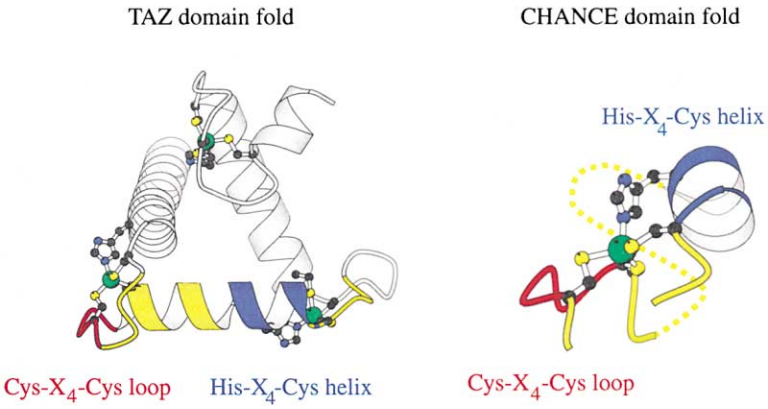

Figure 1. Alternative Zinc Binding Modes

In the TAZ domain structure, the Cys-C4-Cys loop and the His-X4-Cys helix bind to different metal ions. In the CHANCE domain fold, these two substructures come together to form a novel zinc-binding module.

ulatory systems is still in its infancy; however, it is possible that their development is part of the story of the evolution of zinc binding domains.

The concept of mixing domains to generate a new structure is not unique to zinc binding domains. Oligomeric proteins can be derived from monomeric proteins via a "3D domain swapping" mechanism [8]. Typically, a portion of the tertiary structure of one monomeric protein is replaced by the analogous structural element of a second monomeric protein, resulting in a dimer. For example, the removal of part of a surface loop in the monomeric staphylococcal nuclease resulted in a dimeric structure, stabilized by a swapping of the C-terminal α helices of two monomers [9]. As with the generation of novel zinc binding folds by mixing and matching metal binding substructures, 3D domain swapping may provide a route for the evolution of oligomeric proteins from monomers, and it has even been suggested to be a route for the formation of amyloid fibrils associated with neurodegenerative diseases [10]. The results of Mackay and coworkers reveal this process in reverse with the conversion of a single-chain trimeric structure of the TAZ domain into a well-defined folded monomer by truncation.

Sarah L.J. Michel and Jeremy M. Berg
Department of Biophysics and Biophysical Chemistry
Johns Hopkins University School of Medicine
Baltimore, Maryland 21205

Selected Reading

1. Berg, J.M. (1995). Zinc finger domains: from predictions to design. Acc. Chem. Res. *28*, 14–18.
2. Sharpe, B.K., Matthews, J.M., Kwan, A.H.Y., Newton, A., Gell, D.A., Crossley, M., and Mackay, J.P. (2002). A new zinc binding fold underlies the versatility of zinc binding modules in protein evolution. Structure *10*, 639–648.
3. De Guzman, R.N., Liu, H.Y., Martinez-Yamout, M., Dyson, H.J., and Wright, P.E. (2000). Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP. J. Mol. Biol. *303*, 243–253.
4. Michael, S.F., Kilfoil, V.J., Schmidt, M.H., Amann, B.T., and Berg, J.M. (1992). Metal binding and folding properties of a minimalist Cys_2His_2 zinc finger peptide. Proc. Natl. Acad. Sci. USA *89*, 4796–4800.
5. Schwabe, J.W.R., and Klug, A. (1994). Zinc mining for protein domains. Nat. Struct. Biol. *11*, 345–349.
6. Laity, J.H., Lee, B.M., and Wright, P.E. (2001). Zinc finger proteins: new insights into structural and functional diversity. Curr. Opin. Struct. Biol. *11*, 39–46.
7. Keefe, A.D., and Szostak, J.W. (2001). Functional proteins from a random-sequence library. Nature *410*, 715–718.
8. Bennett, M.J., Schlunegger, M.P., and Eisenberg, D. (1995). 3D domain swapping: a mechanism for oligomer assembly. Protein Sci. *4*, 2455–2468.
9. Green, S.M., Gittis, A.G., Meeker, A.K., and Lattman, E.E. (1995). One-step evolution of a dimmer from a monomeric protein. Nat. Struct. Biol. *2*, 746–751.
10. Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. (2001). A domain-swapped RNAse A dimer with implications for amyloid formation. Nat. Struct. Biol. *8*, 211–214.

