3.8E-04 M; Average TGI over all cell lines is 1.6E-05 M

Panel	Cell Line	log(TGI)	Least Sensitive	Most Sensitive
Leukemia	CCRF-CEM	-3.4		
Leukemia	HL-60(TB)	-3.4		
Leukemia	K-562	-4.7		
Leukemia	SR	-4.5		
Non-Small Cell Lung	A549/ATCC	-4.2		
Non-Small Cell Lung	EK VX	-4.0		
Non-Small Cell Lung	HOP-18	-4.0		
Non-Small Cell Lung	HOP-62	-4.4		
Non-Small Cell Lung	HOP-92	-4.9		
Non-Small Cell Lung	NCI-H226	-4.1		
Non-Small Cell Lung	NCI-H23	-5.1		
Non-Small Cell Lung	NCI-H322M	-4.2		
Non-Small Cell Lung	NCI-H460	-5.6		
Non-Small Cell Lung	NCI-H522	-4.8		
Non-Small Cell Lung	LXFL 529	-5.1		
Small Cell Lung	DMS 114	-5.5		
Colon	COLO 205	-5.2		
Colon	DLD-1	-5.1		
Colon	HCC-2998	-4.7		
Colon	HCT-116	-5.3		
Colon	HT29	-5.3		
Colon	KM12	-5.2		
Colon	KM20L2	-5.1		
Colon	SW-620	-5.2		
Central Nervous System	SF-268	-3.4		
Central Nervous System	SF-295	-4.4		
Central Nervous System	SF-539	-5.0		
Central Nervous System	SNB-19	-4.1		
Central Nervous System	SNB-75	-5.4		
Central Nervous System	SNB-78	-4.9		
Central Nervous System	U251	-5.2		
Central Nervous System	XF 498	-5.4		
Melanoma	LOX IMVI	-5.2		
Melanoma	MALME-3M	-5.0		

Melanoma	M14	-5.0	█
Melanoma	M19-MEL	-5.1	██
Melanoma	SK-MEL-28	-5.0	█
Melanoma	SK-MEL-5	-5.2	██
Melanoma	UACC-257	-5.0	█
Melanoma	UACC-62	-5.3	███
Ovarian	IGROV1	-4.8	
Ovarian	OVCAR-3	-5.1	██
Ovarian	OVCAR-4	-4.9	█
Ovarian	SK-OV-3	-4.6	██
Renal	786-0	-4.9	█
Renal	A498	-4.7	█
Renal	ACHN	-4.9	█
Renal	CAKI-1	-4.2	████
Renal	RXF 393	-5.1	███
Renal	RXF-631	-5.0	█
Renal	SN12C	-5.0	██
Renal	TK-10	-4.6	██
Renal	UO-31	-4.8	



# TGI Mean Graph for Compound S406034

## Results from the National Cancer Institute Human Tumor Cell Line Screen

These results provided by the Developmental Therapeutics Program, DCTD, NCI

Highest concentration tested is 1.0E-04 M; Average TGI over all cell lines is 3.3E-05 M

Panel	Cell Line	log(TGI)	Least Sensitive	Most Sensitive
Leukemia	CCRF-CEM	-4.5		
Leukemia	HL-60(TB)	-5.4		
Leukemia	K-562	-4.0		
Leukemia	MOLT-4	-4.2		
Leukemia	RPMI-8226	-4.4		
Leukemia	SR	-4.0		
Non-Small Cell Lung	A549/ATCC	-4.0		
Non-Small Cell Lung	EKVX	-4.0		
Non-Small Cell Lung	HOP-62	-4.6		
Non-Small Cell Lung	HOP-92	-4.0		
Non-Small Cell Lung	NCI-H226	-5.0		
Non-Small Cell Lung	NCI-H23	-4.9		
Non-Small Cell Lung	NCI-H322M	-4.0		
Non-Small Cell Lung	NCI-H460	-4.0		
Non-Small Cell Lung	NCI-H522	-4.8		
Colon	COLO 205	-5.3		
Colon	HCC-2998	-5.3		
Colon	HCT-116	-4.3		
Colon	HCT-15	-5.2		
Colon	HT29	-5.3		
Colon	KM12	-4.9		
Colon	SW-620	-4.0		
Central Nervous System	SF-268	-4.6		
Central Nervous System	SF-295	-5.0		
Central Nervous System	SF-539	-5.3		
Central Nervous System	SNB-19	-4.6		
Central Nervous System	SNB-75	-5.1		
Central Nervous System	U251	-4.6		
Melanoma	LOX IMVI	-4.3		
Melanoma	MALME-3M	-4.0		
Melanoma	M14	-4.5		
Melanoma	SK-MEL-2	-4.6		
Melanoma	SK-MEL-28	-4.0		
Melanoma	SK-MEL-5	-4.4		
Melanoma	UACC-257	-4.0		

