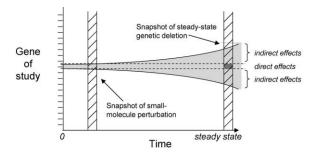Figure 2. The Analysis of Genetic Perturbations Inherently Measures the Steady-State Condition of a Cell and Hence Allows for Indirect Effects

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.

# 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

which you can pluck a signal from the noise. For instance, in one virtual hand we hold the entire genome sequences of many important organisms (even humans!). In the other hand, we either have existing drugs that target specific proteins encoded by those genomes (human or pathogen), or we are seeking new drug candidates to modulate the activity of defined or currently unidentified protein targets. Critically, having a drug in hand is only half of the problem these days—if you don't know the target protein of a bioactive drug, then optimization of drug activity for clinical purposes becomes more onerous for the pharmaceutical chemist. However, even if full knowledge of the target protein is available, side effects that render drugs unusable are often due to action upon additional unknown targets. Thus, the ability to characterize drug-target pairings can be valuable for increasing both efficacy and specificity.

On the other side of the coin we have the issue of creating drugs themselves. Phage display and other library generation technologies have brought "screening" approaches to the research masses. As the ease of generating drug-like candidates has increased, so has the desire to elaborate upon the kinds of constituents from which the libraries are made. What, then, is the best path to (1) develop techniques to enable rapid creation of novel libraries usable by the benchtop research scientist and (2) give researchers a tool that allows them to identify target proteins by defining a mode of drug-target action? In this issue, McPherson et al. and Merryman et al. describe procedures that powerfully address these two points. Merryman and colleagues address the need to directly link gene to function by using modified yeast tRNA, a method that is sufficiently flexible to allow the use of novel chemical groups in peptide side chains. McPherson and colleagues contribute significantly to approaches that could speed up processes by which drug companies might determine relevant target proteins and avoid potentially disastrous side effects.

Ribosome and mRNA display are common features of the approaches described in these reports. These techniques have proven to be of great utility in library generation and library search stratagems over the years. Among the first library-based strategies to gain widespread attention were peptide libraries displayed on phage. Phage display libraries linked the genes carried within the phage to a function—the binding of the peptide displayed on the phage coat to some target moiety. In addition, phage display libraries offered sufficient shape-space complexity to legitimatize investment in their application. But, as we noted above, bigger is often considered better, and thus in a quest for higher complexities in shape-space, researchers have tried to increase the sizes of the libraries to be screened. With in vivo methods such as retroviral libraries or phage display, the complexity is limited by the necessity of cellular transformation after the ligation of the cDNA or peptide insert into the carrier vector. Transformation efficiencies hardly ever gets better than $10^9$. Certainly, library sizes of as few as $10^6$ peptides have shown great successes, but for certain applications this may well be a limitation.

To address this issue, in the last decade several research teams have developed approaches that allow for entirely in vitro-synthesized peptide and polypeptide libraries, physically linked to their encoding nucleic acid. The first such method, termed ribosome display, was described by Mattheakis et al. [1] and was later improved and simplified by Hanes and Plücktun [2]. Roberts and Szostak [3] brought yet another modification that would give rise to mRNA display. In mRNA display, mRNAs encoding randomized peptides or full-length proteins are translated in vitro to yield mRNA-ribosome-peptide complexes. Whereas in ribosome display these three-part complexes are stabilized and directly used in the screen, in mRNA display the mRNA is covalently linked to the peptide only, and this stable conjugate is then used as the screening reagent. The current conjugation method of choice is the addition of a puromycin moiety to the 3′ end of the mRNA by ligation of a DNA-puromycin linker. The ribosome initiates on the mRNA, synthesizes the peptide, and stalls when it encounters the DNA linker. This "pause" allows time for the puromycin to enter the A site of the ribosome (and act as an amino-acylated tRNA mimic) and covalently attach itself to the carboxyl terminus of the peptide or polypeptide chain. In this way, the genetic material encoding the polypeptide is covalently attached to the corresponding polypeptide product. Subsequently, if one can select for a desired peptide function or activity, the corresponding mRNA comes along for the ride and can be characterized by PCR and other standard approaches. Techniques are extant that allow libraries to reach complexities up to $10^{14}$ discrete polypeptide members with high affinity [4, 7].

In their report, Merryman et al. describe a novel approach by which one can attach the mRNA to the synthesized peptide or polypeptide backbone. In this case, instead of using rabbit reticulocyte lysate, they used *E. coli* extracts, which can be easily customized by adding or removing factors. This allowed the use of a novel modified tRNA$^{Phe}$ (tRNA$^x$), which is covalently linked to its amino acid Phenylalanine (Phe), by an amide group that lies in place of the normal amino-acyl linkage. Consequently, during translation, tRNA$^x$ becomes covalently attached to the polypeptide and terminates elongation of the nascent chain. In addition, the natural yeast tRNA$^{Phe}$ contains a unique base, wybutine (Y base), that can be cross-linked to the mRNA near the anticodon by UV treatment. The resultant amide-modified tRNA$^x$ is therefore a bifunctional reagent, enabling the polypeptide to be crosslinked to its corresponding mRNA, with tRNA$^x$ acting as a bridge between the peptide and the nucleic acid. When one mixes different ratios of the tRNA$^x$ to tRNA$^{Phe}$, each time the ribosome reads a Phe codon the modified tRNA$^x$ has a certain chance of being incorporated at the ribosome A site and covalently attaching itself to the growing peptide. Subsequent UV treatment ensures attachment of the tRNA$^x$-peptide to the mRNA.

