[Frontiers in Bioscience 8, d603-619, May 1, 2003]

RETROVIRAL TECHNOLOGY – APPLICATIONS FOR EXPRESSED PEPTIDE LIBRARIES

Roland Wolkowicz, and Garry P. Nolan

Department of Microbiology and Immunology, Baxter Laboratory in Genetic Pharmacology, Stanford University Medical Center, Stanford, California

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. In the beginning: The first phage display
4. The philosophy behind libraries: from cDNA to peptides
5. Current application of peptide phage display libraries
6. A matter of shape — constrained and linear
7. Retroviruses as a means to deliver genetic information for genetic screens in mammalian cells
8. Retrovirus production from retroviral vectors- a tutorial
9. Construction of the peptide library—a tutorial
10. Retroviral expression peptide libraries: from infection to rescue
11. Current application of retroviral cDNA and peptide libraries
12. Peptide libraries—associations with classic and new technologies
13. Perspective
14. References

1. ABSTRACT

The cell is an extremely complex network of interactions between large numbers of molecules. Understanding this entire network and the information arising from it is an overwhelming and challenging task. Reverse genetics has given us the possibility to discover unknown interactions and their related pathways. With the help of peptide libraries, interactions between biomolecules can be disrupted or distorted and the signaling pathways where these proteins are involved, altered. Consequently, novel biological pathways can be discerned.

The peptide libraries become a pool of shapes, some of them might behave as dominant effectors. With the use of retroviral transfer vectors those libraries can be expressed in a stable manner in the mammalian cell. A strong selection and screening process can finally lead to specific peptides. Novel high-throughput approaches might allow for the rapid creation of small-molecule switches in protein-protein interactions. Reverse genetics and as such the expression of small molecules that will have a specific biological outcome, can become an answer to our queries.

2. INTRODUCTION

The information arising from a range of genomic approaches begs the question of how to determine function among the correlations of generated gene expression sets. Although the correlations observed are a glimpse of the potential that genomics offers for therapeutic benefit, the task of determining function still rests with classical genetic and biochemical determination. In addition functional receptor-ligand, substrate-enzyme interaction, signal transduction pathways in general are interactions between biomolecules. These interactions occur through specific motifs contained within such molecules. Determining the nature of these motifs, such as Src-homology two or three (SH2 or SH3) docking domains, zinc fingers, or specific phosphorylation sites, hetero- or homodimerization surfaces has furthered our understanding of biological pathways. Thus, as more effort is applied to decipher the interplay between such molecules the importance of creating tools that effect such interactions becomes more important.

Genetics and genetic selection had given us invaluable tools for the discovery of hundreds of genes and their products, and the understanding of signaling pathways within the cell. Knocking-out genes by direct or indirect approaches, or interfering with their products has revealed in the last decades the complicated and intricate networks of protein-protein or protein-nucleic acid interactions that give rise to the complex environments of the cell.

The purpose of this chapter will be to underscore some of the advances that have been made to determine the function of genes in signaling pathways by applying genetic screening approaches using expression libraries. Such libraries, as will be detailed below, can be expressed stably in the mammalian cell. The libraries possibly deliver a broad spectrum of elements, from whole gene products to random small peptides.

Generically speaking, library searches in mammalian cells fall under the following rubric: Somewhere in a large library of different cDNAs, peptides, RNAs, etc. there is at least one member of the library that will have a desirable phenotype when it is expressed in a target cell. That desirable phenotype is defined by the

focus of the investigation. In a suitably large universe of potential library members (whose shapes and or functions differ appreciably enough) there exists that member of the library that can be revealed to have a required function with the right screen. The art, then, lies in defining a screen that rapidly sifts through all the possibilities and points directly to the object of desire.

As the technologies have grown more sophisticated there has been a realization that one can move from *in vitro* systems of library application to *in vivo* applications. We will detail in this chapter some of the recent progress of note in these areas.

3. IN THE BEGINNING: THE FIRST PHAGE DISPLAY

It was in 1985 when G.P. Smith first showed that a foreign peptide could be expressed as a fusion protein on the surface of a filamentous phage of *E. coli*, introducing the term "fusion phage" (1). He fused peptides to the minor coat pIII adsorption protein (the product of gene III) of the filamentous bacteriophage M13, present in five copies on one terminus of the virus. An enrichment of one thousand fold over the background phage could then be achieved by immunoaffinity, using an antibody against the foreign peptide. Parmley and Smith (2) showed that several different epitopes could be expressed at the N terminus of pIII and, more importantly, that a phage bearing one of those epitopes could be easily rescued. Scott and Smith suggested the possibility of expressing large numbers of random peptide sequences that could be exploited to find specific peptides bound by antibodies and target gene products (3).

Some classical examples of the use of phage peptide libraries more than a decade ago include the use of the very similar fd phage. The aim was to search for epitopes that are recognized by a monoclonal antibody against the N terminus of beta-endorphin (4). In another study a peptide library was generated to screen for phages that bind specifically to streptavidin (5). The use of phage display has been since widely spread and later examples will be presented below.

4. THE PHILOSOPHY BEHIND LIBRARIES: FROM cDNA TO PEPTIDES

Classically, a library is a pool of genetic information with the potential for function when expressed in the right environment. cDNA libraries per definition are pools of DNA sequences representing genes of an organism that are capable of expressing proteins. Thus, cDNA libraries or lambda gt11 cDNA protein expression libraries are actually based on the natural genomic information of the cell. The number of different clones is subsequently limited to the number of genes or genes segments in the organism under study. In contrast to those "natural" libraries there are artificial libraries of shapes and motifs, displayed as peptides in phage display, that were proposed as an alternative means to create ligands for target structures. In part this derives from the understanding that expression of dominant-negative proteins, the injection of antibodies against a specific protein, antisense RNA methods, use of drugs that may alter the function of a gene product, expression of nucleic acids aptamers, all might allow for dominant modification of signaling systems to allow either biological understanding or therapeutic intervention.

A cDNA library from a specific cell or tissue is predicted to contain at least one copy of each gene expressed within this cell or tissue. If the library contains large enough fragments, it will hopefully harbor complete, nondisrupted genes. However, other approaches consider the random fragmentation of one gene, family of genes or subset of genes as the basis for the production of a cDNA library. The fragmentation of a gene may reveal the presence of specific domains or motifs that could be characterized by their function or activity. This is the basis for the establishment of genetic suppressor elements (GSE) libraries. These GSE can encode both truncated or mutated proteins and antisense RNA. Subsequently, they can act as dominant effectors (bigger than random peptides) for the former and as inhibitory molecules for the latter. With the right selection process, it will be possible to find those with GSE properties from a random pool of fragments. As we shall see, the same rational approach can be applied for even further smaller fragments or peptides from genomic DNA that may still remain functional.

Other approaches do not take advantage of the natural genomic information of the organism. Instead, they are based on a random set of sequences. Those sequences or peptides, when expressed within the cell, can actually have biological effects (6) and possibly affect signaling pathways in many different ways.

As mentioned, the value of the selection process is now incredibly vital and is the critical process that will give rise to the finding of a specific peptide of interest. A random peptide library can include millions of sequences translated into millions of shapes. Although a peptide may by chance have biological activity by its own and function as a molecular switch, it is more likely that its effect will be basically physical and work by access to protein-protein interfaces. As a small shape, a specific peptide, in contrast to a gene, may not have any function but may, upon binding to target protein surfaces, modify or perturb the functions of those target proteins or other proteins with which it interacts. As such, peptides have been referred to as "perturbagens" by some (7) or more recently dominant effectors (DE) by others (6). A peptide or DE could cause a biological outcome due to its interaction with other molecules.

The disruption of signaling pathways by peptides is expected to be due to their action on the biochemical interactions occurring within the cell. This is in striking contrast with the classical way of mutations that will act on the genetic information of the cell. As a result, peptides can be directly used as baits for finding their counterparts. Importantly, in other words, the peptide libraries can be used to define function.

5. CURRENT APPLICATION OF PEPTIDE PHAGE DISPLAY LIBRARIES

The use of peptide libraries, in their more classical display as a fusion to phage, has and still is being used in many fields of molecular and cellular biology. Here only some of the many possible examples are presented. As one can see, the same principles can be applied to mammalian cell selections.

For the determination of optimal amino acids sequences within phosphorylation sites oriented peptide library approaches have been extensively used. Examples are the proto-oncogene AKT that bears a Ser/Thr protein kinase domain. Using peptide libraries, amino acids N-terminal and C-terminal to the site of phosphorylation have been determined, giving rise to a consensus sequence substrate optimal for phosphorylation. The authors have further investigated the relevance of this approach by screening a lambda GEX phage cDNA library. Substrates found contained indeed a motif similar to that identified by the peptide library screening. This results shows that primary sequence may provide hints about the AKT enzyme-substrate interaction (8).

Similar work was undertaken to define the *in vitro* substrate specificity of the ataxia telangiectasia mutated (ATM) kinase activity (9). In this study, the peptide substrate selectivity of ATM was compared to that of the ATM related kinase DNA-PK. Findings of the peptide library analyzed defined an ATM motif distinct from that of DNA-PK. Again, the finding of the ATM specific sequence was used to search the data base for previously identified targets of ATM phosphorylation, as well as, and more important, novel candidates.

Another broadly studied protein, or more accurately, family of proteins, is the protein kinase C (PKC) family, whose members play significant roles in many intracellular signal transduction processes. Identification of their biological partners is currently under investigation. Still, little is known about the substrate specificity of each of the PKC members. Again, an oriented peptide library has come to help in order to find optimal peptide substrate sequences for each of the nine studied (currently eleven are known) human PKC isozymes (10). Phosphorylation sequences substrates as well as core binding residues were characterized. Subsequently, synthetic peptides were prepared based on the predicted optimal sequences for some of the PKC members. And indeed, individual PKC isozymes seemed to have distinct optimal substrates. Moreover, analysis of the residues predicted to form the catalytic cleft could become very helpful in discerning the structural basis for PKC isozymes selectivity (10). In the aforementioned study, an oriented library of more than two and a half billion peptide substrates was used (11). This approach is based on ways to determine SH2-binding sites sequences, where a less random, fixed residue sequence library as a phosphopeptide library is used (12). The consensus sequence of optimal substrates is determined by sequencing the mixture of products generated during a brief reaction with the kinase of interest. The optimal sequences for protein kinase substrates could be determined, assuming that one could quantitatively separate the phosphopeptide products from the bulk of nonphosphorylated peptides. This technique predicts an optimal sequence and provides information about the relative importance of each position for selectivity. The approach makes it possible to resolve the site for catalytic recognition of enzymes.