# CHELIDONINE HYDROHLORYDE

Melanoma	UACC-62	-4.6	■
Ovarian	IGROV1	-4.0	■
Ovarian	OVCAR-3	-4.7	■
Ovarian	OVCAR-4	-4.0	■
Ovarian	OVCAR-5	-4.0	■
Ovarian	OVCAR-8	-4.3	■
Ovarian	SK-OV-3	-4.5	■
Renal	786-0	-4.1	■
Renal	A498	-5.2	■
Renal	ACHN	-4.0	■
Renal	CAKI-1	-4.0	■
Renal	RXF 393	-4.6	■
Renal	SN12C	-4.4	■
Renal	TK-10	-4.0	■
Renal	UO-31	-4.0	■
Prostate	PC-3	-4.0	■
Prostate	DU-145	-4.5	■
Breast	MCF7	-4.3	■
Breast	NCI/ADR-RES	-4.7	■
Breast	MDA-MB-231/ATCC	-4.6	■
Breast	HS 578T	-4.0	■
Breast	MDA-MB-435	-5.6	■
Breast	MDA-N	-5.7	■
Breast	BT-549	-4.0	■
Breast	T-47D	-4.0	■

# TGI Mean Graph for Compound S6396

THIO – TEPA

## Results from the National Cancer Institute Human Tumor Cell Line Screen

These results provided by the [Developmental Therapeutics Program, DCTD, NCI](#)

Highest concentration tested is 1.0E-04 M; Average TGI over all cell lines is 4.3E-05 M

Panel	Cell Line	log(TGI)	Least Sensitive	Most Sensitive
Leukemia	CCRF-CEM	-4.1		■
Leukemia	HL-60(TB)	-4.4		■
Leukemia	K-562	-4.1		■
Leukemia	MOLT-4	-4.3		■
Leukemia	RPMI-8226	-4.0		■
Leukemia	SR	-4.5		■
Non-Small Cell Lung	A549/ATCC	-4.3		■
Non-Small Cell Lung	EKVX	-4.4		■
Non-Small Cell Lung	HOP-18	-4.0		■
Non-Small Cell Lung	HOP-62	-4.7		■
Non-Small Cell Lung	HOP-92	-4.3		■
Non-Small Cell Lung	NCI-H226	-4.4		■
Non-Small Cell Lung	NCI-H23	-4.6		■
Non-Small Cell Lung	NCI-H322M	-4.1		■
Non-Small Cell Lung	NCI-H460	-4.5		■
Non-Small Cell Lung	NCI-H522	-4.3		■
Non-Small Cell Lung	LXFL 529	-4.4		■
Non-Small Cell Lung	MLI-019	-6.4		■
Small Cell Lung	DMS 114	-4.4		■
Small Cell Lung	DMS 273	-4.4		■
Colon	COLO 205	-4.4		■
Colon	DLD-1	-4.1		■
Colon	HCC-2998	-4.4		■
Colon	HCT-116	-4.3		■
Colon	HCT-15	-4.1		■
Colon	HT29	-4.2		■
Colon	KM12	-4.2		■
Colon	KM20L2	-4.2		■
Colon	SW-620	-4.2		■
Central Nervous System	SF-268	-4.2		■
Central Nervous System	SF-295	-4.3		■
Central Nervous System	SF-539	-4.3		■
Central Nervous System	SNB-19	-4.1		■
Central Nervous System	SNB-75	-4.2		■
Central Nervous System	SNB-78	-4.0		■

