Chemistry & Biology, Vol. 10, 383–396, May, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S1074-5521(03)00095-4

Multidimensional Chemical Genetic Analysis of Diversity-Oriented Synthesis-Derived Deacetylase Inhibitors Using Cell-Based Assays

Stephen J. Haggarty,[1,2,3] Kathryn M. Koeller,[2,3] Jason C. Wong,[2,3] Rebecca A. Butcher,[2,3] and Stuart L. Schreiber[1,2,3,4,*]
[1]Departments of Molecular and Cellular Biology
[2]Departments of Chemistry and Chemical Biology
[3]Harvard Institute of Chemistry and Cell Biology
Initiative for Chemical Genetics
[4]Howard Hughes Medical Institute
Harvard University
12 Oxford Street
Cambridge, Massachusetts 02138

Summary

Systematic chemical genetics aims to explore the space representing interactions between small molecules and biological systems. Beyond measuring binding interactions and enzyme inhibition, measuring changes in the activity of proteins in intact signaling networks is necessary. Toward this end, we are partitioning chemical space into regions with different biological activities using a panel of cell-based assays and small molecule "chemical genetic modifiers." Herein, we report on the use of this methodology for the discovery of 617 small molecule inhibitors of histone deacetylases from a multidimensional screen of an encoded, diversity-oriented synthesis library. Following decoding of chemical tags and resynthesis, we demonstrate the selectivity of one inhibitory molecule (tubacin) toward α-tubulin deacetylation and another (histacin) toward histone deacetylation. These small molecules will facilitate dissecting the role of acetylation in a variety of cell biological processes.

Introduction

Models of complex biological systems highlight the interdependent and robust nature of biochemical networks [1–4]. The development of experimental methods to modulate selectively the functions of individual nodes (representing proteins) in these networks is a central aim of chemical genetics. Although chemical genetics is modeled after classical genetics, especially with respect to the logic of phenotype-based screening, it differs from classical genetics in the use of small molecules, rather than mutations, to perturb the function(s) of gene products [5–10]. Given the temporal control offered by small molecules and the ability to use combinations of small molecule modulators, chemical genetics promises to complement the use of pure genetic analysis to study a wide range of biological systems. Although the use of ribonucleic acid interference (RNAi) and related phenomena now provides a powerful reverse genetic approach for functional genomics, the inability to selectively target individual functions of proteins, to directly disrupt protein-protein interactions, and the extended temporal scale required for RNAi limits the general application of this approach.

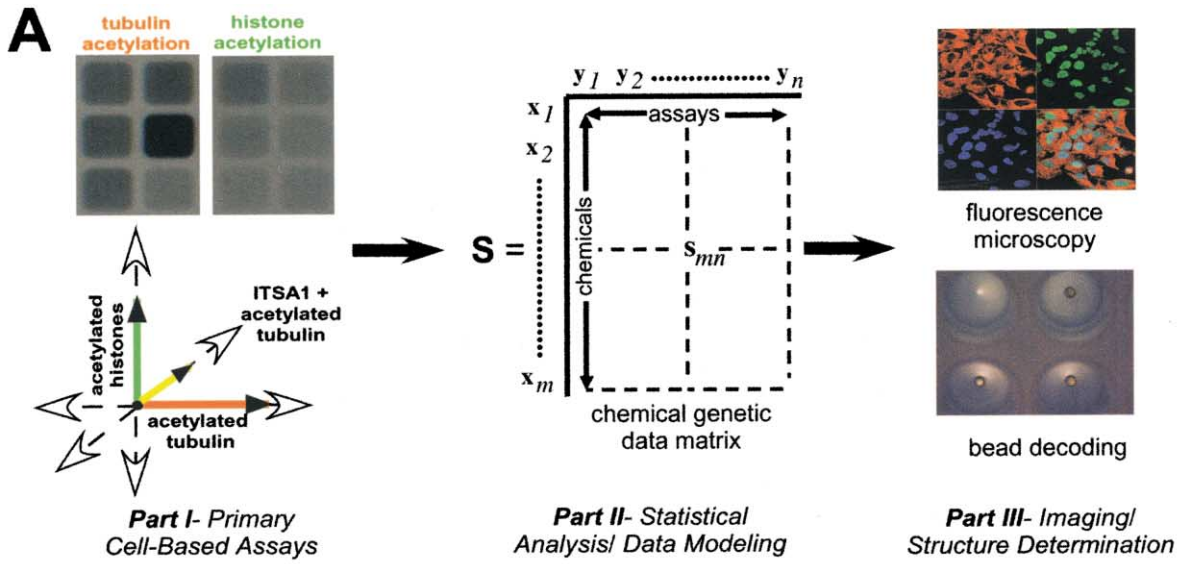
Diversity-oriented organic synthesis is increasingly providing complex and effective small molecule modulators of biological processes [8, 10]. Challenges for chemical genetics include: (1) determining which of these molecules have specific effects upon biological systems (at various levels of resolution from proteins to whole organisms), and (2) determining the structural and physiochemical properties of molecules that specify associated biological activities. These efforts will benefit from the systematic "mapping" of small molecules to positions in a multidimensional chemical space derived from molecular and/or biological descriptors. Toward this end, a multidimensional chemical genetic (MDCG) approach is being developed, using high-throughput phenotypic assays to partition chemical space into regions of different biological activities as well as to partition biological space into regions with different chemical activity.

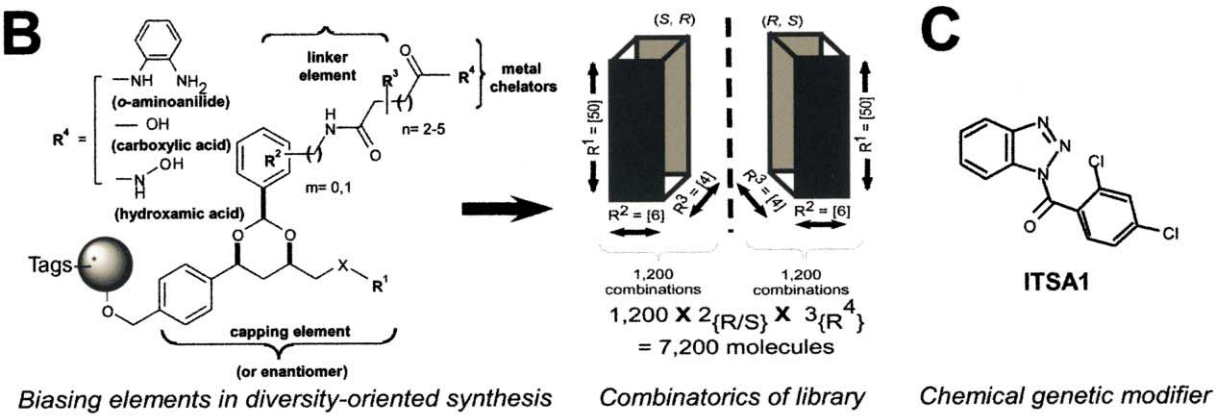
Similar to other research endeavors requiring multidimensional analysis, chemical genetic information can be arranged in a matrix, denoted **S**, consisting of an ordered array of columns and rows (Figure 1A). Each column ($\mathbf{y}_j$) in **S** is a descriptor corresponding to a cell-based or biochemical assay and is denoted by a boldface, lowercase letter subscripted j (where $j = 1$ to n). Similarly, each row ($\mathbf{x}_i$) in **S** corresponds to a chemical and is denoted by a boldface, lowercase letter subscripted i (where $i = 1$ to m). Accordingly, an element (m, n) of **S** encodes information about chemical m for descriptor n.

Geometrically, considering the elements of **S** as coordinates, chemicals (or assays) can be modeled as vectors, $\mathbf{x}_i = [e_1, e_2, \ldots, e_n]$, in an n- (or m-) dimensional vector space. By defining the Euclidean distance D between two vectors (e.g., $\mathbf{x}_1$ and $\mathbf{x}_2$) in this vector space to be $D_{12}^2 = \Sigma\,[(\mathbf{x}_1 - \mathbf{x}_2)^2]$, the space of chemical genetic observation can be considered as a metric space. That means the relative distance D between chemicals $\mathbf{x}_i$ (or assays $\mathbf{y}_j$), in other words their proximity becomes informative with respect to similarity between the particular descriptors considered. Accordingly, in analyzing chemical genetic data obtained from screening small molecule libraries, small molecules $\mathbf{x}_i$ (or assays $\mathbf{y}_j$) can be considered to be *functionally similar* if they are *closely positioned* (i.e., within a specified radius) in the underlying descriptor space. Since similarity between small molecules is determined by the pattern of interaction with biological systems, the corresponding distance metric D complements the definition of similarity obtained from calculated molecular descriptors based on chemical structure. Furthermore, since similarity in cell-based assays results from patterns of small molecules interacting with expressed gene products, the corresponding distance metric D complements the definition of similarity obtained from DNA sequence or gene expression analysis.

As part of this ongoing effort, we describe the devel-

*Correspondence: sls@slsiris.harvard.edu





D

	Assay	Abbreviation	Compounds screened in duplicate
1.	Acetylated tubulin	AcTubulin	7,392
2.	Acetylated tubulin + ITSA1 (chemical genetic modifier)	ITSA1+AcTubulin	2,464 hydroxamic acids
3.	Acetylated lysine	AcLysine	7,392
4.	Acetylated histone H3	AcHistH3	2,464 hydroxamic acids
5.	Acetylated histone H4	AcHistH4	2,464 hydroxamic acids

Table 1.Summary of chemical genetic screens

Figure 1. Overview of Multidimensional Chemical Genetic Analysis

(A) The three-part protocol involved in screening of 7392 diversity-oriented synthesis-derived deacetylase inhibitors. Chemical genetic information is obtained as "object observations" and arranged in a matrix, denoted by **S**. Each column ($\mathbf{y}_j$) in **S** is a descriptor which corresponds to a phenotype from a cell-based or biochemical assay. Each row ($\mathbf{x}_i$) in **S** is an object which corresponds to a chemical. An element (m, n) of **S** encodes information about chemical m for descriptor n.

