A Rational Approach to Personalized Anticancer Therapy: Chemoinformatic Analysis Reveals Mechanistic Gene-Drug Associations

Kerby Shedden,[1] Leroy B. Townsend,[2] John C. Drach,[2,3] and Gustavo R. Rosania[4,5]

Received December 11, 2002; accepted February 28, 2003

Purpose. To predict the response of cells to chemotherapeutic agents based on gene expression profiles, we performed a chemoinformatic study of a set of standard anticancer agents assayed for activity against a panel of 60 human tumor-derived cell lines from the Developmental Therapeutics Program (DTP) at the National Cancer Institute (NCI).

Methods. Mechanistically-relevant gene:drug activity associations were identified in the scientific literature. The correlations between expression levels of drug target genes and the activity of the drugs against the NCI's 60 cell line panel were calculated across and within each tumor tissue type, using published drug activity and gene expression data.

Results. Compared to other mechanistically-relevant gene-drug associations, that of triciribine phosphate (TCN-P) and adenosine kinase (ADK) was exceptionally strong—overall and within tumor tissue types—across the 60 cell lines profiled for chemosensitivity (1) and gene expression (2).

Conclusion. The results suggest ADK expression may be useful for stratifying TCN-P-responsive vs. non-responsive tumors. Based on TCN-P's mechanism of action and the observed TCN-P:ADK association, we contend that catalytic drug activation provides a rational, mechanistic basis for personalizing cancer treatment based on tumor-specific differences in the expression of drug-activating enzymes.

KEY WORDS: triciribine phosphate; adenosine kinase; prodrug; chemoinformatics; microarray; transcriptional profiling; cancer.

INTRODUCTION

To personalize chemotherapy, anticancer drugs should be tailored to the phenotypic characteristics of cancer cells. To study gene:drug activity relationships, the DTP program has established a database of anticancer agents that have been screened against a panel of 60 human tumor-derived cell lines (3–5). To explore gene:drug associations useful for predicting chemosensitivity (6), genome-wide mRNA expression data derived from these 60 cell lines (their gene expression profiles) have been related to the sensitivity of these cells to a battery of anticancer agents (their chemosensitivity profiles) (2,7,8). Correlation analysis yields gene:drug associations that may be useful for predicting the responsiveness of cancer cells to anticancer agents.

Previous work (2) found statistically strong predictive relationships between gene expression and chemosensitivity for many anticancer agents; however, inspection of the prediction relationship does not provide a clear connection to drug mechanism of action. There may be many reasons for this. First, intracellular biochemical pathways are intrinsically complex and the relationship between drug target inhibition and cell growth inhibition is complex. In addition, the relationship between gene transcription, protein levels, and protein activity may not be directly proportional. Moreover, many anticancer drugs preferentially target specific tumor tissue types (e.g., leukemia vs. solid tumors), such that coincidental associations between drug activity and tissue-specific differentiation markers may prevail.

Using public datasets derived from the NCI's 60 cell line panel, we report on correlations between gene expression and chemosensitivity for gene:drug pairs that are known to have a mechanistic relationship. We considered correlations across the 60 cell line panel, and within the tumor tissue types represented in the panel. Out of all possible associations between anticancer drugs and known targets, that of triciribine phosphate (TCN-P) and adenosine kinase (ADK) was exceptionally strong across and within tumor tissues. This result is consistent with the hypothesis that catalytic drug activation may provide a rational mechanistic approach to the development of personalized medicinal chemistry, and suggests that TCN-P should be reconsidered as a drug for personalized anticancer therapy.

MATERIALS AND METHODS

Data Sources and Pre-Processing

DNA oligonucleotide measurements derived from the NCI 60 cell lines covering nine tumor tissues (breast, central nervous system, colon, leukemia, melanoma, non-small cell lung cancer, ovary, renal, and prostate) were obtained from http://www.genome.wi.mit.edu/MPR/NC160/NC160.html. The data were originally analyzed in (2), where the experimental protocols are described. Measurements were scaled so that each array had mean expression equal to 1500 U, and the transform $\log_2$ (50 + max (X + 50), 0)) was applied.