Central Nervous System	U251	-4.4	
Central Nervous System	XF 498	-4.4	■
Central Nervous System	SF-767	-4.8	■
Central Nervous System	SF-763	-4.6	■
Melanoma	LOX IMVI	-4.4	
Melanoma	MALME-3M	-4.5	■
Melanoma	M14	-4.5	■
Melanoma	RPMI-7951	-4.5	■
Melanoma	M19-MEL	-4.2	■
Melanoma	SK-MEL-2	-4.3	
Melanoma	SK-MEL-28	-4.1	■
Melanoma	SK-MEL-5	-4.4	
Melanoma	UACC-257	-4.2	■
Melanoma	UACC-62	-4.6	■
Ovarian	IGROV1	-4.1	■
Ovarian	OVCAR-3	-4.3	
Ovarian	OVCAR-4	-4.4	
Ovarian	OVCAR-5	-4.3	■
Ovarian	OVCAR-8	-4.2	■
Ovarian	SK-OV-3	-4.3	
Renal	786-0	-4.5	■
Renal	A498	-4.4	
Renal	ACHN	-4.4	
Renal	CAKI-1	-4.1	■
Renal	RXF 393	-4.4	
Renal	RXF-631	-4.3	■
Renal	SN12C	-4.1	■
Renal	TK-10	-4.0	■
Renal	UO-31	-4.4	
Renal	UOK-57	-4.7	■
Renal	UOK-57LN	-5.0	■



## Contents

- [Screening Procedures](#)
- [Special Concentration Parameters](#)
- [The Mean Graph](#)
- [The COMPARE Algorithm](#)
- [Building the COMPARE Database](#)
- [Performing a COMPARE Analysis](#)
- [Average Difference Method](#)
- [Correlation Coefficient Method](#)

## Screening Procedures

The NCI screening procedures were described (9) as were the origins and processing of the cell lines (9, 10, 11, 12). Briefly, cell suspensions that were diluted according to the particular cell type and the expected target cell density (5000-40,000 cells per well based on cell growth characteristics) were added by pipet (100  $\mu$  L) into 96-well microtiter plates. Inoculates were allowed a preincubation period of 24 h at 37° C for stabilization. Dilutions at twice the intended test concentration were added at time zero in 100- $\mu$  L aliquots to the microtiter plate wells. Usually, test compounds were evaluated at five 10-fold dilutions. In routine testing, the highest well concentration is 10E-4 M, but for the standard agents the highest well concentration used depended on the agent. Incubations lasted for 48 h in 5% CO<sub>2</sub> atmosphere and 100% humidity. The cells were assayed by using the sulforhodamine B assay (13, 14). A plate reader was used to read the optical densities, and a microcomputer processed the optical densities into the special concentration parameters defined later.

## Special Concentration Parameters GI50, TGI, and LC50

The NCI renamed the IC<sub>50</sub> value, the concentration that causes 50% growth inhibition, the GI<sub>50</sub> value to

emphasize the correction for the cell count at time zero; thus, GI50 is the concentration of test drug where  $100 \times (T - T_0)/(C - T_0) = 50$  (3, 9). The optical density of the test well after a 48-h period of exposure to test drug is T, the optical density at time zero is T<sub>0</sub>, and the control optical density is C. The "50" is called the GI50PRCNT, a T/C-like parameter that can have values from +100 to -100. The GI50 measures the growth inhibitory power of the test agent. The TGI is the concentration of test drug where  $100 \times (T - T_0)/(C - T_0) = 0$ . Thus, the TGI signifies a cytostatic effect. The LC50, which signifies a cytotoxic effect, is the concentration of drug where  $100 \times (T - T_0)/T_0 = -50$ . The control optical density is not used in the calculation of LC50.

These concentration parameters are interpolated values. One uses the concentrations giving GI50PRCNT values above and below the reference values (e.g., 50 for GI50) to make interpolations on the concentration axis. Currently, about 45% of the GI50 records in the database are "approximated". In 42% of the records, the GI50PRCNT for a given cell line does not go to 50 or below. For mean graph (see the discussion later) and COMPARE purposes, the value assumed for the GI50 in such a case is the highest concentration tested (HICONC). Similar approximations are made when the GI50 cannot be calculated because the GI50PRCNT does not go as high as 50 or above (3% of total). In this case, the lowest concentration tested is used for the GI50. Corresponding approximations are made for the TGI and for the LC50.

We use these "approximated" GI50 (TGI and LC50) values in the mean graph and in COMPARE because they represent valued information even though the information is less exact than the measured values would be if the measured values were available. In an extreme case where a compound is essentially inert and the GI50s are all represented by the HICONC approximation, the mean graph becomes a flat vertical line (the mean line) and COMPARE has no pattern to correlate. The opposite extreme case is where a compound is so potent that the lowest concentration tested is used to approximate all of the GI50s. In this case, the mean graph is also a flat vertical line and COMPARE has nothing to correlate. The difference in the two extreme cases is in the retests that are done. The inert compound would not be retested. The

potent compound would be retested at a more appropriate concentration range.