Finding the right phosphorylation substrate can, in addition, be helpful in identifying potential kinases targets. Still, mutations that occur in those sequences *in vivo* are providing some hints on how cellular transformation occurs. For instance, the cytosolic tyrosine kinases preferentially phosphorylate peptides recognized by their own SH2 or closely related domains, whereas receptor tyrosine kinases preferentially phosphorylate peptides recognized by subsets of group III SH2 domains (13). A point mutation in the RET receptor-type tyrosine kinase, which causes multiple endocrine neoplasia type 2B, resulted in a shift in peptide substrate specificity. Other, maybe more sophisticated, combinatorial peptide libraries, includes the one used to find the substrate specificity of cAMP- and cGMP-dependent protein kinases (PKA and PKG) exploiting acetyl-coupled peptide libraries on cellulose paper. The library is subjected to phosphorylation by the kinases and the resulting sublibraries are further investigated for kinetic data. The recognition motif of PKA was confirmed by this approach, and a novel substrate sequence for PKG was identified (14).

The above examples show the enormous potential of the use of peptide libraries for finding downstream target substrates in gene-of interest-directed signaling pathways. Considering a signaling pathway a cascade of interaction events that give rise to a flow of biological information, the use of peptide libraries to study crossing points in those pathways might be of extreme importance. Docking proteins, as natural protein scaffolds that interact with a number of different proteins via different surfaces, are therefore interesting target proteins for the use of peptide libraries.

As an example, the 14-3-3 docking proteins seem to regulate diverse biological processes as cell cycle checkpoint controls, proliferation, or apoptosis, through interaction with different proteins. Using random peptide phage display, a peptide with high affinity to 14-3-3 was found. This peptide inhibited the interaction with Raf-1, a main substrate of 14-3-3, and was able to diminish its protective role against phosphatase-induced Raf-1 inactivation (15). This is an example of how candidate antagonists can teach us about the biological functions of a protein and the processes were it is involved in. Peptide libraries have an enormous value in finding downstream target substrates in gene-of interest-directed signaling pathways. And as members in such signaling pathways, the search of proteins or peptides that can block them could be as important.

Peptide libraries have also been used in the search for inhibitors of enzymatic activities. After the selection of

enzymes as targets for phage display, peptides that bound specifically to those enzymes have been isolated. Analysis and competition assays subsequently indicated that peptides screened with each target had similar amino acid sequences and bound to one or two sites per target. Peptides identified by phage display and specific for *Haemophilus influenzae* tyrosyl-tRNA synthetase were analyzed, showing inhibitory properties. The peptidic decoys were indeed preferentially targeted to specific sites that inhibit enzyme function (16).

With the use of peptide libraries, competition between several proteins for the actual binding site could bring to the discovery of novel binding partners or novel binding pockets in a specific protein. A random phage display library was used to identify peptides that bind to laminin-1 and further eluted with heparan sulfate or peptide 11 (17). Three classes of mimetopes for different regions of the laminin binding protein (LBP) were recovered. These regions were a palindromic sequence; the G peptide, a predicted helical domain, and TEDWS-containing C-terminal repeats. All elution conditions also yielded phage with putative heparin binding sequences. The predicted helical structure of the LBP domain under study likely contains two surfaces, a heparin-binding one on one side of the helix, and a peptide 11 hydrophobic region on the other. The QPATEDWSA peptide could inhibit tumor cell adhesion to laminin-1. These data supported that LBP can bind the beta-1 laminin chain at the peptide 11 region, and suggested that heparin sulfate might be an alternate ligand.

Peptide libraries have also been exploited in the HIV-1 field. A study was undertaken to find peptides from a peptide library display approach that specifically bind to the virion-associated protein Vpr. A motif present in the uracyl DNA glycosylase enzyme (UDG) seemed to be incorporated into the virion via its interaction with Vpr (18).

Peptide libraries are used to find sequences with higher affinity to their binding partners than the natural occurring amino acid stretch. For the determination of the critical residues within the turn regions of the beta-lactamase inhibitory protein (betaLIP) that are critical for binding TEM-1 beta-lactamase, phage display has been used (19). Some mutant forms of the inhibitory protein were found to have higher affinity to lactamase than the natural inhibitor.

A more recent approach of display has been the use flagella display. Although the most frequent application of this method has been used for the production of novel recombinant vaccines, it has also been used in peptide display as an alternative to the more classical phage display technology (20). This technique is based on the fusion of peptides into the surface-exposed region of flagellin. Because flagellin; the major subunit of the flagellum, is present in thousands of copies, the level of expressing hybrid flagella bearing the heterologous peptides, can be highly polyvalent. The display on flagellum has been also exploited for constrained libraries. A basic example includes fusion into a disulfide loop of *E. coli* thioredoxin that has been inserted into flagellin, enabling expression of random peptides in a conformation-constrained manner (21, 22). This random peptide library has been applied in antibody epitope mapping and is suitable for biopanning procedures in the study of ligand-receptor interactions.

6. A MATTER OF SHAPE — CONSTRAINED AND LINEAR

After it was understood that linear peptides could be effectively displayed and screened on the surface of phage the question arose as to the limitations of linear peptides. Linear peptides are rotationally flexible. Therefore their free energy of binding is greater and would be expected to have, on average, a lower affinity for a target molecule than would a rigid molecule. How do polypeptides achieve rigid structures? In general, this is accomplished using intramolecular interactions, either through sufficiently large surfaces (coiled-coils, hydrophobic or charge-charge surfaces) or through covalent interactions (disulfide bonds). For a small peptide aptamer the range of possible configurations is currently limited: disulfide bonds are labile under varying physiological conditions and are not suitable for expression in the interior of eukaryotic cells.

Other methods of presenting peptides have, however, been devised. They range from small self-annealing interfaces to larger protein scaffolds. The choice of scaffold depends upon the outcome desired. Some would expect that larger protein scaffolds would be more stable under a variety of circumstances. Others would suggest that small peptide scaffolds are better since they could allow for the random peptide to access more interesting nooks and crannies of molecular interfaces. A scaffold, per definition, will retain its structure or conformation following insertion or fusion of the peptide, with hopefully no further disruption of its function.

The bottom line is that one system, we believe, is not necessarily better than any other and that to achieve the range of outcomes desired a variety of library structures should be applied. The range of possible manners by which to present peptides is significant, and we list a few representatives of the more interesting approaches here.

LaVallie *et al* (23) first described a protein scaffold based on the active-site loop of *E. coli* thioredoxin (TrxA) that is easy to express and purify. As a fusion to flagellin it can be displayed as a scaffold for peptide libraries (21). Colas *et al* (24) used the TrxA scaffold to express a combinatorial library of constrained peptides with the subsequent use of a two-hybrid system to select peptides that bound to human Cdk2. These peptide aptamers were designed to mimic the recognition function of the complementarity-determining regions of immunoglobulins. The aptamers recognized different epitopes on the Cdk2 surface. Furthermore, those tested inhibited Cdk2 activity.

The green fluorescent protein (GFP) from jellyfish has also been extensively exploited as a scaffold for the display

of constrained peptides. GFP has a beta-can structure with several protruding loops that have been useful as sites for the integration of ectopic sequences. Furthermore, its fluorescence property can be easily tracked as a sign for readiness of the protein to support changes. Then different loops of GFP exposed to the solvent were analyzed as candidates for the insertion of a peptide library. Three of those were found to be more suitable for that purpose and did not decrease significantly the fluorescent properties of the protein. This approach actually allowed the screening for sequences expressing correctly folded protein (25). The GFP scaffold has been used in both rational and non-rational approach. The former includes the introduction of a C-terminal fusion of an average of 400 nucleotides sequences from a yeast genomic library. The latter included a 15 amino acid random peptide library introduced within a GFP loop. Both libraries permitted for the screening of peptides or fragments of proteins that could overcome the alpha factor induced cell cycle arrest in *Saccharomyces cerevisiae* (7).

In yet another recent study, the hemagglutinin (HA) epitope tag of influenza was inserted in different loops of GFP to analyze both fluorescence by fluorescence-activated cell sorting (FACS), and accessibility by anti-HA antibodies. Based on the results one of the loops was chosen for the introduction of a 12 or 18-mer peptide library and further compared to a C-terminal 20-mer fusion. As a search for function, several subcellular localization sequences were discovered. As expected, wild type nuclear localization signal (NLS) conferred nuclear localization, although different clones showed different localization patterns (26). This study enhances the importance of choosing not only the right scaffold but also the need to find the best spot within the scaffold to introduce the peptide or library of interest.

The Staphylococcal nuclease shows also some important features of a scaffold (27). It is small, it folds spontaneously and it is easily highly expressed in both prokaryotic and eukaryotic cells. A prominent protruding loop in its structure seems perfectly suitable for the insertion of a peptide library. In a study pursued in yeast cells, such a library was expressed for the selection of inhibitors of both spindle checkpoint and mating pheromone response pathways. Indeed, peptides were found that reduce the Mps1 checkpoint protein level. Others interfered downstream the pheromone receptor-interacting G protein Ste4 and upstream the cyclin-dependent kinase (CDK) inhibitor Far1, a direct inducer of cell cycle arrest (28).