(B) Summary of small molecules based upon a 1,3-dioxane structure and combinatorics of the full library.

(C) Structure of ITSA1, a chemical genetic modifier that suppresses TSA-induced histone and α-tubulin acetylation [21].

(D) Summary of screens performed and abbreviations used for the assays.

opment of new cytoblot assays that use posttranslational chemical modifications as markers of cellular state [11]. In this system, MDCG analysis has been applied to identify selective inhibitors of the family of zinc-dependent deacetylase enzymes typified by the histone deacetylases (HDACs) [12–20]. In a companion paper

to this one [21], we describe the motivations behind the use of chemical genetic modifiers (e.g., ITSA1, "*i*nhibitor of *t*richostatin *A*-1") in dissecting the role of HDACs in regulating cell cycle progression. In another paper [22], one selective inhibitor of α-tubulin deacetylation (tubacin) discovered in this analysis is used to dissect the function of HDAC6 as an α-tubulin deacetylase [23] in mediating cell cycle progression, microtubule stability, and cell motility. Finally, we have begun dissecting the structural and physiochemical basis for the in vivo selectivity of HDAC inhibitors discovered in this analysis [24].

Results

Library Design and Outline of Multidimensional Chemical Genetic Analysis

The stereoselective synthesis of a small molecule library based on a 1,3-dioxane diversity element has been described previously (Figure 1B) [25]. Each small molecule contains a metal-chelating functional group (*o*-aminoanilide, carboxylic acid, or hydroxamic acid) at position $\mathbf{R^4}$. Since HDACs are zinc-dependent hydrolases [26], metal chelators "bias" the library toward deacetylase inhibition. Prior to screening the full 7,392-member library, a pilot screen of five representative small molecules indicated that these compounds inhibited HDACs 1, 4, and 6 in vitro, and four out of five also induced α-tubulin acetylation (low μM range; data not shown) [27]. Consequently, since determining the relative activity and selectivity of these small molecules was a main objective, we focused on identifying compounds that caused a differential increase in α-tubulin acetylation and histone acetylation as indicators of cytoplasmic and nuclear deacetylase inhibition, respectively.

Based on our previous success using miniaturized cytoblot cell-based assays [9, 11], we developed a three-part screen compatible with the amount of compound available from a "one bead-one stock solution" approach to chemical genetics [29–31]. Using antibodies specific to the acetylated state of α-tubulin and acetylated histones (Figure 1A, part I) [27], assay conditions were optimized separately using HDAC inhibitor trichostatin A (TSA) as a positive control for inhibition of both histone and α-tubulin deacetylation [13, 21]. The antibody utilized to evaluate the level of histone acetylation (anti-acetylated lysine antibody) reacts with a variety of proteins by Western blotting. However, under cytoblot conditions, this antibody recognized predominantly nuclear proteins, as judged by fluorescence microscopy. Thus, the acetylated lysine cytoblot is effectively a measure of histone acetylation, as histones are the predominant nuclear acetylated proteins. The chemical genetic modifier ITSA1 (Figure 1C), which suppresses (reduces) the ability of TSA to increase acetylation levels of both α-tubulin and histones [21], was introduced as another variable in our screening strategy. ITSA1 suppresses hydroxamic acid-based (such as TSA) but not epoxy ketone-based (such as trapoxin) deacetylase inhibitors. Thus, using ITSA1 in our screen allowed us to identify both "TSA-like" and "trapoxin-like" small molecule deacetylase inhibitors, based on the ability to be suppressed by ITSA1.

Following cytoblot analysis, statistical properties were calculated for each of the screening datasets (Figure 1A, part II). Using an empirically determined threshold for bioactivity, we constructed a discrete model of these data in the form of a chemical genetic network and calculated topological properties of the resulting graph. To determine global relationships between assay variables and regions of different bioactivity, both clustering and principal component analysis were used. Subsequently, small molecules of interest were retested from the sample stock solution by fluorescence microscopy (Figure 1A, part III). This secondary assay allowed us to assess the observed selectivity in an independent manner and to determine whether small molecule "hits" had other effects on cellular morphology, viability, and chromatin conformation.

Screening Results

A summary of the cell-based assays performed is shown in Figure 1D. Following robotic pin transfer [29], the 1,3-dioxane-based small molecules (~2–5 μM, depending upon efficiency of synthesis and amount of compound pin transferred) were incubated with cells for a total of 18 hr. The entire collection of 7392 molecules was screened (in duplicate) in the AcTubulin and AcLysine cytoblot assays (e.g., Figure 2A). A subset of the library was then further evaluated in three other acetylation-based assays, including one involving the suppressor ITSA1. Correlation between replicates was strong, with $r = 0.84$ (p value < 0.05) representing the minimum correlation between duplicate plates containing hydroxamic acids (Figure 2B). Values from replicates were standardized to a control from each plate, averaged, and Log_2 transformed to reduce the skewness and kurtosis prior to fitting to a normal distribution. Statistical properties pertaining to each $\mathbf{R^4}$ biasing element in the library are shown in the box plots in Figure 2C. Since diversity position $\mathbf{R^4}$ was encoded spatially as well as by tagging, we were able to determine the relative distribution of bioactive small molecules amongst the three biasing elements. Under the assumption of equal synthetic efficiency and purity, the hydroxamic acids were assessed to be most active and the *o*-aminoanilides least active in both the AcTubulin and AcLysine assays (Figure 2D).

After separately fitting relative acetylation data from the AcTubulin and AcLysine assays to a normal distribution, a value corresponding to a 1.5-fold or greater increase was chosen as the criterion for bioactivity (see Supplemental Table S1 at http://www.chembiol.com/cgi/content/full/10/5/383/DC1 or write to chembiol@cell.com for a PDF). Accordingly, 617 small molecules were active in at least one assay, with 475 small molecules active in the AcTubulin assay and 344 small molecules active in the AcLysine assay (Figure 2D). Overall, a significant correlation ($r = 0.39$, p value < 0.05) exists between the normalized acetylation level in the AcTubulin versus the AcLysine assays (Figure 3A). Indeed, deconvolution of the 617 hit compounds into those specific for the AcTubulin (273) or the AcLysine (142) assay revealed that 202 (33%) molecules scored in both assays (Figures 3B and 3C).

Because there was an unequal distribution of posi-

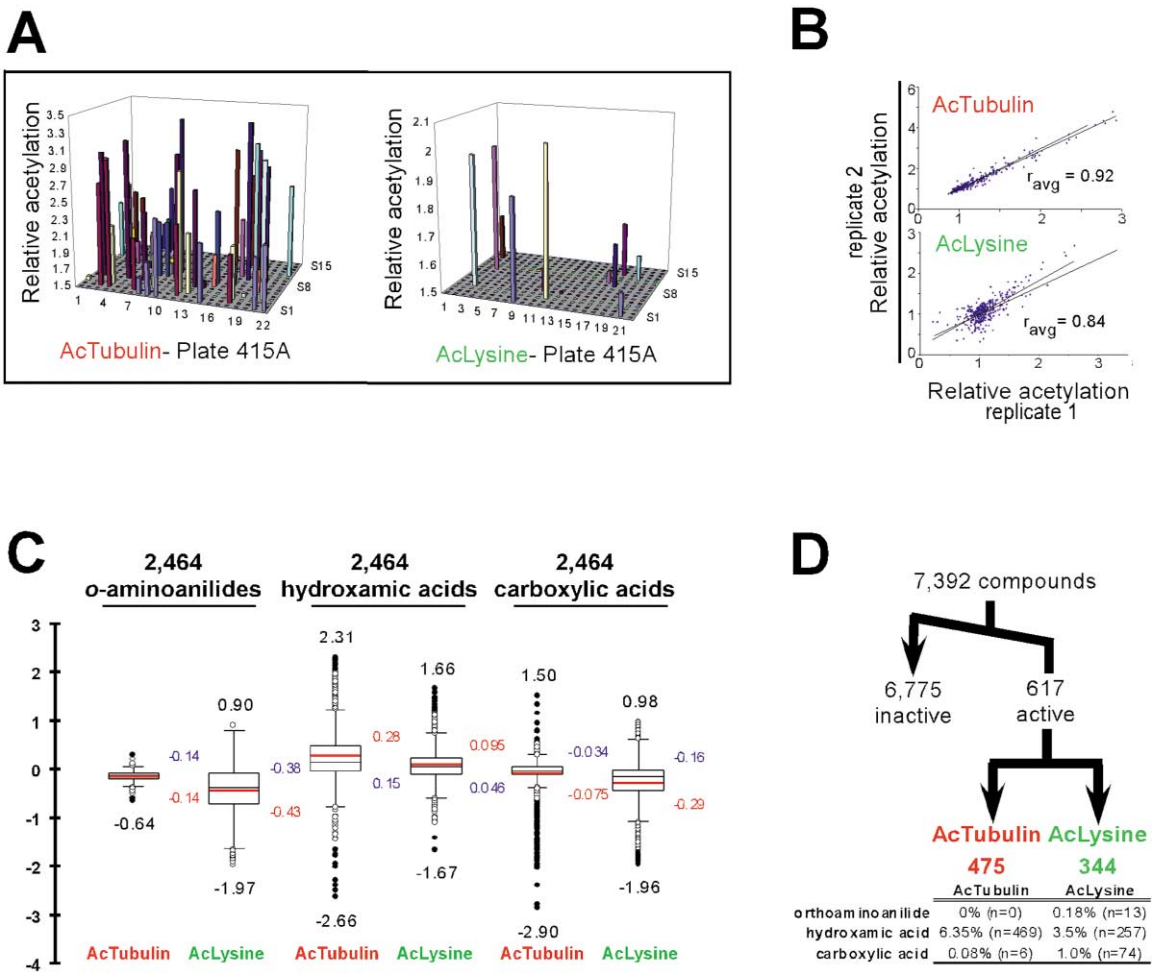


Figure 2. Statistical Analysis of the Biological Activity of the 1,3-Dioxane Library

(A) Example of raw screening data of one 384-well plate from the AcTubulin and AcLysine assays with the relative acetylation level corresponding to the fold change in signal compared to background.