Between the two extremes are examples with few or many approximated GI50s. These can give good results in COMPARE, but the presence of the approximated GI50 requires an additional strategy in the database preparation. The strategy is to treat the data for a given compound in groups defined by the range of concentrations used in the experiment. These ranges are conveniently labeled according to the HICONC. Thus, if multiple tests of a compound are present in the database, only those experiments with the same HICONC are averaged.

This strategy results in compounds having more than one entry in the database. There are differences in the "approximated" GI50 content of the averaged data, and the averages are calculated from different experiments. Therefore, one should expect that the COMPARE-generated correlation coefficients may be different for the same compound tested at different HICONCs. Moreover, at run time the COMPARE user has the option to choose any one of the HICONC sets for the probe pattern averaging, or the user may choose to average all seed data regardless of the HICONC. The consequences of these options and strategies will be apparent later in the examples provided under Applications of COMPARE. For instance, the probe may find itself several times in the COMPARE list at less than 1.00 correlation coefficient.

## The Mean Graph

The discussion of COMPARE presented in this chapter requires an understanding of the mean graph, a means of presenting the *in vitro* test results developed by the staff of DTP (2, 4, 5) to emphasize differential effects of test compounds on various human tumor cell lines.

A "mean graph" is a pattern created by plotting positive and negative values generated from a set of GI50, TGI, or LC50 values. The positive and negative values are plotted along a vertical line that represents the mean response of all the cell lines in the panel to the test agent. Positive values project to the right of the

vertical line and represent cellular sensitivities to the test agent that exceed the mean. Negative values project to the left and represent cell line sensitivities to the test agent that are less than the average value.

The positive and negative values, called deltas, are generated from the GI50 data (or TGI or LC50 data) by a three-step calculation. The GI50 value for each cell line tested against a test compound is converted to its log10 GI50 value. These log10 GI50 values are averaged. Each log10 GI50 value is subtracted from the average to create the delta. Thus, a bar projecting 3 units to the right denotes that the GI50 (or TGI or LC50) for that cell line occurs at a concentration 1000 times less than the average concentration required for all the cell lines used in the experiment.

The complete presentation and organization of the mean graph data were intended to optimize subpanel specific effects, for the listing of cell lines is by disease type, but this presentation of the data is only incidental to the COMPARE concept.

Experience with a wide variety of test compounds has led to the conclusion that a presentation of the mean graphs at all three special concentrations, the GI50, TGI, and LC50, is most useful [\(3\)](#).

## The COMPARE Algorithm

COMPARE analyses are rank-ordered lists of compounds. Every compound from one of several specially prepared databases is ranked for similarity of its in vitro cell growth pattern to the in vitro cell growth pattern of a selected seed or probe compound. To derive COMPARE rankings, a scalar index of similarity between the seed compound cell growth pattern and the pattern for each of the COMPARE database compounds must be created. Two indexes of similarity, the average difference between deltas and the correlation coefficient, are described later, but others are possible [\(3\)](#). The average difference method (ADM) was developed and reported [\(4\)](#) before the correlation coefficient method (CCM), and the database preparation



procedure for the ADM gives a database usable by either method. Therefore, the databases described later are those required by the ADM. The CCM can use these databases, but the mathematical characteristics of this method make the computation of deltas unnecessary.

Individual cell growth patterns can be represented by delta values. The number that is depicted in the mean graph as the length and direction of a bar on the graph is called delta. The ADM utilizes the same description of in vitro data as the mean graph (for important details, see the [discussion of "approximated" GI50, TGI, and LC50](#)).

## Building the COMPARE Database.

To facilitate routine use of COMPARE, several types of COMPARE databases are precalculated with values and stored in SAS data sets (SAS Institute, Inc.). These COMPARE databases are automatically updated (usually each week) as additional compounds are tested. The following detailed description illustrates the processes for collecting data for an ADM database.

The main Oracle database table that contains the GI50 data of the in vitro anticancer screen was searched by Structured Query language (SQL) (the search language for Oracle and many other relational database management systems, SQL). The resulting file was output as an SAS data set ready for further processing by SAS programs. Each test was represented in the SAS data set by six variables: NSC, Panelnbr, Cellnbr, HICONC, GI50, and Discreet status. The NSC identifies which compound was tested. The Panelnbr and Cellnbr together describe the cell line used in the test. HICONC defines the highest dose used in the five-dose assay. The Discrete status denotes whether the compound is confidential or not. The log10 of the GI50 was taken. The database was then sorted by NSC, HICONC, and cell line (Panelnbr and Cellnbr), and the log10 GI50 values were averaged for the same NSC, HICONC, PANELNBR, and CELLNBR. The averaged log10GI50 were named M\_GI50. For preparation of a database suitable for the CCM (but not the ADM), it is only necessary to sort the database at this stage by cell line (e.g., by Panelnbr and Cellnbr). The database would then be ready for use.