Cell adhesion has become a widely investigated field, specially keeping in mind the importance of homing possible pharmaceutical compounds to specific cell types or/and tissues. The very well characterized RGD sequence has a lot to do with the success in this arena. The RGD sequence first identified as the tetrapeptide cell recognition signal RGDS, was actually determined as part of a larger 30 amino acid sequence (29). This peptide and further smaller peptides (30) were then synthesized and shown to be the minimum signal within fibronectin used as a cell recognition signal that mediates adhesion to the extracellular matrix.

In that same line of research, phage libraries expressing cyclic peptides have been used to select for specific ligands to the cell surface receptors of fibronectin, vitronectin and fibrinogen (31). A mixture of libraries expressing peptides flanked by a cysteine residue on each side (CX5C, CX6C, CX7C) or only on one side (CX9) was used. Many of the integrin-binding sequences derived from the CX9 library contained another cysteine, indicating that the selection gave rise to constrained cyclic peptides. Different ring sizes and flanking residues around the RGD motif were selected as ligands to integrin receptors. Some of the peptides displayed two disulphide bonds (four cysteines) and were shown to be stronger inhibitors of alphav beta5 and alphav beta3-mediated cell attachment to vitronectin than peptides bearing a single disulphide bond and even stronger binders than linear RGD peptides (31). These results emphasize the fact that by introducing more cyclizing bonds it is possible to achieve further structural constrain.

Single immunoglobulin fold-based scaffolds are also considered as molecular constraints due to their globular nature. An example of this constrain is the human cytotoxic T lymphocyte-associated antigen 4 (CTLA-4). A fusion of the extracellular domain of CTLA-4 with the phage coat protein pIII is able to interact with its natural ligands B7-1 and B7-2. Moreover, by replacing the nine amino acid complementarity determining region 3-like loop of this domain with the sequence XXX-RGD-XXX (see below for RGD), several CTLA-4 variants gained the ability to bind the human alpha beta3 integrin, one even *in situ* (32). This data suggested indeed that the CTLA-4 scaffold is suitable for the production of peptide libraries.

The scaffold for a library or a specific sequence can be either from the natural world or artificially engineered. An effective scaffold will keep the functionality of the insert when introduced into it. The leucine zipper, a good example of that approach, has been taken as a structural template for the design of peptides that mimics the function of a larger protein. By introducing four lysines of platelet factor 4 (PF4) into a leucine zipper, it adopted a stable helical conformation with heparin binding activity, typical from PF4 (33). With the aid of computerized modeling others have designed a leucine zipper domain with two helix coiled coil peptide mimetics of interleukin-4 (IL-4). By doing so, the engineered zipper was able to recognize and bind its high affinity receptor (IL-4R-alpha) (34).

Telomeric DNA is broadly under investigation. Proteins that bind to the guanine reach quadruplexes structures of telomers may give some hints about telomere length regulation and senescence and possibly be of therapeutic value. The phage display technology has been used to prepare degenerative libraries based on the three zing fingers of the Zif268 DNA binding domain (DBD). The binding selection process gave rise to clones that specifically recognized both the sequence and structure of

G-quadruplex DNA (35). The zinc finger motif, previously widely used in search of specific fingers that bind double stranded DNA or RNA, is considered an important scaffold due to its modular structure.

These studies demonstrate the importance in finding natural occurring scaffolds, like the leucine zipper and IG-like domains or Zing fingers, the formers for protein-protein interactions and the latter for protein-DNA interaction. These motifs and others waiting to be discovered will be crucial in the design of protein or peptide mimetics with biological function. When used as scaffolds for peptide libraries, it will be possible to recognize specific set of proteins, or target biological compounds within specific signaling pathways.

In another approach, the peptide EFLIVKS, that has been shown to form dimers in solution (36), has been attached to the amino and carboxyl termini of different peptides. The low self-affinity of the duplicated EFLIVKS sequences is sufficient in a tethered peptide structure to drive the polypeptide into a constrained format. In the example reported an unstructured sequence from the inhibitor loop of barley chymotrypsin inhibitor 2 was inserted between the EFLIVKS peptides. A stable monomeric tertiary structure could then be produced that could further be either destabilized or significantly stabilized by the introduction of point mutations in the dimerizing peptides (37).

A more adventurous approach has been the use of split inteins or internal proteins. Inteins have been considered for the intracellular catalysis of peptide backbone cyclization as a method for generating proteins and small peptides that are stabilized against cellular catabolism (38). The DnaE split intein from *Synechocystis sp*. PCC6803 was used to cyclize the *E. coli* enzyme dihydrofolate reductase and to produce the cyclic, eight-amino acid tyrosinase inhibitor pseudostellarin F in bacteria. This novel approach will enable to generate and screen for backbone cyclic products *in vivo* for the ultimate production of intracellular cyclic peptides.

The search for new scaffolds for the display of peptides or peptide libraries that could be expressed in eukaryotic cells is still a matter of taste and the requirements of the screen. When expressed in a eukaryote, scaffolds from prokaryotes such as thioredoxin will be advantageous in the sense that may be inert to the cell and may not have nonspecific unwanted effects. Conversely, they may have the inconvenience of not being correctly folded or might be too bulky to achieve access to desirable surfaces. On the other hand, eukaryotic scaffolds like small globular domains (Ig-like), docking motifs such as SH2 or SH3, or even Zing fingers, may be expressed as desired, folded correctly, but may have numerous effects on the intricate network of signaling within the cell. This could hamper the detection of real effects of the different peptides of the library members. Small linear peptides can be expressed within mammalian cells and library members can be selected that have effects upon interesting signaling systems (Kinoshita and Nolan, unpublished). So, while the range of possible ways to express peptides is broad, there appear to be few rules left unbroken.

7. RETROVIRUSES AS A MEANS TO DELIVER GENETIC INFORMATION FOR GENETIC SCREENS IN MAMMALIAN CELLS

The examples described above demonstrate the importance of the context in which one displays libraries in order to find targets. The systems can be applied in nearly any signaling system or biological process under study. As seen, the natural environment of the cell will have huge impact on the behavior of intra- or intermolecular interactions. It may support or dismiss the use of different scaffold types.

Technically what arises next is the manner in which one can best deliver to, and subsequently recover, genetic information from a mammalian cell system. Though there are several ways in which one can deliver such information to a cell, the requirements for an efficient genetic screen seem to be matched by only a single class of agents. What are the requirements? First, the DNA must be delivered efficiently to majority of the target cells under conditions that do not unnecessarily perturb the cells during the delivery process. Second, the genetic information must be well expressed or its expression regulated. Third, daughter cells that arise after the selection process must contain the genetic information. Fourth, the genetic information must be rescuable to allow for a transferring of phenotype to confirm the effect of the genetic information on target cells.

Though the options for gene delivery are many the answer comes from the world of retroviruses. Retroviruses have been used for many years as vectors for a variety of gene delivery purposes. (For reviews on the applications of retroviruses please see Ref. (39-41)).

Retroviridae are enveloped plus stranded RNA viruses. Retroviruses start their life cycle by attachment to target cells, entry, partial uncoating, reverse transcription, and then integration into the target cell DNA. After taking up residence in the cell they initiate transcription, undergo post-translational processing of their coat proteins, begin assembly at the surface of the host cell, undergo budding and final maturation into infectious virions. It is primarily their ability to efficiently integrate into host cellular DNA that makes them so attractive for delivery of genes into cells. This allows for the integrated and transferred gene(s) to be stably expressed as part of the chromosomal DNA and inherited from generation to generation. The genetic selection process is essentially based on that property.

Retroviruses are distinguished from most other virus classes by their use of the reverse transcriptase (RT) enzyme, from where their name is derived, that synthesizes DNA from the incoming infectious viral RNA template. During reverse transcription the two RNA strands are converted into one double stranded DNA (dsDNA) molecule known as the provirus (42, 43). As a result of that process, which includes annealing of a tRNA primer,

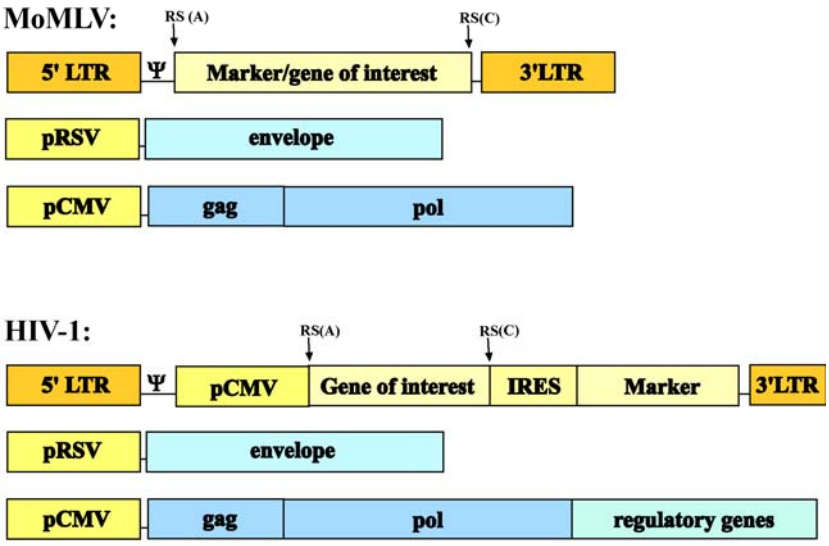

Figure 1. Basic constructs used for retrovirus production. The features depicted are interchangable and are shown just as examples. The 3' LTR in HIV-1 is depicted shorter suggesting a sinLTR (see text). RS: restriction site from the A (A) or C (C) region (see figure 3 and text).