Chemistry & Biology, Vol. 9, June, 2002, PII S1074-5521(02)00158-8

Dissecting Histone Deacetylase Function

Recent research describes the use of chromatin immunoprecipitation and intergenic chromosomal-DNA microarrays to analyze HDAC function genome-wide. The next step in realizing the full potential of these analyses will be to develop specific and temporal control over HDAC perturbation.

Histone acetylation was initially observed and characterized in the 1960s by Allfrey and colleagues. However, the unexpected catalyst for recent advances in our knowledge of histone deacetylase (HDAC) function was not initiated until 1990, when Itazaki and coworkers at Shionogi Laboratories set out upon a chemical genetic screen to look for small molecules with the ability to

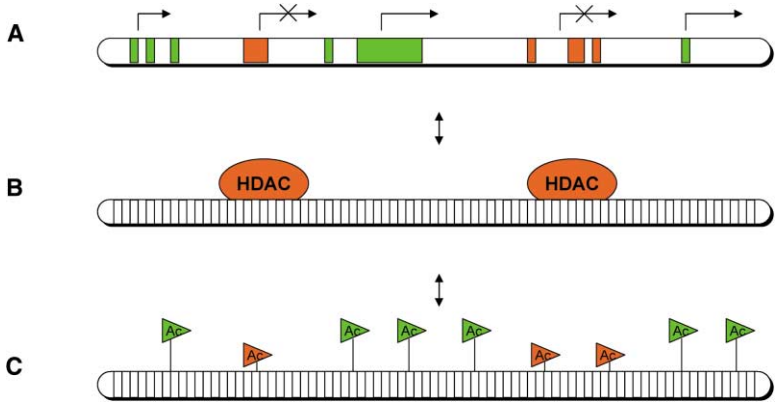

Figure 1. The Interplay between Expression Arrays, Enzyme Binding Arrays, and Acetylation Arrays in the Chromosomal Mapping of Histone Deacetylase Site-of-Action

(A) Expression arrays.
(B) Enzyme binding arrays.
(C) Acetylation arrays. Green acetylation triangles represent areas of high histone acetylation, and red triangles represent areas of low histone acetylation.

morphologically revert v-sis-transformed NIH 3T3 cells from a spindle-like morphological phenotype to the more normal fibroblast-like shape [1]. Without hypothesizing upon the nature and the corresponding protein targets of the bioactive molecules that the screen would uncover, they identified trapoxin, a potent epoxyketone-containing cyclic tetrapeptide. In 1993, Beppu and colleagues would identify another phenotype of trapoxin-treated cells. This time, it was a molecular phenotype; histones extracted from cells treated with trapoxin were globally hyperacetylated [2]. From their new finding, Beppu would demonstrate that trapoxin was capable of inhibiting the biochemical activity of histone deacetylation. It was not until 1996 that Schreiber and colleagues used the same small molecule as the basis for a chemical approach to identify the protein target of trapoxin. This study resulted in the cloning of HDAC1 and, for the first time, physically linked the histone deacetylase enzyme with a known transcription factor [3].