Measurements of GI50 (the drug concentration decreasing cell growth by 50%) on the standard anticancer agent database were obtained from the DTP program website at the NCI (http://dtp.nci.nih.gov/docs/cancer/searches/standard_agent_table.html). GI50 values were $\log_2$ transformed, and missing values were imputed with the mean over all cell lines for a given drug.

[1] Department of Statistics, The University of Michigan, 428 Church Street, Ann Arbor, Michigan 48109.
[2] Department of Medicinal Chemistry, College of Pharmacy; The University of Michigan, 428 Church Street, Ann Arbor, Michigan 48109.
[3] Department of Biologic and Materials Science, School of Dentistry; The University of Michigan, 428 Church Street, Ann Arbor, Michigan 48109.
[4] Department of Pharmaceutical Sciences, College of Pharmacy; The University of Michigan, 428 Church Street, Ann Arbor, Michigan 48109.
[5] To whom correspondence should be addressed. (e-mail: grosania@umich.edu)

ABBREVIATIONS: ADK, adenosine kinase; TCN-P, triciribine phosphate; NCI, National Cancer Institute; DTP, Developmental Therapeutics Program.

Identification of Genes Relevant to Drug Mechanism of Action

To identify genes relevant to drug mechanism of action, each drug was used as key word for searching abstracts of articles in PubMed. In addition, published articles that have comprehensively explored the relationship between drug mechanism of action and chemosensitivity profiles of the standard anticancer agent database were used as references (2,4–7). Genes and proteins identified as playing a primary role in drug mechanism of action were matched to specific genes on the DNA oligonucleotide array, based on gene name. For example, doxorubicin is a topoisomerase II inhibitor, so expression of the topoisomerase II genes was used when analyzing doxorubicin. Drugs with unknown mechanism of action—or mechanism of action that does not directly involve a protein (such as DNA alkylators)—were not considered.

Statistical Analysis

The Pearson correlation coefficient was used to measure the relationships between the log-transformed gene expression and GI50 data, across the 60 cell lines (the "overall correlation"). To isolate within-tumor type associations, correlation coefficients were computed separately for each tissue, then absolute values of these coefficients were averaged using weights proportional to the number of samples in a given tissue.

To assess the significance of the correlation coefficients, the gene expression levels were randomly permuted 1000 independent times, over the 60 cell lines. For each permutation, the correlation with the GI50 levels was recomputed. The proportion of these correlations that are larger in magnitude than the actual correlation corresponds to the p value for that particular gene:drug pair. The distribution of the largest permuted correlation among all 364 gene:drug pairs is used to produce a p value for the largest observed gene:drug correlation.

Results

The concentration at which different anticancer drugs inhibit cell growth by 50% (the drug's GI50 concentration) was graphed relative to the expression level of genes involved in drug mechanism of action. By visual inspection, it is evident that cell growth inhibition by TCN-P parallels ADK gene expression (Fig. 1). The TCN-P:ADK correlation—overall and within-tissues—was greater than that of the other gene:drug pairs examined (Supplementary table). Beyond a simple correlation, the trend covered a substantial magnitude of variation, in that there is at least an 8-fold spread between the least and greatest value in ADK gene expression and TCN-P sensitivity. The correlations within tumor tissue types are also strong and consistently negative for both copies of ADK, suggesting that the TCN-P:ADK interaction is specific and occurs in all tissue types studied.

TCN-P is a water-soluble and plasma membrane-impermeant, prodrug derivative of triciribine (9). In serum, TCN-P is de-phosphorylated extracellularly to triciribine (10), which is then able to traverse the plasma membrane and inhibit DNA synthesis (11). Nevertheless, TCN-P and triciribine are inactive unless phosphorylated by intracellular ADK, at the 5′ position (9) (Fig. 2). Intracellular phosphorylation by ADK is necessary for activation of TCN-P and its parent compound triciribine. The ADK inhibitor 5-iodotubercidin decreases the activity of either compound by 200-fold (9). Also, an ADK-deficient leukemic cell subline is 5000 fold less sensitive to the drugs than the parent cell line (9,12). Moreover, 5′-deoxy derivatives of triciribine are inactive because they cannot be phosphorylated (13). In HIV-growth inhibi-