(B) Average (n = 7 plates) correlation of duplicate plates within the set of hydroxamic acids.

(C) Box plot showing summary of statistical properties of the screening data from the three structural classes of deacetylase inhibitors after averaging duplicate data and Log_2-transformation. Top and bottom numbers (black) are the highest and lowest extremes, respectively. Blue numbers are the median values and red numbers the mean value for each distribution. The inner box represents the upper and lower bound of the third and first quartile, respectively, which contains 50% of the distribution. The upper and lower bars represent 1.5-times the spread of the third to first quartile with black/white dots representing small molecules outside this range.

(D) Number of bioactive small molecules and distribution amongst the three structural classes using a normalized acetylation value of 1.5 as the criterion.

tives toward those small molecules containing hydroxamic acids at diversity position R^4 in both the AcTubulin and AcLysine cytoblots, we focused subsequent screening efforts on the 2464-member hydroxamic acid subset of the 1,3-dioxane library. These compounds were screened in the AcTubulin assay in the presence of chemical genetic modifier ITSA1, which suppresses trichostatin-induced histone and α-tubulin acetylation [21]. Two additional assays were employed to measure acetylation of histone H3 (AcHisH3) and histone H4 (AcHisH4). Using the same statistical analyses and threshold used for the AcTubulin and AcLysine assays, 137 small molecules caused increased α-tubulin acetylation in the presence of ITSA1, 229 caused an increase in AcHisH3, and 231 caused an increase in AcHisH4 (see Supplemental Table S1 at URL above).

Suppression of α-Tubulin Acetylation by the Chemical Genetic Modifier ITSA1

Over the hydroxamic acid subset, the correlation between the AcTubulin and AcTubulin + ITSA1 assays remained significant ($r = 0.65$, p value <0.05). However, in the AcTubulin + ITSA1 cytoblot, reduction in the number of positives (137 versus 475) indicates that ITSA1 effectively suppressed many of the inducers of α-tubulin acetylation. ITSA1-resistant small molecules are, in general, among the most potent inducers of α-tubulin acetylation (Figure 3D). Interestingly, not all of the hydrox-

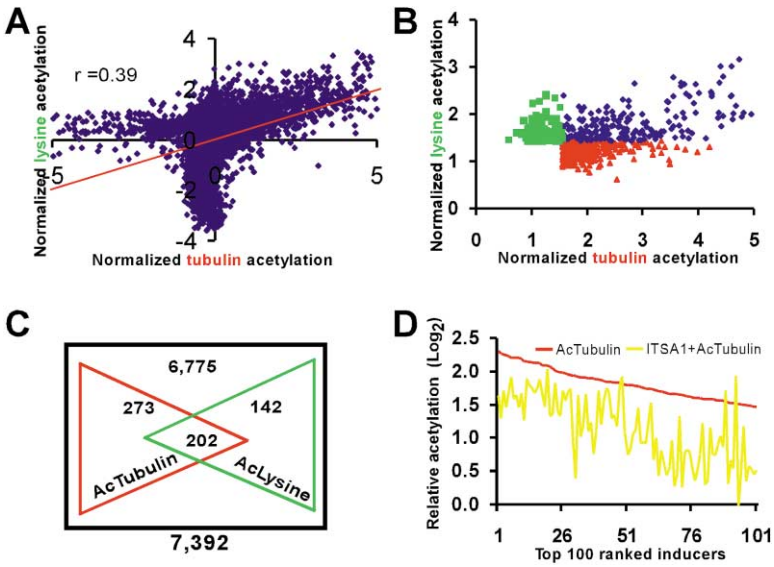


Figure 3. Correlation between AcTubulin and AcLysine Assays and Suppression of α-Tubulin Acetylation by ITSA1

(A) Linear correlation of the normalized acetylation values of the 7392 members of the 1,3-dioxanes in the AcLysine and AcTubulin assays.

(B) Six hundred seventeen of the total library members with bioactivity using a 1.5-fold increase in the normalized acetylation value as the criterion (blue, active in both the AcTubulin and AcLysine assays; green, active in only the AcLysine assay; red, active in only the AcTubulin assay).

(C) Venn diagram of the 617 bioactive small molecules deconvoluted into selective and overlapping sets.

(D) ITSA1 (50 μM) suppressed the increased α-tubulin acetylation (yellow line) induced by most but not all of the top 100 ranked positives in the AcTubulin assay (red line).

amic acids that induced α-tubulin acetylation were suppressed by ITSA1, making it unlikely that ITSA1 reacts directly with this functional group.

Representation of Screening Results as a Chemical Genetic Network

To visualize patterns and compute global properties of MDCG screening data, a general approach for analyzing networks of interactions would be useful. Accordingly, using principles from graph theory [32] (see http://vlado.fmf.uni-lj.si/pub/networks/pajek/), the screening data were used to construct a symmetric, square adjacency matrix ($\mathbf{A}_{nxn}$). Here, n equals the number of vertices/nodes ($\boldsymbol{V}$), representing an assay or one of the 617 total active small molecules. In $\mathbf{A}_{nxn}$, each element a_{ij} = 1 if the vertices v_i and v_j are connected by an edge (E), which indicates the activity of compound v_i in assay v_j. Alternatively, a_{ij} = 0, which indicates the inactivity of compound v_i in assay v_j. From the information encoded in this adjacency matrix, the corresponding bipartite graph, $\boldsymbol{G}_{deac}$ = ($\boldsymbol{V}$, $\boldsymbol{E}$), was constructed (Figure 4A). To aid in visualization of the graph, we applied the Fruchterman-Reingold algorithm (see URL above) [33]. This algorithm is a "spring embedder" that considers the graph as a physical system composed of charged masses (vertices $\boldsymbol{V}$) repelling each other and connecting springs (edges $\boldsymbol{E}$) pulling adjacent vertices together. The algorithm minimizes the "energy" of the system when embedding the graph into a two- or three-dimensional space. Overall, using a 1.5-fold minimum normalized acetylation value as the criterion for 1,3-dioxane bioactivity, the resulting graph ($\boldsymbol{G}_{deac}$) contains 620 nodes $\boldsymbol{V}$ (617 from chemicals, 3 from assays), 956 edges $\boldsymbol{E}$, and has a bipolar, spindle-like structure (Figure 4A). Here, edges on the ends of the graph generally represent small molecules selective for either the AcTubulin or AcLysine assays, as they radiate singly from their respective nodes. In the center of $\boldsymbol{G}_{deac}$ are the most highly connected nodes of the chemical genetic network. These are the subset of small molecules that scored in both the AcTubulin and AcLysine assays, a portion of which also scored in the ITSA1 + AcTubulin assay. Small molecules not suppressed in ITSA1 + AcTubulin are either the most potent α-tubulin deacetylase inhibitors or act mechanistically distinct from TSA. Thus, by further subdividing small molecules based upon interaction with ITSA1, five classes of deacetylase inhibitors with different patterns of bioactivity are observable in $\boldsymbol{G}_{deac}$.

Small molecules are now routinely compared by decomposing their structures into components using molecular descriptors (e.g., cLogP, molecular weight, Chi connectivity indices). By analogy, in order to quantify the information obtained in each assay and to compare the structure of the resulting networks, the adjacency matrix $\mathbf{A}_{nxn}$ was recursively decomposed into six submatrices. To quantitatively compare these six networks, we computed a set of three graph-theoretical descriptors for the corresponding networks (Figures 4B and 4C; see http://vlado.fmf.uni-lj.si/pub/networks/pajek) [32]. Based on the relative differences in these descriptors, the most complex subgraph was that of AcTubulin + AcLysine (A + B). Since the ITSA1 + AcTubulin assay (C) did not contribute new nodes to the graph (compare A + B to complete graph A + B + C in Figure 4A), all small molecules that scored in the ITSA1 + AcTubulin assay (C) also scored in the AcTubulin assay (A). Overall, the A + C and B + C subgraphs were similar with respect to the adjacency and Randic connectivity indices, but B + C had a lower value for the Zagreb index, indicating lower vertex connectivity.

Global Analysis of Screening Data: Hierarchical Clustering and Principal Component Analysis

A primary objective of this study was to determine the similarity/differences of chemicals based upon their biological interactions or of biological assays based upon their chemical interactions. Toward this end, standardized covariance matrices were computed from the origi-