Since the initial identification of the first histone deacetylase enzyme and the subsequent cloning of homologs and orthologs in a number of species, the challenge for the field has been to dissect the function of the individual histone deacetylases. Key questions remain unanswered, including (1) what are the biological substrates of the different enzymes? (2) how are the different enzymes spatially, temporally, and functionally regulated? and (3) in which biological processes and pathways are the different enzymes relevant? In a pre-genomic era, answers to these questions were obtained on an *ad hoc* basis from which generalizations were sought. With the advent of new tools and reagents to broadly interrogate biology on a genome-wide basis, a more fully developed understanding of global histone deacetylase function is emerging.

In a recent issue of *Cell*, Grunstein and coworkers have tried to address question number 3 through the genome-wide analysis of changes in the pattern of histone acetylation observed in strains of *Saccharomyces cerevisiae* lacking individual HDACs [4]. The authors first generated a pool of DNA fragments enriched for their association with specifically acetylated histones by a chromatin immunoprecipitation. Subsequently, by hybridizing these fragments to an intergenic DNA microarray, Grunstein and colleagues produced the first chromosomal map capable of localizing changes in histone acetylation status (acetylation arrays). Although such maps are subject to indirect effects from the time allowed for cells to habituate to the lack of a given gene product, they are nonetheless highly informative. The authors note that the content of these maps is synergistic with other genome-wide maps that measure not acetylation status, but instead histone deacetylase binding sites (enzyme-binding arrays) as well as the transcriptional effects resulting from the loss-of-function of specific histone deacetylases (expression arrays) (Figure 1; [5, 6]). Together, these maps allow the direct correlation of enzyme binding sites, enzyme action, and resultant transcriptional activity and physical mapping of this cumulative information to the genome.

By analyzing all three sources of data, the authors have demonstrated new and interesting roles for the yeast histone deacetylases. The most significant observation may be the elucidation of a role for HDA1 in the regulation of a class of sub-telomeric genes that are normally transcriptionally repressed but become active under stressful environmental conditions, such as growth on non-glucose carbon sources. Although the effect of the cell's carbon source on levels of HDA1 mRNA was first noticed from HDA1 expression arrays, it was the combination of acetylation array data with expression array data that led to the finding and definition of a group of HAST genes (HDA1-affected subtelomeric regions). Although it had previously been known that the Sir2 histone deacetylase was responsible for sub-telomeric transcriptional repression of the most distal 3–4 kb of genes, it now appears that HDA1 assumes responsibility for repression of the HAST genes located approximately 10–25 kb from the telomeres. The fact that a large number of HDA1-sensitive genes are physically clustered together on chromosome ends suggests the possibility that gene expression is regulated by higher-order chromatin structure in part mediated by HDA1-dependent histone deacetylation. In addition to identifying a new role for HDA1, Grunstein and colleagues have also defined roles for Hos1/Hos3 and Hos2 in regulating ribosomal DNA and the genes encoding ribosomal proteins, respectively. It is clearly a testament to the power of the methods available in the postgenomic era that such global analyses are able to reveal protein functional patterns not discernable from a more narrow perspective.

What are the remaining challenges in dissecting histone deacetylase function? Although highly refined tools

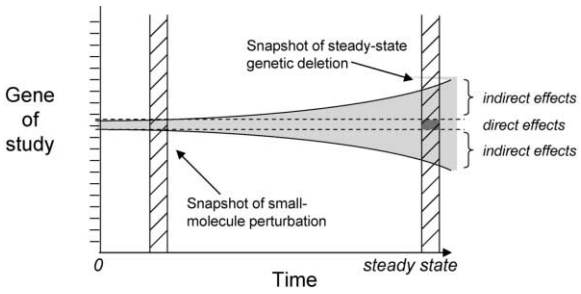

Figure 2. The Analysis of Genetic Perturbations Inherently Measures the Steady-State Condition of a Cell and Hence Allows for Indirect Effects

Small-molecule perturbations, on the other hand, have the element of temporal control and can thus, in principle, dissect biological pathways with greater precision.