DNA synthesis by RT and strand displacement, two tandem repeats at both 5' and 3' of the dsDNA molecule are created. Those characterize retroviridae and are referred to as the long terminal repeats (LTR). The provirus is then integrated into the genome of the host cell. In the process of transcription, the viral RNA molecules are synthesized by the cellular transcription machinery. Elements within the LTR are mainly responsible for the transcription regulation of the viral genome. The viral RNA will be subsequently used for two main objectives: the production of genomic RNA further incorporated into novel viral particles, and mRNA for the production of viral proteins.
As mentioned, integration is the major reason for the use of retroviral vectors. Within this category, two classes of retroviral vectors are important for consideration. The first are based on the Moloney Murine Leukemia Virus (MoMLV). Additionally, it is possible to deliver genes and libraries to nondividing cells via lentiviral vectors such as those based on human immunodeficiency virus-1 (HIV-1) and feline immunodeficiency virus (FIV) (44, 45). This feature is based on the ability of the preintegration complex of lentivirus to penetrate the nuclear membrane. The development of vectors derived from the retroviral family of viruses enables the delivery of genetic information for genetic screens into mammalian cells. A variety of safe gene delivery platforms that incorporate many novel features of retroviruses in general and lentiviruses in particular have been created. The reader is directed to the extensive literature on the subject, some of it referenced in the next paragraph.

8. RETROVIRUS PRODUCTION FROM RETROVIRAL VECTORS- A TUTORIAL

Mainly, the difference between a wild-type infectious virus and its vector counterpart is that the latter may be rendered replication deficient, and non-pathogenic, by deletion some certain viral functions while maintaining the ability of the vector to transfer genetic information. In most of the broadly applied retroviral gene transfer systems applied to date the genetic information of the virus is divided into more than one vector (Figure 1). This is true for HIV-1 based vectors. In this case, the delivery construct, the structural proteins and the regulatory/accessory genes of HIV-1 are separated onto three different plasmids. The latter, specifically nef, vpr, vpu, vif and tat have been shown to be largely dispensable for gene delivery to most cell types. The relevance of the need for the different proteins of HIV-1, particularly for gene therapy, namely infection efficiency of the target cell and expression level of the transgene, is beyond the scope of this chapter. The genes gag-pol and env are also separated into two independent plasmids. Although the likelihood of recombination events between three independent plasmids giving rise to a replication competent recombinant (RCR) virus is extremely small, ways to increase the safety features of the vectors to decrease even further this probability when dealing with HIV-1 based vectors have been developed. This is the basis for the development of self-inactivating (SIN) vectors (46-51).

Basically, most of the U3 region of the 3'LTR is deleted. This site harbors the major transcriptional functions of the HIV genome. During the process of reverse transcription, the 3'LTR is copied to the 5'LTR. By deleting non-replicative portions of the 3'LTR, the genomic viral DNA is inserted into the target genome as a promoter-less sequence. The lack of active viral promoter avoids both the possible transcription of the viral sequence and detrimental effects on eukaryotic gene expression by insertional effects.

The most commonly used retroviral vector is derived from the MoMLV. The basic system employs a

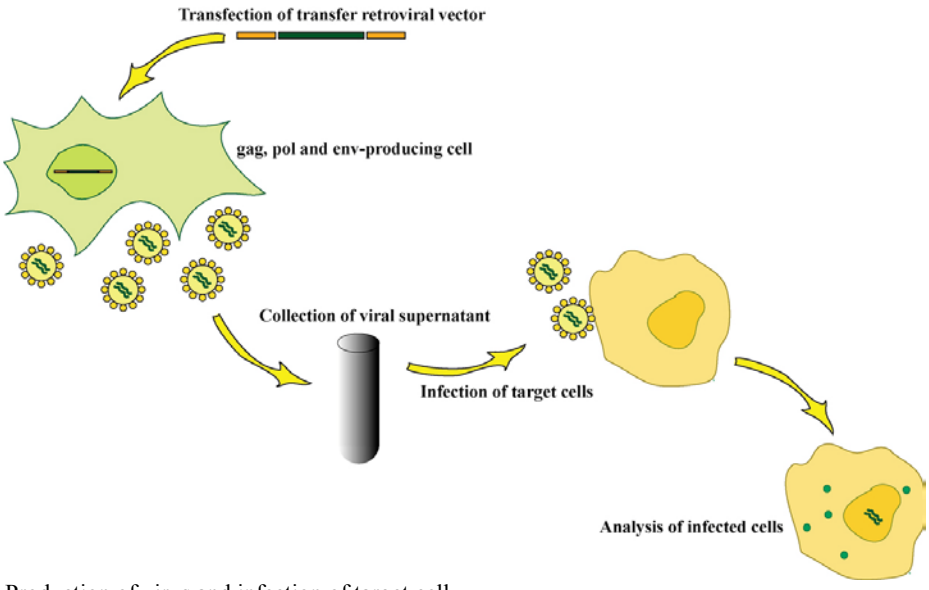

Figure 2. Production of virus and infection of target cell.

For the creation of a competent infectious virion, the three proteins Gag, Pol and Env have to be made in a cell in trans together with the packaging vector. Cell systems have been created, called "packaging cells", in which 293T cells (a highly transfectable cell type) have been engineered to stably express Gag-Pol and Env in trans. The purpose of such cell is to provide the gag, pol and env genes products needed for the complete production of an infective viral particle. Several such cell lines exist, as the Phoenix cells (52). By transfecting the packaging cell with the vector containing the gene of interest, a virus representing the DNA construct will be produced. This virus will subsequently infect different target cells according to the envelope with which it was provided.

In that respect, MoMLV is divided in the ecotropic and amphotropic subgroups according to the receptors their envelopes interact with. The former binds to several membranes spanning transporters for cationic amino acids (53, 54). The latter interact with transporters for inorganic phosphate. Packaging cell lines bearing different envelope genes have been developed, permitting the production of both amphotropic and ecotropic MLV particles, enabling infection of both human and murine cells, respectively Use of different envelopes other than the natural envelope proteins to change the tropism is referred to as pseudotyping the virion core structure with a foreign envelope. For a review see Ref. (55).

After the construction of the retroviral transfer vector carrying the genetic information of interest has been accomplished, the system is ready for the production of the virus. The packaging cells are transfected with the transfer vector (Figure 2). Two days following transfection, the virus-containing supernatant is collected and used to infect the target cells (52). Assuming the production of the virus by the packaging cell was effective and the target cell was infectable, measurable infection efficiency will be possible given the presence of a surrogate marker. A marker such as the green fluorescent protein (GFP), a selectable marker like antibiotic resistance or an envelope receptor is important for the discrimination of peptide library expressing clones versus background. This might be of relevance for the search of a specific biological effect within a huge pool of clones. A positive population (positive for the marker) is a positive population for peptide expression as well, and might request less rounds of enrichment / selection to find the one peptide relevant for the specific biological assay in search.

Surrogate markers that co-express with an insert of interest in one bicistronic element are possible using internal ribosomal entry sites (IRES) (56, 57). IRES sequences act in cis in a cap independent translation mechanism. Translation mediated by the cis-acting element IRES recruits the translational machinery to an internal initiation codon of the surrogate marker (note: the ATG of the initiating codon must be positioned according to the rule(s) of the IRES being used. Since several different IRES sequences from different organisms and genes are in current use the surrogate marker should be encoded and positioned appropriately).

9. CONSTRUCTION OF THE PEPTIDE LIBRARY—A TUTORIAL

Understanding that we can have a retroviral vector that, upon transfection of a packaging cell line, will enable the production of an infectious viral particle we now turn to the details of making such a library suitable for the incorporation into a retroviral vector.

A random peptide library per definition should allow the incorporation of nearly any of the four DNA nucleotides at each position within the sequence. This, in turn, will be translated into a polymer of amino acids

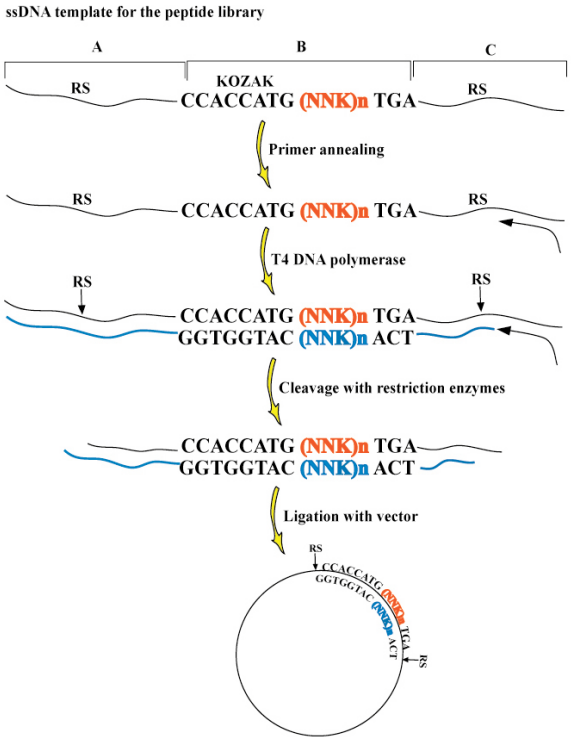

Figure 3. Construction of the peptide library. The template for a linear peptide library is depicted as an example. A: constant region with RS. B: Kozak and ORF for the library. C: As with annealing site for primer. RS: restriction enzyme site.

construct that contains the crucial features for transcription and encapsidation of retroviridae, namely the long terminal repeats (LTRs) and packaging signal necessary for the production and packaging of a competent virion. This is generally referred to as the transfer vector as it usually contains the insert of interest to be transferred. The insert of interest is placed within the vector and will be subsequently transcribed by the 5'LTR of the vector. Alternatively there are designs that allow for an internal promoter that drives transcription of the gene product (see below).

according to the codon usage. But a $(NNN)_n$ sequence (were N represents any nucleotide in the triplet and n the number of codon in the sequence) will not be completely random at the amino acid level as the amino acids used by the random DNA codons are dictated by the translational machinery vis-à-vis the genetic code. By introducing a (NNK)n sequence instead (representing K guanine or thymidine) resulting in the lack of adenine at the third position of the codon, two out of the three eukaryotic stop codons are ruled out. This lowers the possibility for premature termination of the peptide. Moreover, by avoiding both adenine and cytosine at this position, half of the amino acid codon usage is restricted. As a result, the distribution of the different amino acids will be one to three instead of one to six, and thus slightly more uniform. For example, the ratio between leucine or serine to methionine will be three to one instead of six to one.