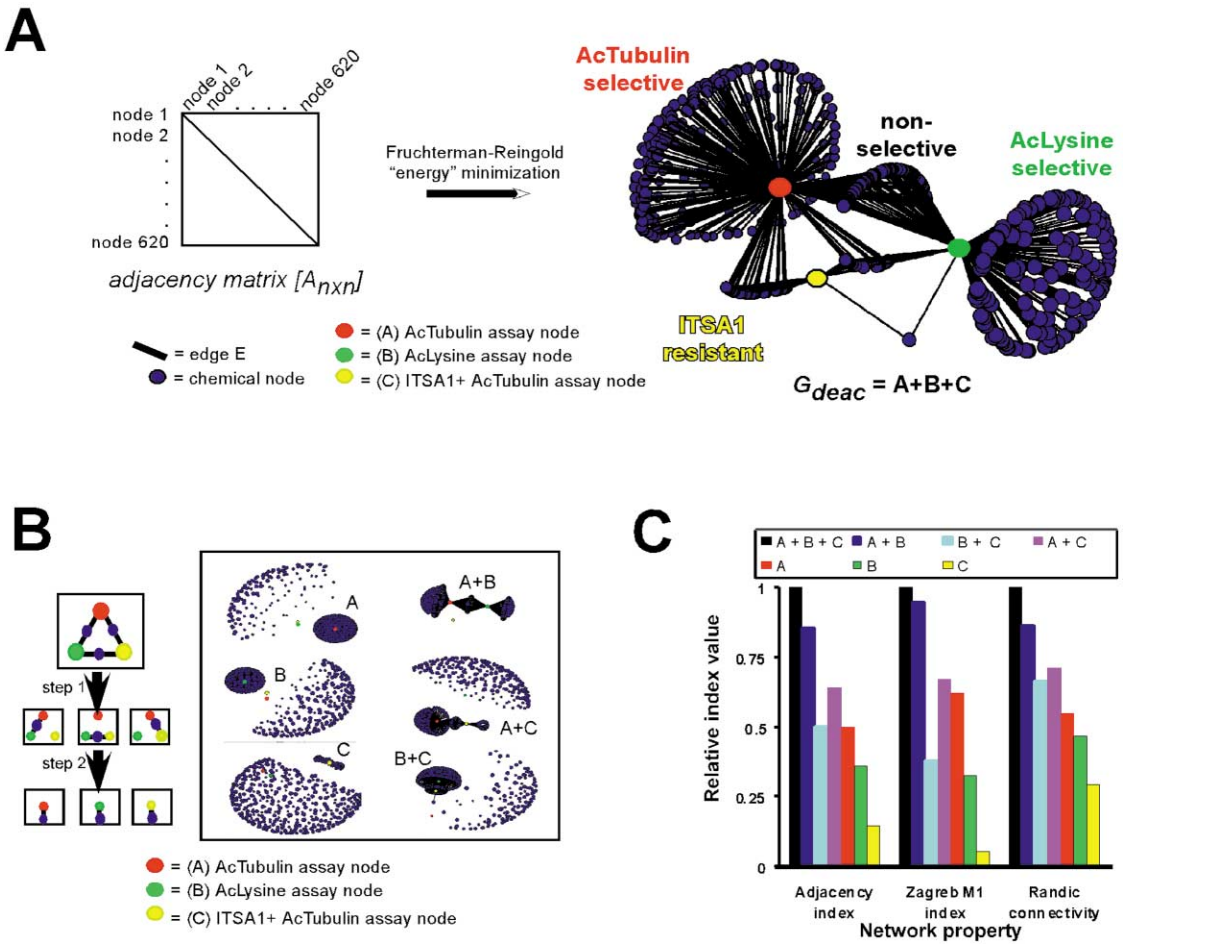


Figure 4. Chemical Genetic Network from Screening Data

(A) Adjacency matrix and resulting graph (G_{deac}) after applying the Fruchterman-Reingold "energy" minimization algorithm. Nodes represent either assays or small molecules according to the indicated colors. Edges (black lines) connect bioactive small molecules to the corresponding assay.

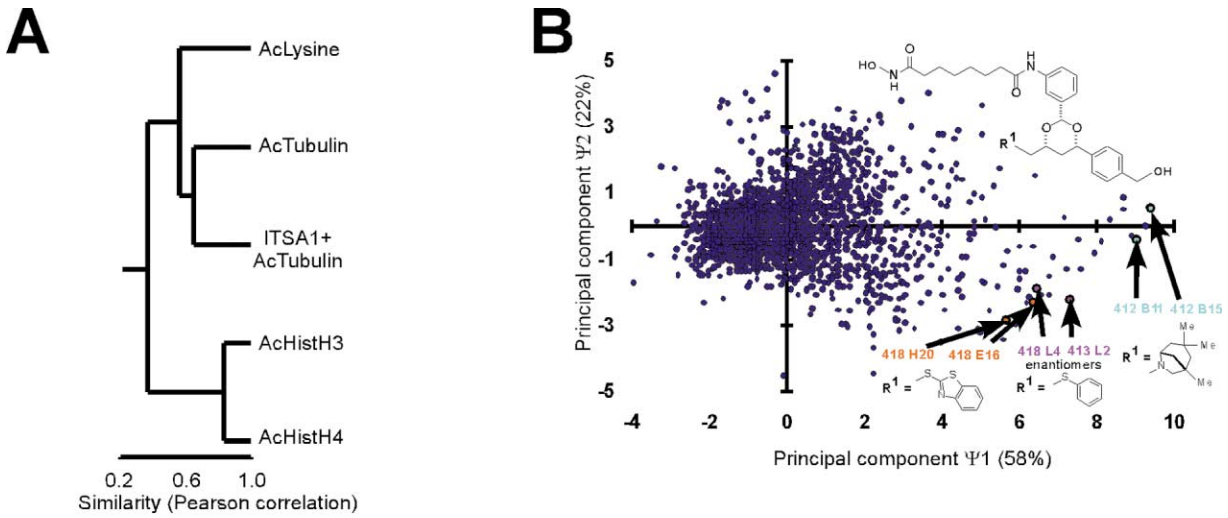
(B) "Retrosynthetic analysis" of G_{deac} through decomposition into the subgraphs from all six combinations of assay data.

(C) Comparison of the information content in the six subgraphs, relative to the full graph, using graph-theoretical descriptors. The adjacency index is defined to be the sum of the lengths of minimal paths between all pairs of vertices. Since all edges are of equal length in these particular cases, this is equal to half the sum of all entries in the adjacency matrix: $W = [\Sigma\Sigma A_{nxn}]/2$. The Zagreb M1 index is defined as the sum of the squares of the vertex connectivities: $Z = \Sigma c_n^2$, where c_n is the number of edges from a particular vertex n. The Randic connectivity index, which encodes information about branching, is defined as ${}^1\chi_p = \Sigma\,(c_n c_m)^{-1/2}$, where c_n and c_m are the number of edges of adjacent vertices joined by each edge summed over all edges. For all three indices, a high value corresponds to a high percentage of the particular type of topological property in the graph being analyzed.

nal chemical genetic data matrix using the Pearson correlation coefficient. Covariance matrices are square, symmetric matrices with off-diagonal elements representing the correlation between descriptors. For all observed phenotypes, MDCG assigns an n-dimensional operator, Σ_{bio}, represented by a standardized covariance matrix that is computed globally across the set of n biological assays. Similarly, for all observed chemicals, MDCG assigns an m-dimensional operator, Σ_{chem}, represented by a standardized covariance matrix that is computed globally across the set of m small molecules. This information can be further analyzed using clustering algorithms and other methods of pattern finding.

Clustering was applied to data from the hydroxamic acid subset of the library across five different cell-based assays. This analysis indicated the greatest similarity between the AcTubulin and ITSA1 + AcTubulin data sets when compared to any of the assays measuring histone acetylation levels (Figure 5A). Regarding the three acetyl-histone assays, the AcHisH3 and AcHisH4 assays are more correlated to each other than either is to the AcLysine assay. Overall, these results demonstrate that while there is redundancy between the assays, global differences exist in the window of biology described.

We next computed a principal component model ([34, 35]; see Data Analysis under Experimental Procedures) from the standardized covariance matrix of the hydrox-



C

	λ1	λ2	λ3	λ4	λ5
Eigenvalue	2.92	1.11	0.51	0.30	0.16
variance %	58.3	22.1	10.3	6.0	3.3
cumulated %	58.3	80.4	90.7	96.7	100.0

Eigenvalues

	Ψ1	Ψ2	Ψ3	Ψ4	Ψ5
AcTubulin	20.5	20.4	0.3	57.8	1.0
AcLysine	18.9	11.4	52.9	16.7	0.1
ITSA1+AcTubulin	20.1	9.3	46.7	23.4	0.5
AcHist3	19.9	30.6	0.0	0.1	49.4
AcHist4	20.6	28.3	0.1	2.0	49.0

Contribution of variables to principal components Ψ1 - Ψ5

Figure 5. Hierarchical Clustering and Principal Component Analysis of 2464 Hydroxamic Acids from a Five-Dimensional Chemical Genetic Data Matrix

(A) Dendrogram from clustering the standardized covariance matrix of chemical genetic assay data (Σ_{bio}) using the unweighted pair-group average method and Pearson correlation as the distance metric.

(B) The principal components are obtained by solving the algebraic eigenvalue problem: $\Lambda = \Psi^T \Sigma_{bio} \Psi$, where Λ is the eigenvalue matrix, Ψ is the eigenvector matrix, Σ_{bio} is the standardized covariance matrix of chemical genetic assay data, and T denotes a transpose of a matrix. Ψ defines a coordinate transform (rotation) that best decorrelates the data into orthogonal linear subspaces. The chart shown plots the location of the hydroxamic acids (blue dots) on the reduced space formed by the first and second principal components. Position and structures of three of the decoded small molecules (colored) chosen for their overall activity are shown.

(C) Summary of variance accounted for by each eigenvalue ($\lambda 1 - \lambda 5$) and the contribution of each of the five original assay variables to the principal component (Ψ_1–Ψ_5).

amic acid subset (Figure 5B). The result is a global model that provides a visualizable representation of a chemical space that minimizes the information lost upon projection of the elements into a reduced space of one to three dimensions. Accordingly, principal components can be used to position chemicals, with respect to the new system of coordinate axes, in terms of a linear combination of the original assay variables. Using this method, principal component analysis (PCA) revealed that 58% of the information in the data set is accounted for in a one-dimensional space represented by the first principal component (Ψ_1). The second principal component (Ψ_2) represents a total of 22% of the information. Examination of the relative contribution of the different assays and small molecules to the information represented on Ψ_1 indicated that this component represents a composite measure of the activity of small molecules in all five assays (Figure 5C; see below for details of structures and use of this model in determining regions of selectivity).

Correlating Chemical Structure with *Activity*: Hydroxamic Acids Are the *Most Bioactive* Small Molecules across Assays

Seventeen small molecules exhibiting the highest levels of activity in both the AcTubulin and AcLysine assays were selected for bead decoding and structure determination (see Supplemental Figure S1 at http://www.chembiol.com/cgi/content/full/10/5/383/DC1 or write to chembiol@cell.com for a PDF) [25, 30]. Notably, of the 15 structures determined successfully, all are hydroxamic acids with six-carbon linkers. In addition, two molecules were present in duplicate within the set of 15, and both enantiomers of another molecule were identified. When the position of each decoded compound was plotted in the reduced space formed by

principal components Ψ_1-Ψ_2, these identical or enantiomeric pairs were within very close proximity (Figure 5B). Thus, the projection of a compound on Ψ_1 is positively correlated with a composite measure of bioactivity. Although we have not yet determined the structure of all small molecules within this "high activity neighborhood" across all five assays, we expect many of the other molecules in this region to share similar structural properties.