For preparation of a database suitable for both the ADM and the CCM, deltas have to be computed. For each NSC at a particular HICONC, the means of the M\_GI50 values were calculated and named MeanGI50. By subtracting the M\_GI50 from the corresponding MeanGI50, the deltas were calculated. As a quality-control measure, sets of cell line deltas for a compound where there were fewer than 35 deltas were excluded. The number 35 was arbitrary, but experience suggested that sets with too few delta values sometimes gave spurious COMPARE matches. It would certainly be better to determine the cutoff value statistically, but this has not been done. Finally, for preparation of this data set for the eventual merging with data sets derived from seed data, this data set was sorted by cell line order (Panelnbr and Cellnbr). The database was ready for use.

## Performing a COMPARE Analysis

To run COMPARE analyses, one must first select from a menu a COMPARE program appropriate to the analysis desired. The selection determines which database will be analyzed. The current options are the standard agent database with 171 compounds, the synthetic compound database with >40,000 compounds and growing (this includes synthetic compounds and natural products of known structure), the natural product crude extract database with >20,000 screened extracts and growing, and other special-purpose databases.

The selection also determines which method of comparison will be used (CCM or ADM), what type of seed may be entered (standard agent or any screened compound or extract), and the level of the analysis (GI50, TGI, LC50, or all three simultaneously). The analyst then enters the NSC number of the desired seed compound and decides if all, some, or just one of the experiments performed on the seed should be averaged. Once the choices are made, the analyst executes the program. The execution begins with the collection, at the time of analysis, of the seed data containing GI50 (or TGI or LC50 or all three) values

directly from the master Oracle database. The seed data collected from the master Oracle database is converted to an SAS data set. If more than one experiment was collected for the seed, the data are then averaged as described previously for building the database. For the ADM, the log10GI50 values from the seed must be converted to a set of delta values just as deltas were calculated for the COMPARE database.

In the next step, pairs of delta values are created (by using an SAS MERGE data step). Each pair consists of the delta value from the seed for a particular cell line and the delta value of a database compound for the same cell line. For example, the delta value calculated for HL-60 data from the seed is paired with the HL-60 delta value calculated for each database compound. For inclusion in the similarity index calculation, both methods (ADM and CCM) require that the delta value (or optionally the M\_GI50 for the CCM) be present in both the seed and the database compound. Thus, if HL-60 was not successfully tested against the seed, no use would be made of any HL-60 data present in the database compounds. If HL-60 data are available for the seed, they will be used only in those cases where HL-60 data are also present in the database compounds. Thus, the seed data determine the maximum number of pairs of delta values that will be used to calculate the index of similarity of each database compound. Variations in the number of cell lines tested against individual database compounds will determine if the number of pairs in the seed data or some lesser number of pairs is used for the particular similarity index calculations.

## **Average Difference Method**

The first step in calculating this index of similarity is to take the difference between paired deltas. An average of these differences, by compound, is computed for each compound. The compounds are sorted by their average difference. The compound with the smallest average difference is the most similar to the seed compound.

## **Correlation Coefficient Method**

A pairwise correlation coefficient ( PCC ) with the seed is calculated for each compound in the database. Those compounds with the highest correlation coefficient are most similar to the seed.

We use a commercial statistical package procedure (the SAS procedure PROC CORR) to obtain PCCs. The PCC provides an excellent index of similarity as judged by the many successful examples provided under Standard Agents and the Standard Agent Database, but other types of correlation coefficients could potentially work as well.

Analysts are usually interested in finding only those compounds in the database that are most similar to the seed. Thus, the list is truncated to 100 of the most similar compounds. Common chemical names, if they are on file, are added automatically to the truncated answer list.

Surprisingly similar, but totally independent, work ([97](#), [98](#)) was published essentially concurrently with the early publications of mean graph and COMPARE in an entirely different area of application.

[Home](#)[Discovery](#)[Development](#)[Pathways](#)[Site Search](#)[Data Search](#)