First, one prepares the DNA template representing the peptide insert. This is accomplished by synthesizing an oligonucleotide that has three regions. Regions A and C (Figure 3) are constant regions containing restriction sites that flank the peptide library sequence to be inserted. Region B contains the Kozak sequence, the peptide insert, and the stop codon. Region C has an extended region that contains a primer annealing site. A primer is prepared complementary to constant region C and polymerization is carried out using T4 DNA polymerase and all four nucleotides. This takes a single-stranded template containing regions A, B, and C to a double stranded form. The dsDNA sequence carrying the random library is then digested with restriction enzymes matched to the retroviral vector into which they are to be inserted. The DNA is cut with appropriate restriction enzymes, gel-purified and then ligated into the vector (which has been cut with matching enzymes). The DNA is then electroporated into suitable competent *E. coli*.

While one clone is sufficient when a single DNA insert is introduced into a vector during general laboratory construction of DNA vectors, the nature of a library by definition requires a very large number of independent clones. This simple fact points to the need of high efficiency in every step of the library production, from synthesis of dsDNA insert to ligation with the vector and transformation into bacteria. Without additional care during the preparation of the library, the outcome may be a qualitatively but not quantitatively perfect library. If the complexity of the library is not large enough to have the statistical chance to contain the sequence of interest, the event we are searching for may not occur. cDNA libraries in general contain around ten million independent clones. Random peptide libraries that have been used successfully contain at least ten times that number: 10^7-10^8 independent clones.

In some cases it is desirable to express the library in the context of what is termed a presentation structure or scaffold. Presentation structures are generally larger proteins that provide either stability or a biological context for the peptide. In such cases some modifications of the vectors are required as well as modifications to the design of the peptide insert. In the case of a fusion protein, the peptide will be introduced in frame to the open reading frame (ORF) of interest. Thus, the library will have a matching set of restriction sites to the ORF into which it is to be inserted. If the peptide is at the N terminus it will contain an ATG but no stop codon (relying on the stop codon of the presentation ORF structure). If the peptide is in the middle of the ORF then the peptide obviously should have no ATG and no stop codon and only matching restriction sites to the point of insertion. Lastly, if the peptide library is to be c-terminal to the ORF then matching sites and only a stop codon within the peptide are required. For reference to a GFP scaffold expressed in mammalian cells where peptides are either displayed within the loops of GFP or at its C terminus, the reader is directed to Peele *et al* (26).

10.RETROVIRAL EXPRESSION PEPTIDE LIBRARIES: FROM INFECTION TO RESCUE

Now that the library is ready, the viral particles harboring the individual members of the library must be made. The following steps will include the choice of cells

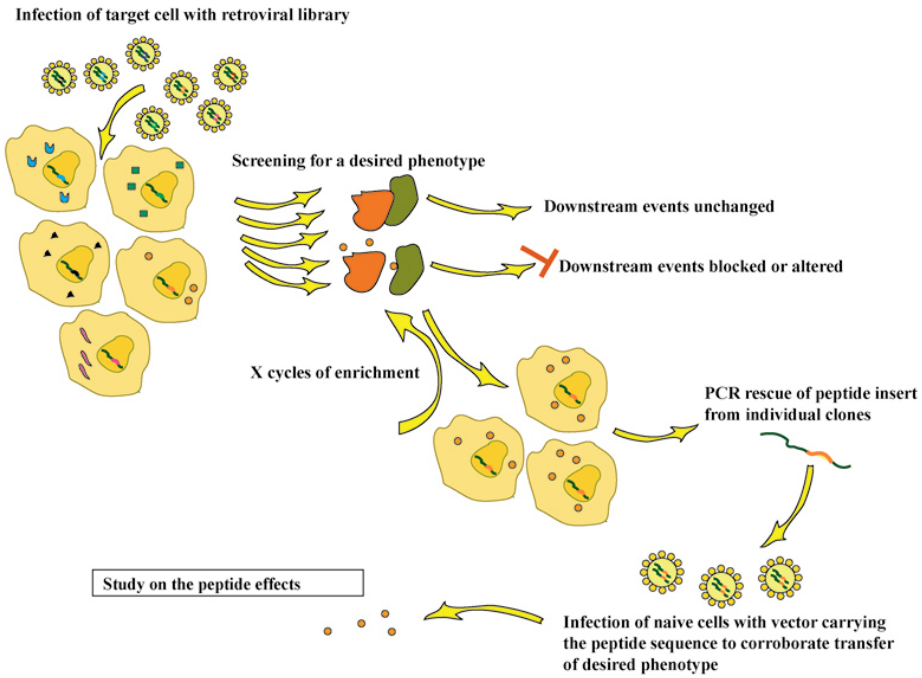

Figure 4. Screening of peptide library.

to infect and of course the screening process of the experiment. For an efficient screening, the selection process is critical. This mainly includes:

- Selection of the cells that consequently will be expected to express the ectopic genetic information.
- Establishment of selection criteria within that cell type. This often requires subcloning of a cell from the original mixed population (if obtained from another laboratory or ATCC it is likely that there is significant variation in the cell population. Subcloning can focus upon a clone within the population that can give rise to a population of cells that has the desired characteristics for a screen).
- Insertion, if needed, of DNA constructs that facilitate a selective phenotype. For instance a promoter element driving a selection marker in the case where peptides that change the regulatory behavior of that promoter are selected.
- Infection of the library and waiting approximately 24-48 hours for peptide expression to occur.
- The selection process itself, which requires a search for particular cells within the population that bear a specific and desired phenotype. This is achieved by cycles of enrichment under a "selective pressure".
- Rescue and analysis of the peptides expressed in these particular clones.
- Transfer of the rescued peptides into "naïve" cells to corroborate their biological effect.

Once the peptide library has been transferred into the target cell via the viral particles, the remaining effort is not distinguishable from any other library screening protocol (Figure 4). Several rounds of selection are often required. The number of those cycles will depend entirely on the screening strategy in the specific experiment and the discrimination between background and real events in each round. The marker will enable one to select firstly the cells infected as well as helping to distinguish non-infected cell cells that might transiently have the phenotype of interest (a form of background). Once selection of apparently positive clones is achieved the next step is the rescue of the peptide sequence expressed in those individual clones. For that purpose, it is of great importance to have considered during the original vector design an easy way to rescue the sequence of the candidate clones. By introducing unique sequences within or close to the edges of the peptide sequences, it is possible to use them as templates for polymerase chain reaction (PCR) rescue, via primer annealing.

In order to make the PCR rescue more efficient it is important to attempt to infect one target cell with not more than a single vector virus. Multiple inserts can complicate recovery and verification. This is accomplished by ensuring that not more than 30% of the cells in the target population are infected as it is generally accepted (though not formally proven) that retroviral infection proceeds randomly and thereby follows Poisson statistics. A 30% infection efficiency would predict that one has a grater than 95% confidence that only a single virion has established a genomic integration in any given cell. This will ensure a clear resolution of the peptide sequence during the PCR genomic rescue.

11. CURRENT APPLICATION OF RETROVIRAL CDNA AND PEPTIDE LIBRARIES

Retroviral technology has been used for several years to transfer cDNA libraries, cDNA fragment libraries,

ribozyme libraries, etc. into different target cells (58-61). And more recently, with the development of peptide libraries, the application of the systems is expanding. One can now see the utility of delivery and expression of a library in a mammalian cell. The cellular microenvironment, indeed, will have a great impact on a small molecule such as a peptide. A traditional example is the ability of cyclic peptides to form a loop only under oxidative conditions where crosslinking occurs between the cysteines present. In an *in vitro* assay such a phage display this will not take place. Examples of retroviral delivery are given below to direct the reader to relevant techniques and methods for comparison and instructive purpose.

In two independent settings, retroviruses expressing cDNA derived from mRNA of human T-cells or interleukin (IL)-3-dependent TF-1 cells were used to infect murine IL-3-dependent Ba/F3 cells. The infected cells were selected for the expression of CD2 using flow cytometry, or the alpha subunit of the human IL-3 receptor (hIL-3Ralpha) by factor-dependent growth. As expected, cDNA of CD2 and hIL-3Ralpha were detected in the target cell. With the help of a selection assay for a specific expression or function, the retroviral expression cloning system has been shown to be an efficient way to transfer low-abundance cDNAs (59).

An approach pioneered by Roninson and colleagues (58) allowing transfer of fragments of cDNA, termed genetic suppressor elements (GSE) has been employed to find dominant fragments of cDNAs with function that is independent of that of the entire cDNA. In this manner functional subdomains within cDNA might be characterized as well as the targets upon which they can act. As a consequence one is actually applying a library of evolved shapes in an environment in which they arose resulting in an enriched interaction library. A GSE library based strategy has been used for the search of resistant clones to cytotoxic anticancer drugs known to act on topoisomerase II (topo II). This resistance seems to be triggered by topo II downregulation (transcription and/or translation). A retroviral library expressing random fragments of human topo II-alpha cDNA was used to isolate GSE that inferred resistance to etoposide. GSEs encoding peptides matching short segments of topo II-alpha or antisense RNA sequences were rescued. The expression of antisense RNA-encoding GSE decreased the cellular expression of topo II-alpha. Moreover, the GSE induced resistance to different topo II drugs but not to unrelated drugs not acting on topo II (58).

In the field of Acquired Immunodeficiency syndrome (AIDS) the GSE technology has been exploited. GSEs based on the genome of HIV-1 were screened for deficient virus production and infectivity. The GSE library was delivered into the target cell via a retroviral vector, combining the retroviral and GSE technologies. The selected GSEs clustered in seven specific regions of the HIV-1 genome, keeping their function intact (62). This approach could shed some light on the molecular biology of many viruses to find novel candidates for the generation of antiviral drugs.