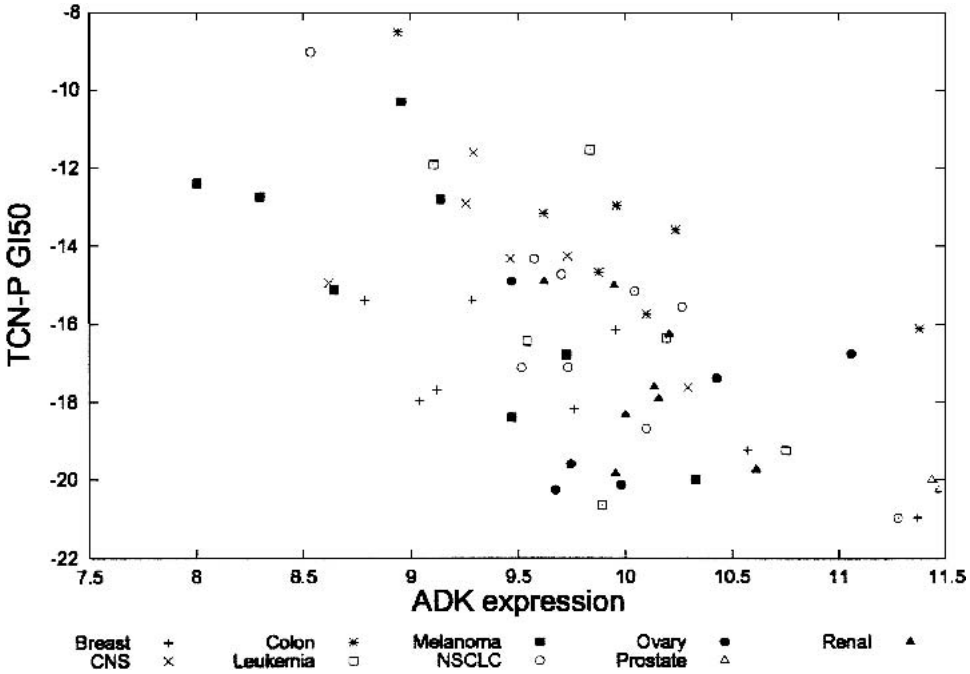

Fig. 1. Plot of $\log_2$ transform of TCN-P activity (based on the GI50—the drug concentration inhibiting cell growth by 50%) and adenosine kinase mRNA levels, for each cell line in the NCI's 60 cell line panel.

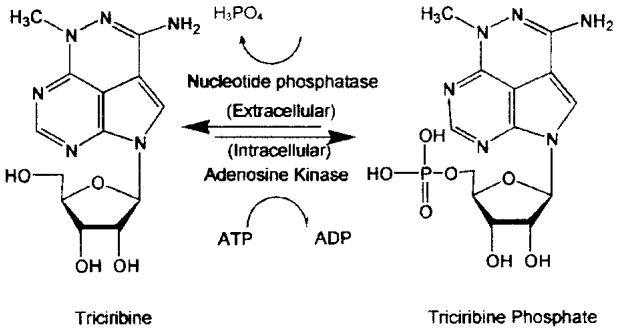

Fig. 2. Triciribine phosphate (TCN-P) is de-phosphorylated extracellularly, traverses the plasma membrane, and is activated by intracellular re-phosphorylation catalyzed by ADK.

tion assays, phosphorylation of TCN-P by ADK is essential for antiviral activity (12,13). In fact, these and other related experiments establishing the activation of TCN-P by ADK led to its inclusion in the NCI's standard anticancer agent database.

According to Michaelis-Menten kinetics, the rate of drug activation—and consequently, growth inhibitory activity—should be proportional to the relative amounts of activating enzyme concentration, for a set intracellular substrate concentration. To better analyze the within-tissue type association, the correlation between ADK expression and TCN-P cell growth inhibitory activity was calculated independently for each tumor tissue type, and compared to that of other gene:drug associations. Based on this analysis, ADK expression exhibits a strong negative correlation with the concentration of TCN-P needed to inhibit cell growth (Fig. 3). This finding indicates that ADK expression can be predictive of chemosensitivity across and within tumor tissues, as expected from a true mechanistic relationship between ADK expression and TCN-P activation.