Correlating Chemical Structure with *Selectivity*: the *Most Selective* Class of Small Molecules Varies between Assays

In order to identify the most selective small molecules in both the AcTubulin and the AcLysine assays, PCA was performed on each structural class of 1,3-dioxanes defined by the three functionalities incorporated at position R^4 (Figure 1B). As this analysis used data from the two assays in which the entire library was screened (AcTubulin and AcLysine), the input space was two-dimensional. As such, PCA centered the data based on the distribution centroid and rotated the original axes as specified by the eigenvectors, without any information loss. As shown (Figure 6A), for both *o*-aminoanilides and carboxylic acids, the two eigenvalues ($\lambda 1$- $\lambda 2$) account for roughly equal amounts of the variance. This corresponds to low activity of these sets of small molecules and low correlation between bioactivity in one assay versus the other. Indeed, none of the 2464 *o*-aminoanilides showed any activity toward inhibiting α-tubulin deacetylation, while 0.5% of these small molecules inhibited histone deacetylation (Figure 2D). The *o*-aminoanilide functionality may therefore provide an effective means to target histone deacetylation without affecting α-tubulin acetylation levels (see [22] for more details). In contrast, $\lambda 1$ accounted for more than 80% of the variance in the set of hydroxamic acids, which were on average more active than either the *o*-aminoanilides or carboxylic acids. This difference in eigenvalue magnitude corresponds to the high degree of bioactivity and strong correlation between the AcTubulin and AcLysine assays. Consequently, one principal component is capable of representing the majority of the information. The resulting bioactivity distributions on Ψ_1-Ψ_2 are shown in Figure 6B. Based upon the average activity of the class of small molecules and position within the rotated space of the principal components, a set of small molecules was chosen for bead decoding and structure determination (see Supplemental Figures S2 and S3 at http://www.chembiol.com/cgi/content/full/10/5/383/DC1 or write to chembiol@cell.com for a PDF) [23, 28] Twenty-four of these were selective for the AcTubulin assay, and 22 were selective for the AcLysine assay. Calculated masses of the inferred chemical structures matched those determined by liquid-chromatographic mass spectrometric analysis of the corresponding stock solutions (data not shown). However, because chemical-encoding strategies record chemical history only, these structural assignments should be considered tentative.

Following decoding, we noted that one structure, from two different plate positions, was common to both the AcLysine (418 B2) and AcTubulin (416 N3) selective lists. This anomalous result may be due to statistical variation in the assay results, differences in stock solution concentration, impurities, or an insufficiently stringent measure of selectivity. Since the majority of the selective inhibitors in both the AcTubulin and AcLysine assays, and all of the most potent deacetylase inhibitors, were hydroxamic acids, the relative position of most of these compounds could be obtained using the three-dimensional PCA model derived above using five assay descriptors (Figure 6C). In this model, which accounts for 90.7% of the variance in the data, the AcTubulin-selective (red) and AcLysine-selective (green) sets clustered together and contributed less to the variation described by Ψ_1 then did either the full set of 2464 small molecules (inset; black spheres) or the most potent inhibitors (blue; not all depicted).

Tubacin, a Selective Inducer of α-Tubulin Acetylation

Two small molecules were chosen as selective inducers of α-tubulin acetylation, based upon their position in the rotated space of Ψ_1-Ψ_2. No apparent selectivity was observed for one of these molecules (413 D10) when a sample was tested in the secondary fluorescence microscopy assays for acetylated α-tubulin and acetylated lysine. This discrepancy might arise from higher compound concentrations in the secondary assay of ($\sim$10 μM versus a $\sim$2–5 μM screening concentration) or increased sensitivity in the fluorescence microscopy assay. However, the second compound 415 N3 (Figure 7A), here named tubacin (**1**) (*tub*ulin *ac*etylation *in*ducer) strongly increased α-tubulin acetylation with no effect on lysine acetylation. This small molecule was resynthesized to allow a systematic assessment of the effect on α-tubulin and histone acetylation. Treatment of A549 cells with tubacin (19 hr) strongly increased α-tubulin acetylation levels (Figure 7B) at concentrations as low as 125 nM (4 hr, data not shown). Consistent with the original screening data, tubacin did not affect histone acetylation and was partially suppressed by ITSA1 (Figure 7B). Thus, tubacin is the first known selective inhibitor of α-tubulin deacetylation.

In the three-dimensional PCA model (Figure 6C), the most proximal compound to tubacin (415 N3) was 415 O6. This compound had the identical structure as another decoded compound, 418 D6, which was also among the closest in proximity to tubacin in the three-dimensional PCA model (Figure 6C). Notably, 415 O6/418 D6 is one of the $\sim$1.4% of the total library members (not considering stereochemistry) that share the same building blocks at three of the four diversity positions (R^2-R^4) as tubacin. Furthermore, at the position that is different from tubacin (R^1), the 2-mercaptopyridine group shares some structural similarity to the core of the 4,5-diphenyl-2-oxazolethiol building block present in tubacin. In agreement with these structural similarities, retesting of the original stock solution of 418 D6 using fluorescence microscopy confirmed its selective inhibition of α-tubulin deacetylation (data not shown).

Histacin, a Selective Inducer of Histone Acetylation

Two small molecules were chosen as selective inducers of histone acetylation based upon their position in the

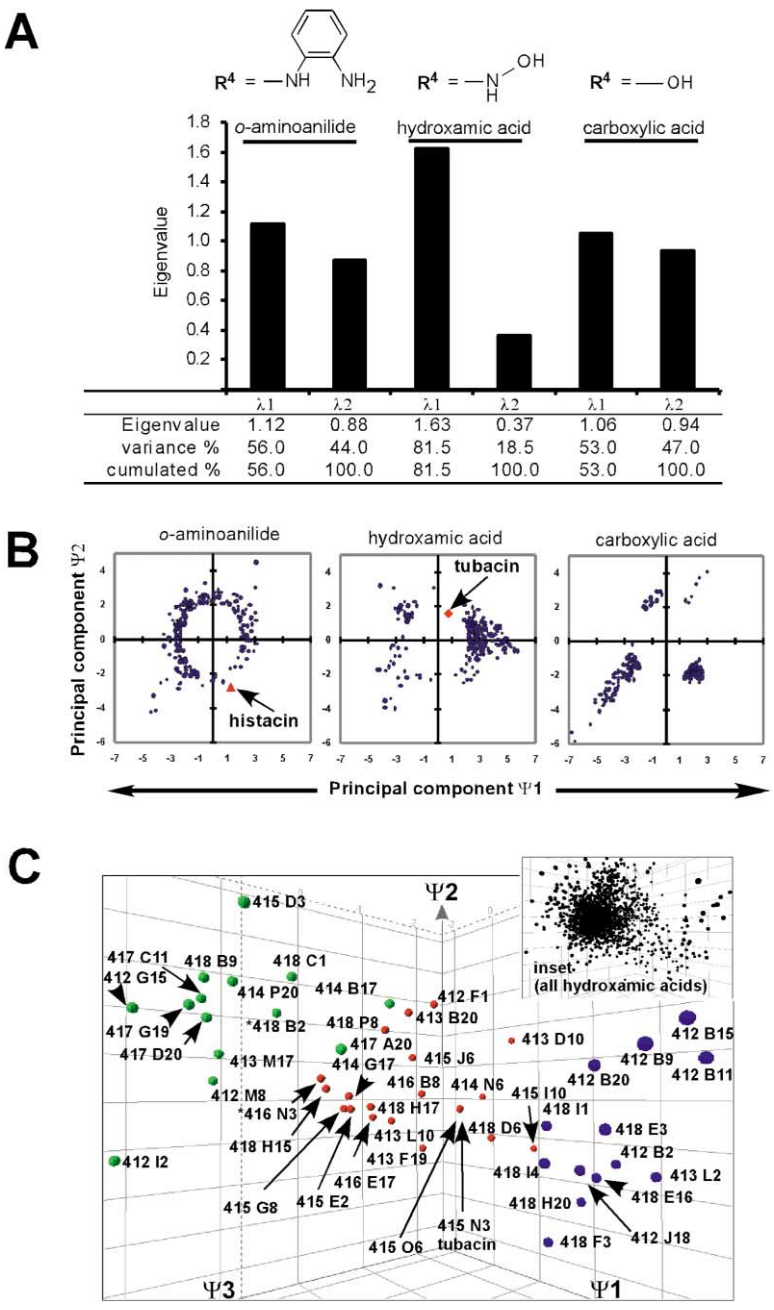


Figure 6. Principal Component Analysis of Diversity Position R^4

(A) Eigenvalue spectrum and variance associated with the three structural classes of deacetylase inhibitors.

(B) Charts plotting a random sample of 10% of the small molecules (blue dots) of each structural class on the rotated space of Ψ_1 and Ψ_2. The location of histacin and tubacin, selective inhibitors of histone and α-tubulin deacetylation, respectively, are shown.

(C) Relative position of decoded structures (see Supplemental Figures S1–S3 at http://www.chembiol.com/cgi/content/full/10/5/383/DC1 or write to chembiol@cell.com for a PDF) in a PCA model computed from five cell-based assay descriptors (see Figure 5C). Ac-Tubulin selective, red; AcLysine selective, green; and most potent, blue.

rotated space of Ψ_1-Ψ_2. Again, no selectivity was observed for one of these small molecules (417 A20) when tested in the secondary fluorescence microscopy assays for acetylated α-tubulin and acetylated lysine. However, the second compound 410 F1 (Figure 7C), here named histacin (**2**) (*hist*one *ac*etylation *in*ducer), strongly increased lysine acetylation with no effect on α-tubulin acetylation (Figure 7D). After resynthesis, treatment (14 hr) of cells with histacin (20 μM) increased acetylated histone levels without affecting α-tubulin acetylation levels. Thus, histacin, like the epoxy ketone-containing HDAC inhibitors trapoxin and HC toxin, is a selective inhibitor of histone deacetylation. Furthermore, like trapoxin/HC toxin, histacin's effects on histone acetylation cannot be suppressed by ITSA1 (data not shown).