A similar approach was undertaken with random fragments of cDNA from the tumor suppressor gene p53. Again, a retroviral delivery system was used to transduce cells, in that case from rat. Interestingly, the GSEs selected for a p53-related function felt into four categories comprising the three well defined domains of: transcription regulation (class I), DNA-binding (class II) and C terminus (class III), and the 3' untranslated region of the mRNA (class IV). All the GSEs promoted primary cells immortalization. Transformation in collaboration with ras occurred only in the presence of classes I and III. And finally, only class III conferred resistance to etoposide and had a strong inhibitory effect on the transcriptional activity of p53 (63). The above mentioned studies enhance not only the elegance of the GSE approach, but also the ease with which one can apply the retroviral technology to deliver into mammalian cells a variety of functional products.

In a beautiful recent application of that approach full-length death-associated protein kinase (DAPK) cDNA was subjected to partial DNase I in order to obtain fragments ranging between 50 bp and 2 Kb. DAPK is a Ca(+2)/calmodulin-regulated serine/threonine kinase with a multidomain structure that participates in apoptosis. To identify regions critical for its proapoptotic activity a genetic screen was undertaken based on the functional selection of short DAPK-derived fragments that could protect cells from apoptosis. A library of the fragmented DAPK cDNA was transfected into HeLa cells treated with the proapoptotic IFN-gamma. Following survival selection, functional cDNA fragments were recovered and further examined in a secondary screen for their ability to protect cells from DAPK-dependent TNF-alpha induced apoptosis. Four biologically active peptides that mapped to the ankyrin repeats, the "linker" region, the death domain, and the C terminus of DAPK were found (64).

The downside of the GSE strategy is that one cannot be entirely sure that the target upon which the cDNA fragment acts is a natural target. The upside is that this might not be important as one is interested in the effect. Thus, the reader is cautioned not to necessarily link a native biological function to the cDNA from which the fragment is generated (ie, sea polyp cDNA expressed in mammalian cells in a T cell screen could likely generate functional cDNA fragments, but one might be hard-pressed to link a T cell function to the sea polyp cDNA from which the fragment was derived. In this case the fragments might be more expected to be a library of shapes taken out of a context of evolved functional interactions).

GSE has been extended to application in fusion libraries. In that approach, a domain or motif is expressed as a fusion protein and its function compared to the parental protein. The cell cycle inhibitor protein p21, known to arrest cells at the G1/S checkpoint, has been taken as an example. By delivering 24 C-terminal amino acids of p21 fused to the GFP protein the cells were arrested, interestingly not only at the G1/S boundary but also at G2 (65). In another example with primary bone marrow-derived mast cells, the role of kinases in IgE-mediated exocytosis was studied. The cells were transduced with

retroviral vectors encoding kinases inhibitory motifs and analyzed by flow cytometry for their effect on exocytosis. A peptide derived from protein kinase A (PKA) inhibited the stimulatory effect of IgE on mast cell exocytosis (65). Microdomains could harbor novel properties unseen in the parental proteins, a property that can be easily exploited for the pharmacological/therapeutic research.

Retroviral technology has been used to sort trafficking signals in the mammalian cell. Three different combinatorial libraries based on the endoplasmic reticulum (ER) retention/retrieval signals were prepared. They were based on both cytosolic and luminal retention/retrieval ER signals (dilysine -KKXX-COOH and -KDEL-COOH, respectively), and the endocytosis and transGolgi network (TGN) trafficking motifs (tyrosine-based and dileucine). Small alterations in the signals changed their affinity, consequently creating all kinds of trafficking phenotypes, as seen by flow cytometry and quantitative trafficking assays (66).

Ribozyme libraries are also exploited as a tool for functional genomics. Those libraries may include members able to target and cleave specific RNA substrates. Hairpin ribozyme gene libraries with randomized target recognition sequences have been constructed in retroviral vectors. In one study, such a library was used to select for cellular factors involved in the function of the hepatitis C virus (HCV) IRES. Rescued ribozymes inhibited IRES-mediated translation of HCV core protein but did not inhibit cap-dependent protein translation or cell growth. Their targets were identified as gamma subunits of human eukaryotic initiation factors 2B (61). In another study, a selection was carried out in order to find ribozymes that promote cell transformation in tissue culture. Ribozymes rescued in those experiments acted on the gene for telomerase reverse transcriptase (mTERT) (60).

Lastly, peptide libraries have also been delivered by retroviral transfer technology. By studying induced cell death, for example, one can ask if specific clones within a population of cells expressing a stable peptide library can circumvent the deleterious effects of a drug. In a system based in the retroviral delivery of a random peptide library within the EFLIVKS-EFLIVKS scaffold (see above: A matter of shape — constrained and linear) a search was undertaken to find peptides that could confer resistance to Taxol-induced apoptosis in HeLa cells. Some peptides with characterized specific motifs were found. Among them, one was related to ABCB1 upregulation (also known as multiple drug resistance (MDR1) (6).

In an attempt to take the *in vitro* phage peptide display to an *in vivo* setting, phage libraries have been injected into mice (67). This strategy has been applied for the search of organ-selective targeting based on *in vivo* screening of random peptide sequences, The fact that specific tumor cells home to specific organs has told us that tissues may bear unique marker molecules accessible to circulating cells. “Address molecules” on endothelial surfaces have been described for both lymphocyte homing to various lymphoid organs and tissues undergoing inflammation, and for tumor homing to the lungs. By injecting phage libraries intravenously into mice peptides (actually phages) were rescued from specific organs as brain or kidney. One of the peptides recovered could even compete with the homologous phage for binding. Furthermore, when red blood cells coated with the peptide were injected intravenously, they accumulated in the brain. Much of the success of drug/peptide or gene delivery may rest on finding those “address” markers. For that purpose, peptide libraries seem to have come to our aid.

Yeast has not only been the basic tool for classical forward genetics. It seems know to be subject to the most modern technologies. In a nice example of reverse genetics in yeast, a peptide library was expressed in order to explore the complex phenotype of arrest in response to mating pheromone. As mentioned in above studies, this system is well characterized and allows screenings with clear biological outputs. Several aptamers were isolated, rending the yeast organism resistant to □-factor induced arrest. Further identification of the peptides protein targets by a modified two-hybrid based system revealed known proteins as well as Cbk1, a protein previously unknown to have any effect on cell cycle arrest (68).

A peptide aptamer from a combinatorial library was shown to be able to discriminate between different interactions. This peptide bound to cyclin-dependent kinase 2 (Cdk2) and inhibited its kinase activity although with distinct substrate specificities. This is in clear contrast with the natural p21Cip1 inhibitor, which has a general effect. When the peptide was expressed in human cells, the cell cycle G1 phase was retarded. This suggested that the peptide inhibited cell cycle progression by blocking Cdk2 activity, specifically on the substrates needed for the G1/S transition (69). Still, the advantages of peptide libraries expressed in mammalian cells are based on the more natural physiological environment were the peptide is produced. As such, one cannot exclude post-translational modifications of the peptide, as glycosylation or specially crosslinking between cysteines in cyclic peptides. Peptides expressed in the cytosol may have diverse effects on different signaling pathways (70). When introduced in the living cell via retroviral vectors, specific peptides could be screened by the differential biological outcome due to their expression, or, in other words, may change the phenotype of the cell (65).

Intracellular expression of random peptide libraries within mammalian cells has been shown recently to be useful for the study of signal transduction pathways and the search of molecular elements that may perturb them. Peptides that rescued from Taxol-mediated apoptosis were selected in HeLa cells (6). Kinoshita and Nolan have expressed retroviral linear peptide libraries in T cells to define processes that interfere with T cell dependent HIV-1 signaling (manuscript in preparation). Other novel approaches based on the expression of the peptide library as a fusion protein to surface proteins, are also under investigation (Wolkowicz and Nolan, unpublished). Now that the tools are in hand, and suitably harnessed, it is

possible that *in vivo* technologies could have the same impact that phage display has had on *in vitro* studies.

12. PEPTIDE LIBRARIES—ASSOCIATIONS WITH CLASSIC AND NEW TECHNOLOGIES

The creation of the complete protein interaction map of a cell is an ultimate goal in biomedicine. High throughput technologies are becoming increasingly powerful for such purposes—and are likely the only means by which with limited resources we can attempt to accomplish the goal.

Peptides and peptide libraries have been used not only for studying protein-protein interactions, but also for their ability to bind DNA. In a study aimed to discern the DNA-binding characteristics of peptides, peptides ligands from solid-phase library beads were displayed and screened with the use of target DNA-conjugated magnetic beads (71). The analysis of the peptide sequences identified in this study revealed that peptide ligands contained mainly hydrophobic amino acids. Zucconi proposes that domain repertoires with recognition specificities can be produced for a number of different protein recognition studies (72). With combinatorial methods, each amino acid position in the binding site of a given protein recognition domain is exchanged in such a way that thousands of possible combinations are analyzed. Only those combinations that keep the active site are considered for further analysis. From the interaction of this domain repertoire with peptide phage display techniques and from the primary structure revealed in the binding site, one could postulate ways in which proteins interact in the cell.

In a two hybrid system based approach, the search of peptides that interact with a peptide displayed by a library has been recently undertaken. An example includes the use of the lambda repressor model developed by Hu *et al* (73, 74) based on the fact that the repressor is active only upon dimerization of the DBD. A peptide library is fused to one of the DBDs, while the peptide of interest is fused to a second one in an independent vector. Both vectors are used to transform *E. coli* and upon the right selection, a peptide member of the library can be screened via its ability to bind the target peptide, conferring full lambda repressor activity (75). Chemistry and peptide libraries had always gone glove in hand. Many of the peptide libraries are being chemically modified. But, more importantly, rescued peptides from genetic libraries, have been, ultimately, chemically synthesized for further investigation. As a consequence, high throughput technologies and search for peptide analog pharmacophores are closer to achieve their goal.