By calculating the average within tumor correlation across the eight different tumor tissues analyzed and plotting it against the overall correlation, the strength of TCN-P:ADK association can also be compared to the other gene:drug pairs examined. This analysis underscores the strength of the TCN-P:ADK correlation, relative to those of other clinically-relevant drugs with known intracellular targets (Fig. 4). Overall, the within-tissue correlations are slightly stronger than the across tumor tissue correlations. This suggests that the relationship is not due to the confounding effects of tissue specific gene expression or tissue specific chemosensitivity. Moreover, the within tissue correlations for TCN-P:ADK are consistently negative (17/18 of the correlations are negative in the two replications). Both of these observations are consistent with the hypothesis that enzymatic activation of TCN-P by ADK in cancer cells is a common and specific phenomenon, across and within tissue types.

In addition to TCN-P:ADK, there are two other outliers in the overall vs. within-tumor correlation plot (Fig. 4). One of them (0.63, 0.05) is for ara-6-mp/RNA polymerase II. This is just one of 14 probes of RNA polymerase II on the array, however, and the other 13 do not show correlation. The other outlier (0.31, 0.5) is for rhizoxin/beta-tubulin, but again there are many beta tubulin probes on the array and only one shows correlation. Although these apparent associations may be experimental noise, it is possible that the expression of a specific tubulin or polymerase II isoform may determine the activity

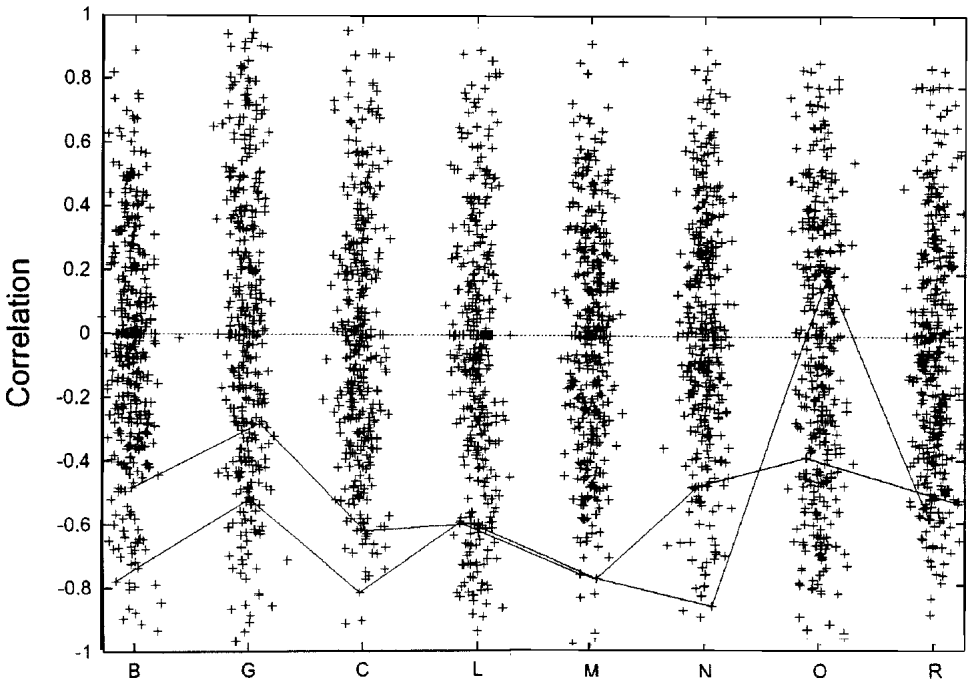

Fig. 3. Within-tumor Pearson correlations between the NCI's standard anticancer drug set and genes associated with drug mechanism of action. A total of 364 drug/gene pairs were considered, based on experimental evidence for mechanistic drug-gene association found in the scientific literature. The solid lines connect the correlations between TCN-P and two probes for ADK. The tumor types are: B: breast; G: lung; C: CNS; L: leukemia; M: melanoma; N: colon; O: ovarian; R: renal. Prostate was excluded from this analysis, as there are only two cell lines in the prostate panel.