for the global analysis of histone-modifying enzymes are now available in the form of enzyme-binding arrays, expression arrays, histone acetylation arrays, and even histone methylation arrays [7], the tools with which the HDACs themselves are perturbed are still relatively crude by comparison. If we are to further our understanding of HDAC function, a method for temporally controlling gene expression is needed to discern the difference between direct, early effects and indirect, late effects (Figure 2). Current methods for genetic perturbation, by definition, directly alter the cell at the level of the DNA. However, because the effects of such perturbations are normally manifest at the protein level, new protein synthesis is necessary; consequently, there is enough time for indirect effects to accumulate before the mutant phenotype can be assessed. Methods that do not require new protein synthesis in order to perturb protein function (e.g., existing temperature-sensitive alleles) require massive perturbations to the cellular environment as a whole. For both gene deletions as well as temperature-sensitive alleles, one is likely to have a mix of both direct and indirect effects in any analysis.

Chemical biologists are now presented with both enormous opportunities and challenges. At present, the most promising method for addressing the issue of temporal protein function perturbation lies in the use of small molecules. The challenge, however, is in developing molecules with the specificity to selectively perturb individual deacetylases. Will chemical biologists be able to generate molecular probes of histone deacetylase function that are both highly potent and specific in order to harness the powerful suite of analytical tools available to measure protein function at the global level? Can chemical biologists not only ultimately dissect the function of the different deacetylase enzymes themselves but also generate the tools to dissect the role of a given deacetylase in different protein complexes (e.g., through the generation of molecules that disrupt protein-protein interactions)? Analogous to a saturating mutagenesis for genetic screens, novel libraries of chemical diversity directed against histone deacetylases will need to be generated and efficiently screened to further dissect histone deacetylase function.

Jeffrey K. Tong
Infinity Pharmaceuticals
Boston, Massachusetts 02118

Selected Reading

1. Itazaki, H., Nagashima, K., Sugita, K., Yoshida, H., Kawamura, Y., Yasuda, Y., Matsumoto, K., Ishii, K., Uotani, N., Nakai, H., et al. (1990). Isolation and structural elucidation of new cyclotetrapeptides, trapoxins A and B, having detransformation activities as antitumor agents. J. Antibiot. (Tokyo) *43*, 1524–1532.
2. Kijima, M., Yoshida, M., Sugita, K., Horinouchi, S., and Beppu, T. (1993). Trapoxin, an antitumor cyclic tetrapeptide, is an irreversible inhibitor of mammalian histone deacetylase. J. Biol. Chem. *268*, 22429–22435.
3. Taunton, J., Hassig, C., and Schreiber, S. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science *272* 408-411.
4. Robyr, D., Suka, Y., Xenarios, I., Kurdistani, S., Wang, A., Suka, N., and Grunstein, M. (2002). Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. Cell *109* 437-446.
5. Bernstein, B., Tong, J., and Schreiber, S. (2000). Genomewide studies of histone deacetylase function in yeast. Proc. Natl. Acad. Sci. USA *97* 13708-13713.
6. Kurdistani, S., Robyr, D., Tavazoie, S., and Grunstein, M. (2002). Genome wide binding map of the RPD3 histone deacetylase in yeast. Nat. Genet., in press.
7. Bernstein, B., Humphrey, E., Erlich, R., Schneider, R., Bouman, P., Liu, J., Kouzarides, T., and Schreiber, S. (2002). Methylation of histone H3 lys 4 in coding regions of active genes. Proc. Natl. Acad. Sci. USA, in press.

Chemistry & Biology, Vol. 9, June, 2002, PII S1074-5521(02)00159-X

Recognizing a Something When Your Library Sees It

Advances are needed in random-display technologies to more tightly link drug actions and functions to the genes that control physiological processes. The reports discussed here explore two sides of these issues—generating new library formats and identifying the targets of drug ligands.

In recent years those of us in the biomedical establishment have been led to believe that more is better when it comes to generating information and data. More data, well, that obviously means more drugs, right? But we learn early in life that a vast amount of information is only useful if you have a good referencing system by