By chemical means, approaches have been considered to produce peptide analogs of protein ligands. Such studies include the introduction of peptide backbone modifications, constrained amino acids and cyclic structures to obtain highly potent, selective and metabolically stable analogs of those peptides (76). An example of a chemically modified library includes the search for peptides that serve as substrates for protein kinases. Peptide libraries were chemically synthesized by a modified "divide-couple-recombine" strategy. After reaction with the kinase of interest, the most highly phosphorylated substrate selected from the library was identified using on-line liquid chromatography-electrospray mass spectrometry (LC-ESMS). As predicted, the cAMP-dependent protein kinase was shown to preferentially phosphorylate Kemptide. Still, a peptide library based on the v-Src autophosphorylation site served to find motifs identified in several tyrosine kinases as phosphorylation or autophosphorylation site (77). A chemical approach was also undertaken for the search of the different PKC isoforms. Conventional peptide library approaches have not revealed any major distinction between the amino acid consensus sequences of the isoforms. A wide variety of structurally diverse N-appended alcohol-containing residues, including tyrosine, can serve as substrates for specific PKC isoforms, namely alpha, betaII and gamma. Both natural and non-natural residues may be used for the search of active-site substrate specificity. Furthermore, they can be used for the construction of transition-state analogues and suicide substrates (78).

Molecular modeling is being useful in combination with phage display. The description of map motifs or specific sequences in the surface between two molecules can be resolved in that way. An example includes the epitope mapping of the neutralizing antibody against Crotoxin, the presynaptic neurotoxin rattlesnake *Crotalus durissus terrificus* venon. Comparison of two antibodies that compete for binding to the toxin, A-56.36 and CB, showed that they had overlapping binding sites on the toxin. Random phage display libraries were used to screen for peptides that bound these monoclonal antibodies. The peptides selected competed with Crotoxin for mAb binding, moreover, they were able to induce neutralizing antibodies when injected into mice. Molecular modeling helped to decipher the conformational rather than linear nature of the toxin epitope, mimicked by the peptides (79).

Another novel biosynthetic phage display assay has been developed that combines both genetic and chemical manipulations in one protein engineering tool. Non-natural amino acids are introduced into a protein using native chemical ligation together with phage display targeted mutagenesis. An example includes the modification of the protease inhibitor eglin c. The N terminus was chemically synthesized with a non-natural amino acid at position 25. The C terminus was expressed fused to pIII or pVIII of filamentous phage in a phage display assay. The resulting protein was able to refold into its active form (80). This study shows that peptide synthesis and phage display mutagenesis can be used together in order to investigate novel chemically modified amino acid residues that could produce novel proteins with improved biological activities.

13. PERSPECTIVE

The cell is an extremely complex network of interactions between large numbers of molecules. Those cannot be discerned entirely by *in vitro* assays. Dissection

of isolated pathways or interactions between specific partners may help discern a picture of the cellular network. Artificial proteins or peptides could modify known pathways in original ways via novel interactions —thus potentially informing us about the pathway's nature. The importance of finding agents that block the interaction of one partner with another without disrupting its contacts with third party proteins is an important goal in both understanding of signaling pathways and the search for pharmacophores. The classical approach for the use of peptide libraries, based mainly on phage displays, has been an *in vitro* methodology. Nevertheless, it is clear that the natural *milieu* of the cell and the specific microenvironment of the different loci in the cell (membrane, cytosol or organelles) are critical for the appropriate structure and function of a biomolecule. Retroviral technologies have become a standard approach for the expression of cDNAs, cDNA fragments, and peptide libraries within mammalian cells. This will likely give rise to a more realistic understanding of the molecules under investigation and their effect on biological pathways.

In addition, analysis of peptides and protein functions with the new array technologies (81) is transforming the world of classical library screenings. Synthetic peptide arrays are allowing high throughput screenings to find discrete biochemical properties. Large-scale genetic screening is possible today with the production of proteins within the living cell and is being helpful in discerning unknown biological activities. Protein arrays can be used to examine in parallel the functions of thousands of proteins previously known only by their DNA sequence (82).

As can be seen, high-throughput approaches are allowing for the rapid creation of small-molecule switches in protein-protein interactions. Peptides with specific functions are being used in assays to discover organic small molecule shapes that duplicate their effect, or indeed to deliver small molecules to cells via direct binding to drug-like molecules (Rozinov and Nolan, 1999). It seems that we are only limited by our imagination as to how we apply the evolutionary principles described above to the questions that face us. Peptide libraries approaches, including those expressed in mammalian cells, and other library principles will lead to numerous applications in biomedical sciences. Combinatorial peptide libraries together with combinatorial chemistry (83) offers a path where we can acknowledge our lack of understanding of biochemical processes while using a mercenary application of evolution to provide us answers to our queries.

14. REFERENCES

1. Smith, G.P.: Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 228(4705), 1315-7 (1985)
2. Parmley, S.F. & G.P. Smith: Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* 73(2), 305-18 (1988)
3. Scott, J.K. & G.P. Smith: Searching for peptide ligands with an epitope library. *Science* 249(4967), 386-90 (1990)
4. Cwirla, S.E., E.A. Peters, R.W. Barrett & W.J. Dower: Peptides on phage: a vast library of peptides for identifying ligands. *Proc Natl Acad Sci U SA* 87(16), 6378-82 (1990)
5. Devlin, J.J., L.C. Panganiban & P.E. Devlin: Random peptide libraries: a source of specific protein binding molecules. *Science* 249(4967), 404-6 (1990)
6. Xu, X., C. Leo, Y. Jang, E. Chan, D. Padilla, B.C. Huang, T. Lin, T. Gururaja, Y. Hitoshi, J.B. Lorens, D.C. Anderson, B. Sikic, Y. Luo, D.G. Payan & G.P. Nolan: Dominant effector genetics in mammalian cells. *Nat Genet* 27(1), 6-8 (2001)
7. Caponigro, G., i.M.R. Abed, A.P. Hurlburt, A. Maxfield, W. Judd & A. Kamb: Transdominant genetic analysis of a growth control pathway. *Proc Natl Acad Sci U S A* 95(13), 7508-13 (1998)
8. Obata, T., M.B. Yaffe, G.G. Leparc, E.T. Piro, H. Maegawa, A. Kashiwagi, R. Kikkawa & L.C. Cantley: Use of peptide and protein library screening to define optimal substrate motifs for AKT/PKB. *J Biol Chem*, (2000)
9. O'Neill, T., A.J. Dwyer, Y. Ziv, D.W. Chan, S.P. Lees-Miller, R.H. Abraham, J.H. Lai, D. Hill, Y. Shiloh, L.C. Cantley & G.A. Rathbun: Utilization of oriented peptide libraries to identify substrate motifs selected by ATM [In Process Citation]. *J Biol Chem* 275(30), 22719-27 (2000)
10. Nishikawa, K., A. Toker, F.J. Johannes, Z. Songyang & L.C. Cantley: Determination of the specific substrate sequence motifs of protein kinase C isozymes. *J Biol Chem* 272(2), 952-60 (1997)
11. Songyang, Z., S. Blechner, N. Hoagland, M.F. Hoekstra, H. Piwnica-Worms & L.C. Cantley: Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr Biol* 4(11), 973-82 (1994)
12. Songyang, Z., S.E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W.G. Haser, F. King, T. Roberts, S. Ratnofsky & R.J. Lechleider: SH2 domains recognize specific phosphopeptide sequences. *Cell* 72(5), 767-78 (1993)
13. Songyang, Z., K.L. Carraway III, M.J. Eck, S.C. Harrison, R.A. Feldman, M. Mohammadi, J. Schlessinger, S.R. Hubbard, D.P. Smith, C. Eng, M.J. Lorenzo, B.A.J. Ponder, B.J. Mayer & L.C. Cantley: Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* 373(6514), 536-9 (1995)
14. Tegge, W., R. Frank, F. Hofmann & W.R. Dostmann: Determination of cyclic nucleotide-dependent protein kinase substrate specificity by the use of peptide libraries on cellulose paper. *Biochemistry* 34(33), 10569-77 (1995)
15. Wang, B., H. Yang, Y.C. Liu, T. Jelinek, L. Zhang, E. Ruoslahti & H. Fu: Isolation of high-affinity peptide antagonists of 14-3-3 proteins by phage display. *Biochemistry* 38(38), 12499-504 (1999)
16. Hyde-DeRuyscher, R., L.A. Paige, D.J. Christensen, N. Hyde-DeRuyscher, A. Lim, Z.L. Fredericks, J. Kranz, P. Gallant, J. Zhang, S.M. Rocklage, D.M. Fowlkes, P.A. Wendler & P.T. Hamilton: Detection of small-molecule enzyme inhibitors with peptides isolated from phage-displayed combinatorial peptide libraries. *Chem Biol* 7(1), 17-25 (2000)
17. Kazmin, D.A., T.R. Hoyt, L. Taubner, M. Teintze & J.R. Starkey: Phage display mapping for peptide 11 sensitive sequences binding to laminin-1. *J Mol Biol* 298(3), 431-45 (2000)
18. BouHamdan, M., Y. Xue, Y. Baudat, B. Hu, J. Sire, R.J. Pomerantz & L.X. Duan: Diversity of HIV-1 Vpr