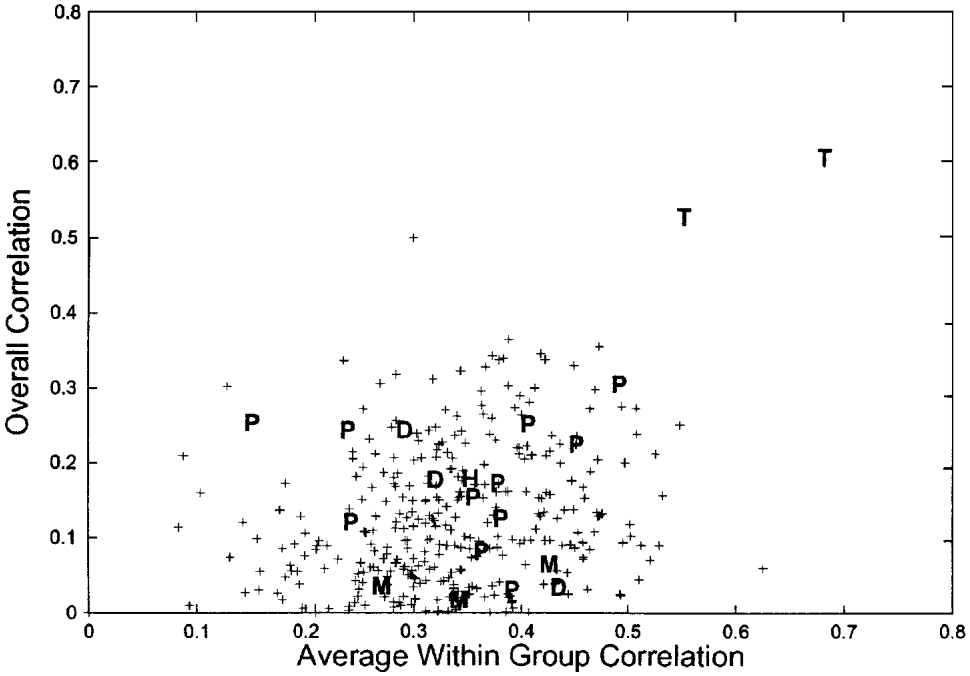

Fig. 4. Bivariate plots of mechanistically relevant gene-drug activity correlations, across and within tumor tissue types in the 60-cell line panel. (T: TCN-P correlated with two ADK probes; P: paclitaxel correlated with multiple α- and β-tubulin probes; D: doxorubicin correlated with multiple topoisomerase II probes; H: hydroxyurea correlated with a ribonucleotide reductase probe; M: methotrexate correlated with multiple dihydrofolate reductase probes. The "+" symbols indicate identified correlations for the other drugs in the database.

of rhizoxin or ara-6-mp respectively, and we are exploring the nature of this association in greater detail. Nevertheless, the strength of the TCN-P:ADK correlation, occurring in two replications across and within tumor tissues, far surpasses that of all other gene:drug pairs.

To establish whether the TCN-P:ADK association is likely to reproduce, we considered several sources of experiment error that may produce false positives when carrying out correlation analysis in a large dataset. To test whether the observed TCN-P:ADK correlation may be caused by random chance, a statistical simulation was carried out as described in the Methods (results shown in the supplementary table). For both copies of ADK on the array, the overall correlation is significant ($p < 0.001$). In addition, we considered whether the strongest of the 364 associations in a simulated dataset (as opposed to the simulated association for TCN-P:ADK in particular) is stronger than the observed TCN-P:ADK correlation. Again, in this case, the p values are small and significant for both copies of ADK ($p < 0.01$). Within each tumor tissue the TCN-P:ADK association is significant for both copies of ADK ($p < 0.01$ for TCN-P:ADK in particular; $p < 0.05$ for the maximum within group correlation over all 364 pairs studied). Finally, to combine the evidence contained in the overall- and within-tumor correlations, we considered the probability that a gene:drug pair would have stronger within-tumor and overall correlations than the observed values for TCN-P:ADK. Using the statistical simulation analysis described in the Methods, this was found to be unlikely for both copies of ADK ($p < 0.001$ for the TCN-P:ADK association in particular; $p < 0.01$ for all 364 gene:drug pairs simultaneously).