Discussion

Classical genetics began annotating genetic factors by observing heritability and determining the linkage of simple phenotypic traits. Using these observations and the frequency of recombination during the first meiotic division as a metric, the relative distance between genes encoding for phenotypes was experimentally defined, one map unit being equal to 1% recombination (measured in cM) [36]. Accordingly, a mutant gene could be "mapped" as a point in a one-dimensional space. Through

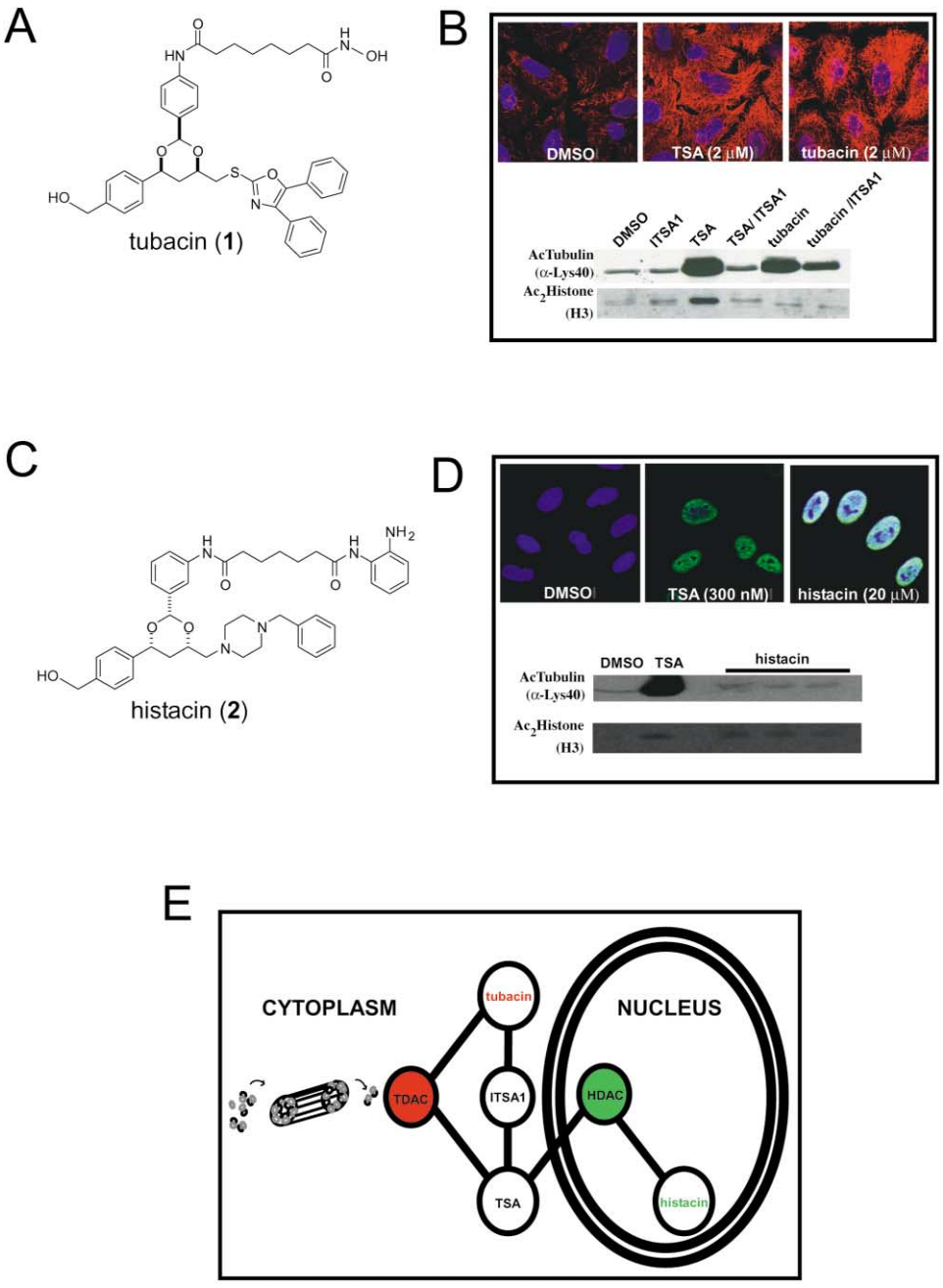


Figure 7. Selective Inhibitors of α-Tubulin and Histone Deacetylation

(A) Chemical structure of tubacin (compound 415 N03).

(B) Effect of tubacin (2 μM) and trichostatin (2 μM) treatment (19 hr) on the acetylation level of α-tubulin (red) in A549 cells measured by immunofluorescence. Western blot analysis of acetylated α-tubulin and acetylated histone H3 (K9, K14) in A549 cells pretreated (2 hr) with tubacin (2 μM) followed by ITSA1 (50 μM) for an additional 2 hr. See [22] for further analysis of tubacin and its inhibition of HDAC6.

(C) Chemical structure of histacin (compound 410 F1).

(D) Effect of histacin (20 μM) and trichostatin (300 nM) treatment (5 hr) on the acetylation level of histone H3 (yellow/green) in A549 cells measured by immunofluorescence. Western blot analysis of acetylated α-tubulin and acetylated histone H3 (K9, K14) in histacin (20 μM) treated (5 hr) A549 cells. See [24] for further analysis of 410 F01 (histacin).

(E) Network of genetic and chemical genetic interactions among 1,3-dioxane-based deacetylase inhibitors.

overlapping distance measurements, a genetic map could be constructed. Although not obvious at the onset, it is now well known that such maps represent the physical arrangement of genes within a linear and continuous sequence of deoxyribonucleic acids.

By analogy to the logic of classical genetics, we are interested in creating "chemical genetic maps" that position chemicals and biological systems in a multidimensional space. Such maps will facilitate systematic analysis of the factors determining interactions between small

molecules and biological systems. At the core of such efforts is the relationship between information present in chemical genetic observations and mathematical quantities that make up the geometric properties of vector spaces. While chemical space is often considered a vector space with dimensions represented by calculated molecular descriptors, we instead focused on measuring the similarity/differences of small molecules using phenotypic properties from cell-based assays. Pioneering work by the group of Dr. John Weinstein demonstrated that data sets representing interactions of small molecules with tumor cell lines were rich in information [35–38]. Although limited to a single phenotype (cellular growth as measured by protein production), this analysis was able to classify small molecules into similar patterns of activity and tumor cell lines into similar groups. Here, the logic of this type of analysis is applied to a set of biochemical phenotypes, specifically acetylation levels of α-tubulin and histones in cells. In addition, the small molecules screened derive from a biased library rather than a "random" collection. Our initial analyses were limited to one genotype (that of a human lung cancer cell line) and thus could be expanded to include varying cell types and/or genotypes.

Genes can have multiple functions, as observed by multiple phenotypic effects resulting from their mutation. Similarly, small molecules can have multiple functions, as observed by multiple phenotypic effects resulting from their interaction with cellular components. The distribution of activities in a given assay, correlation across assays, and the relationships of the assays themselves renders the number of comparisons required to find certain patterns of activity intractable. Accordingly, modeling and visualizing multidimensional data requires the use of dimensionality reduction and pattern finding algorithms. This allows higher-level representation of the information inherent in lower-level relational data.

For this purpose, we introduced the concept of a chemical genetic network for the analysis of screening data. If two small molecules score in the same phenotypic assay, then there is some probability that they are hitting the same target. However, if the same two small molecules exhibit different patterns of biological activity in other phenotypic assays, the probability that they are targeting the same gene product is reduced. Thus, the connectivity of a chemical node can be used to constrain the set of possible targets and to provide a discrete measure of a small molecule's selectivity. Similarly, the connectivity of an assay node can be used to compare the aspects of a biological system being queried and to provide a discrete measure of an assay's sensitivity. Although the network constructed here was based upon an empirically defined threshold, similar networks could be formed at different levels of bioactivity. Furthermore, rather than using a binary edge length (= 0 or 1), the lengths could be made proportional to the bioactivity of a compound or a likelihood score. As the annotation and subdivision of a chemical genetic network increases, predictions of whether a new member of a class will have additional targets within a cell or will interact with a specific chemical genetic modifier may be possible.

As another form of multidimensional data analysis, we used a computational method called principal component analysis [34, 35]. This eigenvalue/eigenvector approach takes into account the global properties of the space spanned by the variables under study. In doing so, it reduced the probability of selecting a compound as selective only due to a technical error or the compound having an undesired cellular effect. Although we considered only a two-dimensional space spanned by Ψ_1 and Ψ_2, the logic of this analysis holds upon the addition of more dimensions. Thus, PCA is well suited for the discovery of molecules with complex phenotypic effects, and not just activity in one particular assay. In the global analysis of the hydroxamic acids in the five assays, the first eigenvalue ($\lambda 1$) accounted for most of the information in the data set and correlated with the bioactivity of the small molecules. In the discovery of *selective* inhibitors of α-tubulin deacetylation or histone deacetylation, the two- and three-dimensional PCA models delineated regions of chemical space with different patterns of bioactivity (Figures 6B and 6C). Besides allowing for a more informative visualization of multidimensional data, PCA has a practical application for data analysis, as the reduced number of dimensions simplifies subsequent computations that may be memory and time intensive.

The possibility exists that both tubacin and histacin affect the acetylation levels of proteins other than α-tubulin and histones, respectively. Future work with the small molecules discovered here will be directed toward determining their potency and testing for differential phenotypes on gene expression, stem cell differentiation, and zebrafish embryogenesis. With the remaining amount from each stock solution, additional screens of the 1,3-dioxane library will be aimed at determining the extent to which the small molecules affect other histone modifications, metalloenzymes (i.e., deubiquitinases and proteases), and other biological processes.