One caveat of using a modified tRNA$^{Phe}$ as a cross-linking agent is that Phe codons are found throughout mRNA sequences and also within the stretches of random sequence that are an unavoidable result of library production. Early incorporation of tRNA$^x$ rather than tRNA$^{Phe}$ will result in shorter, incomplete peptides that will be less likely to have a high affinity for a target

structure. The functional diversity of such a library could be reduced. To solve this problem, the authors propose simultaneously lowering the effective concentration of the tRNA$^x$ and artificially increasing the representation of the Phe codon after the randomized region of the mRNAs. Despite the fact that this could skew the library toward a higher representation of members that encode Phe at the C terminus, compromise should ensure that the library contains as few prematurely terminated members as possible.

The authors prove the utility of this method as a means of screening for small peptides that have a desired activity by starting with a defined pool of mRNAs and generating an enriched subset encoding peptides that bind a specific matrix. The challenge for the authors is now to attempt a real screen against an unknown target structure by using a larger, fully randomized library. If the authors can accomplish this (and the necessary scale-up for the creation of their tRNA$^x$ substrate), the technique could provide a valuable alternative to the puromycin termination approach currently favored by the majority of researchers as a basis for variations on the mRNA display theme.

After isolating a molecule or a peptide that has a desirable phenotype in vivo, one is often faced with the task of finding the cellular target of the compound. McPherson et al. used an interesting variation of mRNA display to find the target of a known drug. In their study, they successfully selected FKBP12 from a mixture of mRNA libraries from various tissues by using an FK506-biotin conjugate. Thus, they demonstrated that mRNA display can be use to isolate, from a cDNA library, the target of a natural compound. Moreover, by implementing a random priming step in the last round of library generation, the authors demonstrate that this technology can be used to map a minimum site of interaction between the drug and the protein.

Because there are about 30,000 genes in the human genome, the average mRNA should be represented around $10^6$ times in any given library. Even extremely rare target species could be represented multiple times, allowing opportunity for their selection. However, the limitations of this particular technique include the need to modify the drug in such a way that one can attach it to a solid support; such modification could potentially occlude its normal binding properties. Not all drugs are modifiable in this manner (because structural, functional, and chemical analyses are not always available). In addition, the library creation step (transcription and translation) will probably be less efficient for longer proteins, thus diminishing the chances of having a functional protein target. By a similar token, if the target of drug a motif is created during the formation of a multimeric protein complex, this approach will not, in its present incarnation, work. Also, if the binding is dependent on posttranslational modifications, mRNA display won't do the trick because in vitro-translated proteins are devoid of such additions. However, one can take advantage of this weakness. In fact, one can select for proteins bearing specific types of modifications, such as phosphorylation, as long as one can induce the modifications in vitro and specifically fish them out afterwards. A good example of this is the recent paper by Cujec et al., who used mRNA display to find targets of the tyrosine kinase v-Abl by phosphorylating their library with v-Abl and then isolating potential target phosphorylated proteins [5].

The studies discussed in this preview, as well as others dealing with incorporation of nonnatural amino acids [6], constitute important steps in demonstrating that we can fool the translation apparatus and harness its reading/polymerizing capabilities to create highly complex libraries of not only peptides and proteins, but also polymers with potentially tailored pharmacological attributes. Also, the fact that mRNA display is entirely done in vitro makes it more amenable to variation in the screening conditions. A bright future lies ahead for in vitro screening methods such as mRNA display, which can be seen as complementing, or logically following, in vivo, phenotypic screenings.

**Jean-Francois Fortin and Garry P. Nolan**
Department of Microbiology and Immunology
Baxter Laboratory of Genetic Pharmacology
Clinical Sciences Research Center
269 Campus Drive
Stanford School of Medicine
Stanford, California 94305

**Selected Reading**

1. Mattheakis, L.C., Bhatt, R.R., and Dower, W.J. (1994). An in vitro polysome display system for identifying ligands from very large peptide libraries. Proc. Natl. Acad. Sci. USA *91*, 9022–9026.
2. Hanes, J., and Pluckthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. Proc. Natl. Acad. Sci. USA *94*, 4937–4942.
3. Roberts, R.W., and Szostak, J.W. (1997). RNA-peptide fusions for the in vitro selection of peptides and proteins. Proc. Natl. Acad. Sci. USA *94*, 12297–12302.
4. Amstutz, P., Forrer, P., Zahnd, C., and Pluckthun, A. (2001). In vitro display technologies: novel developments and applications. Curr. Opin. Biotechnol. *12*, 400–405.
5. Cujec, T.P., Medeiros, P.F., Hammond, P., Rise, C., and Kreider, B.L. (2002). Selection of v-Abl tyrosine kinase substrate sequences from randomized peptide and cellular proteomic libraries using mRNA display. Chem. Biol. *9*, 253–264.
6. Hohsaka, T., Kajihara, D., Ashizuka, Y., Murakami, H., and Sisido, M. (1999). Efficient incorporation of nonnatural amino acids with large aromatic groups into streptavidin in in vitro protein synthesizing systems. J. Am. Chem. Soc. *121*, 34–40.
7. Wilson, D.S., Keefe, A.D., and Szostak, J.W. (2001). The use of mRNA display to select high-affinity protein-binding peptides. Proc. Natl. Acad. Sci. USA *98*, 3750–3755.