DISCUSSION

In pursuit of personalized anticancer treatments, gene-drug associations directly relevant to drug mechanism of action would be useful for tailoring anticancer therapy to the phenotypic characteristics of cancer cells. We have found that ADK gene expression levels, as determined with oligonucleotide microarrays, are predictive of the response of cells to TCN-P across a panel of 60 cell lines derived from different tumor tissues. This result is consistent with the requirement of intracellular ADK for activating TCN-P. It supports the notion that catalytic drug activation can be used to personalize anticancer therapy, by monitoring the expression level of drug activating enzymes.

As a mechanistic hypothesis, the predictive power behind the TCN-P:ADK association resides in the inactivity of the drug in cells that do not express ADK, together with the enzyme-catalyzed intracellular accumulation of the membrane impermeant, activated drug in cells expressing ADK. As a corollary, we predict that our observations are not unique to the TCN-P:ADK association or to phosphorylation reactions and that other enzymatic activation mechanisms could be exploited for similar purposes. For example, ester hydrolysis yields a cell-impermeable carboxylate product. An inactive, cell-permeable ester prodrug could be rationally designed as a substrate for an esterase, such that accumulation of an active carboxylate drug product should be proportional to esterase expression.

Is the discovered correlation between TCN-P and ADK real? Statistically, we have considered the possibility that the association may be a false positive, because of a random ex-

perimental variation producing a spurious association. We calculate that this is unlikely, with a probability <0.001 that both within-tumor and overall correlations as strong as those observed for TCN-P:ADK could be found by chance. This result was reproduced simultaneously for two independent measurements of ADK expression, making it even less likely to be due to chance variation. Moreover -in the particular case of TCN-P:ADK correlation- chemical, biochemical, and pharmacological experiments support the notion that ADK gene expression is mechanistically and causally related to the level of TCN-P activity (9,11–13).

In conclusion, the catalytic drug activation mechanism exemplified by the ADK:TCN-P association may provide a rational approach to the development of personalized therapies. For all mechanistically-relevant gene:drug associations explored in the NCI standard anticancer agent database, the use of ADK expression for stratifying TCN-P responsive vs. non-responsive tumors presents the most promising gene: drug relationship with potential application to personalized therapy. These results are consistent with the fact that TCN-P is inactive unless activated by adenosine kinase, and that the level of adenosine kinase expression determines the activity of TCN-P against different tumor cells. In preliminary phase I and II clinical trials, TCN-P is tolerated (14,15). Based on the rational, mechanistic hypothesis supported by the results presented here, we propose that TCN-P be reconsidered as an advanced, experimental drug candidate for personalized therapy against ADK-over-expressing tumors.

ACKNOWLEDGMENTS

We acknowledge Dr. D. Smith for helpful suggestions. This work was supported by a grant from University of Michigan Bioinformatics Program; Pfizer Global Research and Development-Ann Arbor; Howard Hughes Medical Institute; and the Upjohn-Vahlteich Research Fund.