An outstanding question in the field of chemical biology is how best to quantify molecular diversity in a way that informs biological discovery. Although molecular descriptors based upon a chemical graph (structure) are generally trivial to compute, structure-activity relationships are often complex. As an alternative, the effectiveness of diversity-oriented syntheses can be quantified by the dimensionality of the biological space required to effectively describe the observed phenotype. For example, had the 1,3-dioxane library been composed of *o*-aminoanilides alone, all molecules could have been positioned on a one-dimensional space, since none of them scored in the AcTubulin assay. In contrast, a 1,3-dioxane library composed solely of hydroxamic acids would require two dimensions, one for AcLysine and AcTubulin. Thus, through systematic screens of small molecule libraries in minimally redundant, maximally informative cell-based assays, maps of chemical space can be derived from biologically based descriptors. Such maps may have more distinct geometric and topological properties than those derived from calculated descriptors.

The utility of antibodies to detect specific posttranslational modifications in high-throughput phenotypic assays provides a powerful approach to identifying novel biologically active small molecules [9, 11, 21, 38].

This complete screen consumed less than 20% (~20 nmol) of the compound released from one synthesis bead, leaving the majority of compound for retesting, analytical analysis, or subsequent additional screening. The overall high level of observed activity (~8% of library small molecules active) at a concentration of 2–5 μM reflects prior knowledge guiding library synthesis, as well as apparent sensitivity of HDACs to small molecule modulation. While a number of small molecules (~80) have previously been reported to inhibit HDAC activity in vitro, only a limited set of these appear to be effective inhibitors of HDAC activity within the context of living cells [19, 20]. Since the cell-based screening methodology described in this paper is scalable and can be generalized to other posttranslational modifications, the development of additional assays coupled to important biological processes is possible. In particular, iterative cycles of chemical genetic modifier screens, using newly discovered bioactive small molecules, should facilitate the construction of chemical genetic networks for use in "chemical genomic profiling" experiments. An important direction in such studies will be the development of novel screening methodologies with increased efficiency, information content, and accuracy. Here, the combination of small molecules and RNAi-based perturbations promises to be an illuminating direction of research. In parallel, along with increasing the diversity of small molecule modulators, deconvolution of synthetic variables, such as efficiency and purity, will continue to be required. In the limit, beyond traditional functional characterization, one can imagine annotating proteins encoded within an entire genome using a set of small molecule modulators of their biochemical activities [4, 10].

Significance

The identification of 617 small molecule inhibitors of intracellular deacetylation highlights multidimensional chemical genetics and diversity-oriented organic synthesis as tools for the discovery of probes of complex biological networks. New small molecules with diametrically opposed inhibitory activities (tubacin/tubulin deacetylase inhibitor versus histacin/histone deacetylase inhibitor) provide a means to dissect the role of acetylation in microtubule dynamics and chromatin remodeling [21, 22, 24]. The discovery of potent and especially selective deacetylase inhibitors was possible through computational methods, most notably principal component analysis and consideration of proximity in multidimensional descriptor spaces. Using the principles of graph theory, a discrete model of the screening data was derived in the form of a "chemical genetic network" and analyzed using topological invariants. By systematically mapping the structure-activity relationships of this particular "deacetylase" region of chemical space, the results of this study should assist in the discovery of selective HDAC inhibitors that may be more effective research tools and therapeutic agents.

Experimental Procedures

Trichostatin A (TSA) and anti-acetylated α-tubulin (6-11B-1) antibody were purchased from Sigma. ITSA1 (5253409) was purchased from Chembridge. Anti-acetylated lysine antibody was purchased from Cell Signaling. Anti-acetyl histone H3 (K9, K14) and anti-acetyl histone H4 (K4, K7, K11, K15) antibodies were purchased from Upstate Biotechnology. Anti-mouse IgG horseradish peroxidsae (HRP) conjugated, anti-rabbit IgG-HRP conjugated secondary antibodies, and enhanced chemiluminescent mixture (luminol) were purchased from Amersham Pharmacia. Alexa 594 and Alexa 488-conjugated anti-mouse IgG antibody, anti-rabbit IgG antibody, and Hoechst 33342 were purchased from Molecular Probes.

Cell Culture

A549 human lung carcinoma cells (American Tissue Culture Collection) were cultured at 37°C with 5% carbon dioxide in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% v/v fetal bovine serum (Gibco BRL), 100 units/ml penicillin G sodium (Gibco BRL), 100 μg/ml streptomycin sulfate (Gibco BRL), and 2 mM L-glutamine (Gibco BRL) ($DMEM^+$).

1,3-Dioxane Library Screen and Cytoblot Assays

A549 cells were seeded at a density of 4000 cells/50 μl in $DMEM^+$ in white tissue culture-treated 384-well plates (Nalge Nunc) using a Multidrop 384 liquid dispenser (Labsystems). Library small molecules were pinned twice (100–150 nl/pin) from ~1 mM stock solutions (DMSO) and incubated for 18 hr. For the ITSA1 + AcTubulin assay, ITSA1 (50 μM) was added four hours after pin transfer of library molecules. Media was aspirated with a 16-channel wand. For the acetylated α-tubulin cytoblot (AcTubulin), 40 μl/well of fixation solution (3.7% formaldehyde [1:10 from 37% stock] in Tris-buffered saline [TBS; 0.15 M NaCl, 0.02 M Tris-Cl pH 7.4]) was added, and plates were incubated for 20 min at room temperature. After aspiration, 30 μl 100% methanol (−20°C) was added and incubated for 5 min at room temperature. After aspirating and washing three times with 90 μl of antibody dilution buffer (ADB; TBS/2% bovine serum albumin, 0.1% Triton X-100), the wells were blocked for 20 min in ADB. After aspirating, 20 μl/well of ADB containing anti-acetylated α-tubulin antibody (mouse; 1:1000) and anti-mouse IgG horseradish peroxidase (HRP)-conjugated antibody (1:750) was added and incubated overnight (4°C). For the anti-acetyl lysine cytoblot (AcLysine), 40 μl/well of fixation solution (95% ethanol/5% acetic acid, −20°C) was added, and plates were incubated for 5 min at room temperature. Wells were aspirated, washed, and blocked in ADB as for the AcTubulin protocol. After aspirating, 20 μl/well of ADB containing anti-acetylated lysine (rabbit; 1:375) and anti-rabbit IgG-HRP conjugated antibody (1:750) was added and incubated overnight (4°C). For the anti-acetylated histone H3 (AcHistH3) and anti-acetylated histone H4 (AcHistH4) cytoblots, cells were fixed, washed, and blocked in ADB as for the AcLysine cytoblot. After aspirating, 20 μl/well of ADB containing anti-acetylated histone H3 antibody (rabbit, 1:500) or anti-acetylated histone H4 antibody (rabbit, 1:500) was added and incubated for 2 hr at room temperature. Wells were aspirated and washed twice with 80 μl of ADB. After aspirating, 20 μl/ well of ADB containing anti-rabbit IgG-HRP conjugated antibody (1:1000) was added and incubated for 2 hr (room temperature). For the AcTubulin and AcLysine assays, plates were washed three times with TBS before adding 20 μl/well of enhanced chemiluminescent (ECL) mixture. For the AcHistH3 and AcHistH4 assays, plates were washed three times with 80 μl TBS with 1% Triton X-100 and once with 80 μl TBS before adding 20 μl/well of ECL mixture. Plates were read on an Analyst plate reader (LJL Biosystems) with an integration time of 0.1 s.

Chemical Library and Retesting of Positives

1,3-dioxanes from single polystyrene macrobeads were prepared as 1–2 mM stock solutions in DMSO as described [23, 28]. Bead decoding and structure determination were as described [23, 29]. Small molecules (~2–5 μM) of interest were retested using a sample from the plate of stock solutions by fluorescence microscopy and comparison to the effects of trichostatin (300 nM to 1 μM).

Immunofluorescence

A549 cells were seeded at a density of 50,000/ 200 μl in $DMEM^+$ on the surface of glass coverslips on top of Parafilm in a 10 cm dish and allowed to attach overnight. For detection of acetylated

α-tubulin, cells were permeabilized and fixed with 100 mM K-PIPES (pH 6.8), 10 mM EGTA, 1 mM $MgCl_2$, 0.2% Triton X-100, 0.2% glutaraldehyde for 15 min.Coverslips were aspirated and excess glutaraldehyde was quenched with sodium borohydride (10 mg/m) for 10 min. After washing twice with ADB, coverslips were blocked for 10 min before adding anti-acetylated α-tubulin antibody (6-11B-1, mouse; 1:500) in ADB for 1–2 hr. For detection of acetylated lysine, cells were fixed in 200 μl of fixation solution (95% ethanol/5% acetic acid, −20°C) for 5 min at room temperature. Coverslips were washed three times with ADB and blocked 10 min in ADB. After aspirating, 50 μl/coverslip of ADB containing anti-acetylated lysine (rabbit; 1:375) in ADB was added and incubated for 1–2 hr at room temperature. For detecting both acetylated α-tubulin and acetylated lysine, the coverslips were washed three times with ADB and incubated with Alexa 594 and Alexa 488-conjugated anti-mouse IgG (1:500) and anti-rabbit IgG (1:500) antibodies along with a nuclear counter stain of Hoechst 33342 (1 μg/ml) for 1–2 hr at room temperature. Coverslips were washed three times in TBS-Triton X-100 (0.1%) and mounted in 20 mM Tris (pH 8.8), 90% glycerol containing 0.5% *p*-phenylenediamine and mounted. Images were collected on a Zeiss LSM510 confocal scanning laser microscope at the appropriate wavelengths using the accompanying software and processed with Adobe Photoshop.