REFERENCES

1. M. R. Boyd. The NCI In Vitro Anticancer Drug Discovery Screen. In B. Teicher (ed), *Anticancer Drug Development Guide: Preclinical Screening, Clinical Trials and Approval,* Humana Press, Totowa, New Jersey, 1995, pp. 23–41.
2. J. E. Staunton, D. K. Slonim, H. A. Coller, P. Tamayo, M. J. Angelo, J. Park, U. Scherf, J. K. Lee, W. O. Reinhold, J. N. Weinstein, J. P. Mesirov, E. S. Lander, and T. R. Golub. Chemosensitivity prediction by transcriptional profiling. *Proc. Natl. Acad. Sci. USA* **98**:10787–10792 (2001).
3. R. H. Shoemaker, D. A. Scudiero, G. Melillo, M. J. Currens, A. P. Monks, A. A. Rabow, D. G. Covell, and E. A. Sausville. Application of high-throughput, molecular-targeted screening to anticancer drug discovery. *Curr. Top. Med. Chem.* **2**:229–246 (2002).
4. A. A. Rabow, R. H. Shoemaker, E. A. Sausville, and D. G. Covell. Mining the National Cancer Institute's tumor-screening database: identification of compounds with similar cellular activities. *J. Med. Chem.* **45**:818–840 (2002).
5. O. Keskin, I. Bahar, R. L. Jernigan, J. A. Beutler, R. H. Shoemaker, E. A. Sausville, and D. G. Covell. Characterization of anticancer agents by their growth inhibitory activity and relationships to mechanism of action and structure. *Anticancer Drug Des.* **15**:79–98 (2000).
6. 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, E. A. Sausville, D. W. Zaharevitz, B. Bunow, V. N. Viswanadhan, G. S. Johnson, R. E. Wittes, and K. D. Paull. An information-intensive approach to the molecular pharmacology of cancer. *Science* **275**:343–349 (1997).
7. U. Scherf, D. T. Ross, M. Waltham, L. H. Smith, J. K. Lee, L. Tanabe, K. W. Kohn, W. C. Reinhold, T. G. Myers, D. T. Andrews, D. A. Scudiero, M. B. Eisen, E. A. Sausville, Y. Pommier, D. Botstein, P. O. Brown, and J. N. Weinstein. A gene expression database for the molecular pharmacology of cancer. *Nat. Genet.* **24**:236–244 (2000).
8. D. T. Ross, U. Scherf, M. B. Eisen, C. M. Perou, C. Rees, P. Spellman, V. Iyer, S. S. Jeffrey, M. Van de Rijn, M. Waltham, A. Pergamenschikov, J. C. Lee, D. Lashkari, D. Shalon, T. G. Myers, J. N. Weinstein, D. Botstein, and P. O. Brown. Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* **24**:227–235 (2000).
9. L. L. Wotring, G. W. Crabtree, N. L. Edwards, R. E. Parks Jr., and L. B. Townsend. Mechanism of activation of triciribine phosphate (TCN-P) as a prodrug form of TCN. *Cancer Treat. Rep.* **70**:491–497 (1986).
10. R. B. Schilcher, J. D. Young, and L. H. Baker. Reversed-phase high-performance liquid chromatographic determination of tricyclic nucleoside and tricyclic nucleoside 5′-phosphate in biological specimens. *J. Chromatogr.* **337**:55–62 (1985).
11. L. L. Wotring, J. E. Passiatore, J. L. Roti Roti, J. L. Hudson, and L. B. Townsend. Effects of the tricyclic nucleoside 6-amino-4-methyl-8-(beta-D-ribofuranosyl)- pyrrolo[4,3,2-de]pyrimido[4,5-c]pyridazine on the viability and cell cycle distribution of L1210 cells in vitro. *Cancer Res.* **45**:6355–6361 (1985).
12. R. G. Ptak, K. Z. Borysko, A. R. Porcari, J. L. Buthod, L. E. Holland, C. Shipman Jr., L. B. Townsend, and J. C. Drach. Phosphorylation of triciribine is necessary for activity against HIV type 1. *AIDS Res. Hum. Retroviruses* **14**:1315–1322 (1998).
13. A. R. Porcari, R. G. Ptak, K. Z. Borysko, J. M. Breitenbach, S. Vittori, L. L. Wotring, J. C. Drach, and L. B. Townsend. Deoxy sugar analogues of triciribine: correlation of antiviral and antiproliferative activity with intracellular phosphorylation. *J. Med. Chem.* **43**:2438–2448 (2000).
14. L. G. Feun, N. Savaraj, G. P. Bodey, K. Lu, B. S. Yap, J. A. Ajani, M. A. Burgess, R. S. Benjamin, E. McKelvey, and I. Krakoff. Phase I study of tricyclic nucleoside phosphate using a five-day continuous infusion schedule. *Cancer Res.* **44**:3608–3612 (1984).
15. K. Hoffman, F. A. Holmes, G. Fraschini, L. Esparza, D. Frye, M. N. Raber, R. A. Newman, and G. N. Hortobagyi. Phase I-II study: triciribine (tricyclic nucleoside phosphate) for metastatic breast cancer. *Cancer Chemother. Pharmacol.* **37**:254–258 (1996).