Data Analysis

Raw data files from the Analyst were imported into Excel (Microsoft). Plate values were first standardized by dividing by the mean (n = 16) of the DMSO control on each plate. An average value from both replicates was used as the measure of a small molecule's activity. Testing of various transforms indicated that a Log_2-transformation reduced the skewness and kurtosis of the data. After Log_2-transformation, the data were each fit to a normal distribution to create a normalized acetylation value. All statistical tests, distribution properties, hierarchical clustering, and principal component analysis were performed using XLSTAT-PRO (v5.2). Principal component analysis (PCA) consists of a linear transformation of the original system of axes formed by the *n*-dimensions of the chemical genetic data matrix [34, 35]. This transformation is in the form of a rotation, which preserves Euclidean distances. The directions of rotation are determined by considering the standardized covariance matrix Σ_{bio} as a linear operator and computing a set of eigenvectors and corresponding eigenvalues that satisfy the eigenvalue equation: $\Sigma_{bio}\Psi = \Lambda\Psi$. The resulting eigenvectors (Ψ_n) of the matrix Ψ form a set of new, linearly independent, orthogonal axes, called principal components, each of which accounts for successive directions in the *n*-dimensional ellipsoid spanning the multivariate distribution of the original data. The corresponding eigenvalues ($\lambda 1$–$\lambda 5$) of matrix Λ account for progressively smaller fractions of the total variance in the original data (Figure 5D). The full set of raw screening data can be found in Supplemental Table S1 (see http://www.chembiol.com/cgi/content/full/10/5/383/DC1 or write to chembiol@cell.com for a PDF). Adjacency matrices were constructed using Notepad (Microsoft), and graphs were drawn and analyzed using Pajek (v0.72; see http://vlado.fmf.uni-lj.si/pub/networks/pajek/).

Acknowledgments

We are grateful to Jim Follen for assistance with robotics, Jason Gatlin for bead decoding, and Dr. Timothy Mitchison, Dr. Rebecca Ward, Marty Burke, Dr. Greg Copeland, and Dr. Paul Clemons for helpful discussions. Dr. Scott Sternson is acknowledged for his contribution to the design of the 1,3-dioxane library as HDAC inhibitors. S.L.S. is an Investigator at the Howard Hughes Medical Institute. We thank the National Institute for General Medical Sciences for support of this research; the National Cancer Institute (NCI), Merck KGaA, Merck & Co., and the Keck Foundation for support of the ICCB; and the NCI for support of the Initiative for Chemical Genetics. K.M.K. was supported by a Damon Runyon Cancer Research Foundation Fellowship (DRG-1650).

Received: January 24, 2003
Revised: March 27, 2003
Accepted: March 28, 2003
Published: May 16, 2003

References

1. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., and Barabasi, A.L. (2000). The large scale organization of metabolic networks. Nature *407*, 651–654.
2. Albert, R., Jeong, H., and Barabasi, A.L. (2000). Error and attack tolerance of complex networks. Nature *406*, 378–382.
3. Maslov, S., and Sneppen, K. (2002). Specificity and stability in topology of protein networks. Science *296*, 910–913.
4. Mitchison, T.J. (1994). Towards a pharmacological genetics. Chem. Biol. *1*, 3–6.
5. Schreiber, S.L., and Bernstein, B.E. (2002). Signaling network model of chromatin. Cell *111*, 771–778.
6. Stockwell, B.R. (2000). Chemical genetics: ligand-based discovery of gene function. Nat. Rev. Genet. *1*, 116–125.
7. Specht, K.M., and Shokat, K.M. (2002). The emerging power of chemical genetics. Curr. Opin. Cell Biol. *14*, 155–159.
8. Schreiber, S.L. (2000). Target-oriented and diversity-oriented organic synthesis in drug discovery. Science *287*, 1964–1969.
9. Haggarty, S.J., Mayer, T.U., Miyamoto, D.T., Fathi, R., King, R.W., Mitchison, T.J., and Schreiber, S.L. (2000). Dissecting cellular processes using small molecules: identification of colchicine-like, taxol-like and other small molecules that perturb mitosis. Chem. Biol. *7*, 275–286.
10. Schreiber, S.L. (2003). Chemical genetics. Chem. Eng. News *81*, 51–61.
11. Stockwell, B.R., Haggarty, S.J., and Schreiber, S.L. (1999). High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. Chem. Biol. *6*, 71–83.
12. Boffa, L.C., Vidali, G., Mann, R.S., and Allfrey, V.G. (1978). Suppression of histone deacetylation in vivo and in vitro by sodium butyrate. J. Biol. Chem. *253*, 3364–3366.
13. Yoshida, M., Kikima, M., Akita, M., and Beppu, T. (1990). Potent and specific inhibition of mammalian histone deacetylasse both in vivo and in vitro by trichostatin A. J. Biol. Chem. *265*, 17174–17179.
14. Kijima, M., Yoshida, M., Sugita, K., Horinouchi, S., and Beppu, T. (1993). Trapoxin, an antitumor cyclic tetrapeptide is an irreversible inhibitor of mammalian histone deacetylases. J. Biol. Chem. *268*, 22429–22435.
15. Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator rpd3p. Science *272*, 408–411.
16. Grozinger, C.M., and Schreiber, S.L. (2002). Deacetylase enzymes: biological functions and the use of small molecule inhibitors. Chem. Biol. *9*, 3–16.
17. Polevoda, B., and Sherman, F. (2002). The diversity of acetylated proteins. Genome Biol. *3*, reviews0006.1-0006.6.
18. Khochbin, S., Verdel, A., Lemercier, C., and Seigneurin-Berny, D. (2001). Functional significance of histone deacetylase diversity. Curr. Opin. Genet. Dev. *11*, 162–166.
19. Remiszewski, S.W. (2002). Recent advances in the discovery of small molecule histone deacetylase, inhibitors. Curr. Opin. Drug Discov. Devel. *5*, 487–499.
20. Johnstone, R.W. (2002). Histone-deacetylase inhibitors: novel drugs for the treatment of cancer. Nat. Rev. Drug Discov. *1*, 287–299.
21. Koeller, K.M., Haggarty, S.J., Perkins, B.D., Leykin, I., Wong, J.C., Kao, M.-C.J., and Schreiber, S.L. (2003). Chemical genetic modifier screens: small molecule trichostatin suppressors as probes of intracellular histone and tubulin acetylation. Chem Biol, this issue, 397–410.
22. Haggarty, S.J., Koeller, K.M., Wong, J.C., Grozinger, C.M., and Schreiber, S.L. (2003). Domain-selective small molecule inhibitor of HDAC6-mediated tubulin deacetylation. Proc. Natl. Acad. Sci. USA *100*, 4389–4394
23. Hubbert, C., Guardiola, A., Shao, R., Kawaguchi, Y., Ito, A., Nixon, A., Yoshida, M., Wang, X.F., and Yao, T.P. (2002). HDAC6 is a microtubule-associated deacetylase. Nature *417*, 455–458.
24. Wong, J.C., Hong, R., and Schreiber, S.L. (2003). Structural biasing elements for in-cell histone deacetylase paralog selectivity. J. Am. Chem. Soc., in press.
25. Sternson, S.M., Wong, J.C., Grozinger, C.M., and Schreiber, S.L.

(2001). Synthesis of 7200 small molecules based on a substructural analysis of the histone deacetylase inhibitors trichostatin and trapoxin. Org. Lett. *3*, 4239–4242.

26. Finnin, M.S., Donigian, J.R., Cohen, A., Richon, V.M., Rifkind, R.A., Marks, P.A., Breslow, R., and Pavletich, N.P. (1999). Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. Nature *401*, 188–193.
27. Piperno, G., and Fuller, M.T. (1985). Monoclonal antibodies specific for an acetylated form of alpha-tubulin recognize the antigen in cilia and flagella from a variety of organisms. J. Cell Biol. *101*, 2085–2094.
28. Piperno, G., LeDizet, M., and Chang, X.J. (1987). Microtubules containing acetylated alpha-tubulin in mammalian cells in culture. J. Cell Biol. *104*, 289–302.
29. Walling, L.A., Peters, N.R., Horn, E.J., and King, R.W. (2001). New technologies for chemical genetics. J. Cell. Biochem. Suppl. *37*, 7–12.
30. Blackwell, H.E., Perez, L., Stavenger, R.A., Tallarico, J.A., Cope Eatough, E., Foley, M.A., and Schreiber, S.L. (2001). A one-bead, one-stock solution approach to chemical genetics: part 1. Chem. Biol. *8*, 1167–1182.
31. Clemons, P.A., Koehler, A.N., Wagner, B.K., Sprigings, T.G., Spring, D.R., King, R.W., Schreiber, S.L., and Foley, M.A. (2001). A one-bead, one-stock solution approach to chemical genetics: part 2. Chem. Biol. *8*, 1183–1195.
32. Balaban, A.T. (1976). Chemical Applications of Graph Theory (London: Academic Press).
33. Fruchterman, T.M.G., and Reingold, E. (1991). Graph drawing by force-directed placement. Software-Practice and Experience *21*, 1129–1164.
34. Legendre, P., and Legendre, L. (1998). Numerical Ecology–Developments in Environmental Modeling (New York: Elsevier).
35. Hotelling, H. (1931). Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. *24*, 417–441.
36. Morgan, T.H., Sturtevant, A.H., Muller, H.J., and Bridges, C.B. (1915). *The Mechanism of Mendelian Heredity* (New York: Henry Holt and Company).
37. van Osdol, W.W., Myers, T.G., Paull, K.D., Kohn, K.W., and Weinstein, J.N. (1994). Use of the Kohonen self-organizing map to study the mechanisms of action of chemotherapeutic agents. J. Natl. Cancer Inst. *86*, 1853–1859.
38. Weinstein, J.N., Myers, T.G., O'Connor, P.M., Friend, S.H., Fornace, A.J., Kohn, K.W., Fojo, T., Bates, S.E., Rubinstein, L.V., Anderson, N.L., et al. (1997). An information-intensive approach to the molecular pharmacology of cancer. Science *275*, 343–349.
39. Roberge, M., Cinel, B., Anderson, H.J., Lim, L., Jiang, X., Xu, L., Bigg, C.M., Kelly, M.T., and Andersen, R.J. (2000). Cell-based screen for antimitotic agents and identification of analogues of rhizoxin, eleutherobin, and paclitaxel in natural extracts. Cancer Res. *60*, 5052–5058.