interactions involves usage of the WXXF motif of host cell proteins. *J Biol Chem* 273(14), 8009-16 (1998)
19. Huang, W., Z. Zhang & T. Palzkill: Design of potent beta-lactamase inhibitors by phage display of beta-lactamase inhibitory protein. *J Biol Chem* 275(20), 14964-8 (2000)
20. Westerlund-Wikstrom, B.: Peptide display on bacterial flagella: principles and applications. *Int J Med Microbiol* 290(3), 223-230 (2000)
21. Lu, Z., K.S. Murray, V. Van Cleave, E.R. LaVallie, M.L. Stahl & J.M. McCoy: Expression of thioredoxin random peptide libraries on the Escherichia coli cell surface as functional fusions to flagellin: a system designed for exploring protein-protein interactions. *Biotechnology (N Y)* 13(4), 366-372 (1995)
22. Lu, Z., B.C. Tripp & J.M. McCoy: Displaying libraries of conformationally constrained peptides on the surface of Escherichia coli as flagellin fusions. *Methods Mol Biol* 87, 265-280 (1998)
23. LaVallie, E.R., E.A. DiBlasio, S. Kovacic, K.L. Grant, P.F. Schendel & J.M. McCoy: A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. *Biotechnology (N Y)* 11(2), 187-93 (1993)
24. Colas, P., B. Cohen, T. Jessen, I. Grishina, J. McCoy & R. Brent: Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature* 380(6574), 548-50 (1996)
25. Abedi, M.R., G. Caponigro & A. Kamb: Green fluorescent protein as a scaffold for intracellular presentation of peptides. *Nucleic Acids Res* 26(2), 623-30 (1998)
26. Peelle, B., J. Lorens, W. Li, J. Bogenberger, D.G. Payan & D.C. Anderson: Intracellular protein scaffold-mediated display of random peptide libraries for phenotypic screens in mammalian cells. *Chem Biol* 8(5), 521-34 (2001)
27. Serpersu, E.H., D. Shortle & A.S. Mildvan: Kinetic and magnetic resonance studies of active-site mutants of staphylococcal nuclease: factors contributing to catalysis. *Biochemistr* 26(5), 1289-300 (1987)
28. Norman, T.C., D.L. Smith, P.K. Sorger, B.L. Drees , S.M. O'Rourke, T.R. Hughes, C.J. Roberts, S.H. Friend, S. Fields & A.W. Murray: Genetic selection of peptide inhibitors of biological pathways. *Science* 285(5427), 591-5 (1999)
29. Pierschbacher, M., E.G. Hayman & E. Ruoslahti: Synthetic peptide with cell attachment activity of fibronectin. *Proc Natl Acad Sci U S A* 80(5), 1224-7 (1983)
30. Pierschbacher, M.D. & E. Ruoslahti: Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* 309(5963), 30-3 (1984)
31. Koivunen, E., B. Wang & E. Ruoslahti: Phage libraries displaying cyclic peptides with different ring sizes: ligand specificities of the RGD-directed integrins. *Biotechnology (NY)* 13(3), 265-70 (1995)
32. Hufton, S.E., N. van Neer, T. van den Beuken, J. Desmet, E. Sablon & H.R. Hoogenboom: Development and application of cytotoxic T lymphocyte-associated antigen 4 as a protein scaffold for the generation of novel binding ligands. *FEBS Lett* 475(3), 225-31 (2000)
33. Butcher, D.J., M.A. Kowalska, S. Li, Z. Luo, S. Shan, Z. Lu, S. Niewiarowski & Z. Huang: A natural motif approach to protein design: a synthetic leucine zipper peptide mimics the biological function of the platelet factor 4 protein. *FEBS Lett* 409(2), 183-7 (1997)
34. Domingues, H., D. Cregut, W. Sebald, H. Oschkinat & L. Serrano: Rational design of a GCN4-derived mimetic of interleukin-4. *Nat Struct Biol* 6(7), 652-6 (1999)
35. Isalan, M., S.D. Patel, S. Balasubramanian & Y. Choo: Selection of zinc fingers that bind single-stranded telomeric DNA in the G-quadruplex conformation. *Biochemistry* 40(3), 830-6 (2001)
36. Bodenmuller, H., E. Schilling, B. Zachmann & H. Schaller: The neuropeptide head activator loses its biological acitivity by dimerization. *EMBO J* 5(8), 1825-9 (1986)
37. Gururaja, T.L., S. Narasimhamurthy, D.G. Payan & D.C. Anderson: A novel artificial loop scaffold for the noncovalent constraint of peptides. *Chem Biol* 7(7), 515-27 (2000)
38. Scott, C.P., E. Abel-Santos, M. Wall, D.C. Wahnon & S.J. Benkovic: Production of cyclic peptides and proteins *in vivo*. *Proc Natl Acad Sci U S A* 96(24), 13638-43 (1999)
39. Hu, W.S. & V.K. Pathak: Design of retroviral vectors and helper cells for gene therapy. *Pharmacol Rev* 52(4), 493-511 (2000)
40. Trono, D.: Lentiviral vectors: turning a deadly foe into a therapeutic agent. *Gene Ther* 7(1), 20-3 (2000)
41. Vigna, E. & L. Naldini: Lentiviral vectors: excellent tools for experimental gene transfer and promising candidates for gene therapy. *J Gene Med* 2(5), 308-16 (2000)
42. Baltimore, D.: RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226(252), 1209-11 (1970)
43. Temin, H.M. & S. Mizutani: RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226(252), 1211-3 (1970)
44. Buchschacher, G.L.J. & F. Wong-Staal: Development of lentiviral vectors for gene therapy for human diseases. *Blood* 95(8), 2499-504 (2000)
45. Curran, M.A. & G.P. Nolan: Non-primate lentiviral vectors. In: Lentiviral vectors. Eds: D. Trono, Springer-Verlag, NY (2001)
46. Miyoshi, H., U. Blomer, M. Takahashi, F.H. Gage & I.M. Verma: Development of a self-inactivating lentivirus vector. *J Virol* 72(10), 8150-7 (1998)
47. Zufferey, R., T. Dull, R.J. Mandel, A. Bukovsky, D. Quiroz, L. Naldini & D. Trono: Self-inactivating lentivirus vector for safe and efficient *in vivo* gene delivery. *J Virol* 72(12), 9873-80 (1998)
48. Kim, V.N., K. Mitrophanous, S.M. Kingsman & A.J. Kingsman: Minimal requirement for a lentivirus vector based on human immunodeficiency virus type 1. *J Virol* 72(1), 811-6 (1998)
49. Mangeot, P.E., D. Negre, B. Dubois, A.J. Winter, P. Leissner, M. Mehtali, D. Kaiserlian, F.L. Cosset & J.L. Darlix: Development of minimal lentivirus vectors derived from simian immunodeficiency virus (SIVmac251) and their use for gene transfer into human dendritic cells. *J Virol* 74(18), 8307-15 (2000)

50. Deglon, N., J.L. Tseng, J.C. Bensadoun, A.D. Zurn, Y. Arsenijevic, L. Pereira de Almeida, R. Zufferey, D. Trono & P. Aebischer: Self-inactivating lentiviral vectors with enhanced transgene expression as potential gene transfer system in Parkinson's disease. *Hum Gene Ther* 1(11), 179-90 (2000)
51. Xu, K., H. Ma, T.J. McCown, I.M. Verma & T. Kafri: Generation of a stable cell line producing high-titer self-inactivating lentiviral vectors. *Mol Ther* 3(1), 97-104 (2001)
52. Swift, S., J. Lorens, P. Achacoso & G.P. Nolan: Rapid production of retroviruses for efficient gene delivery to mammalian cells using 293T cell-based systems. In: Current Protocols in Immunology. Unit 10.28: Suppl.31 (1999)
53. Kim, J.W., E.I. Closs, L.M. Albritton & J.M. Cunningham: Transport of cationic amino acids by the mouse ecotropic retrovirus receptor. *Nature* 352(6337), 666-7 (1991)
54. Wang, H., M.P. Kavanaugh, R.A. North & D. Kabat: Cell-surface receptor for ecotropic murine retroviruses is a basic amino-acid transporter. *Nature* 352(6337), 729-31 (1991)
55. Buchholz, C.J., J. Stitz & K. Cichutek: Retroviral cell targeting vectors. *Curr Opin Mol Ther* 1(5), 613-21 (1999)
56. Sachs, A.B.: Cell cycle-dependent translation initiation: IRES elements prevail. *Cell* 101(3), 243-5 (2000)
57. Martinez-Salas, E., R. Ramos, E. Lafuente & S. Lopez de Quinto: Functional interactions in internal translation initiation directed by viral and cellular IRES elements. *J Gen Virol* 82(Pt 5), 973-84 (2001)
58. Gudkov, A.V., C.R. Zelnick, A.R. Kazarov, R. Thimmapaya, D.P. Suttle, W.T. Beck & I.B. Roninson: Isolation of genetic suppressor elements, inducing resistance to topoisomerase II-interactive cytotoxic drugs, from human topoisomerase II cDNA. *Proc Natl Acad Sci U S A* 90(8), 3231-5 (1993)
59. Kitamura, T., M. Onishi, S. Kinoshita, A. Shibuya, A. Miyajima & G.P. Nolan: Efficient screening of retroviral cDNA expression libraries. *Proc Natl Acad Sci U S A* 92(20), 9146-50 (1995)
60. Li, Q.X., J.M. Robbins, P.J. Welch, F. Wong-Staal & J.R. Barber: A novel functional genomics approach identifies mTERT as a suppressor of fibroblast transformation. *Nucleic Acids Res* 28(13), 2605-12 (2000)
61. Kruger, M., C. Beger, Q.X. Li, P.J. Welch, R. Tritz, M. Leavitt, J.R. Barber & F. Wong-Staal: Identification of eIF2Bgamma and eIF2gamma as cofactors of hepatitis C virus internal ribosome entry site-mediated translation using a functional genomics approach. *Proc Natl Acad Sci U S A* 97(15), 8566-71 (2000)
62. Dunn, S.J., S.W. Park, V. Sharma, G. Raghu, J.M. Simone, R. Tavassoli, L.M. Young, M.A. Ortega, C.H. Pan, G.J. Alegre, I.B. Roninson, G. Lipkina, A. Dayn & T.A. Holzmayer: Isolation of efficient antivirals: genetic suppressor elements against HIV-1. *Gene Ther* 6(1), 130-7 (1999)
63. Ossovskaya, V.S., I.A. Mazo, M.V. Chernov, O.B. Chernova, Z. Strezoska, R. Kondratov, G.R. Stark, P.M. Chumakov & A.V. Gudkov: Use of genetic suppressor elements to dissect distinct biological effects of separate p53 domains. *Proc Natl Acad Sci U S A* 93(19), 10309-14 (1996)
64. Raveh, T., H. Berissi, M. Eisenstein, T. Spivak & A. Kimchi: A functional genetic screen identifies regions at the C-terminal tail and death-domain of death-associated protein kinase that are critical for its proapoptotic activity. *Proc Natl Acad Sci U S A* 97(4), 1572-7 (2000)
65. Lorens, J.B., M.K. Bennett, D.M. Pearsall, W.R. Throndset, A.B. Rossi, R.J. Armstrong, B.P. Fox, E.H. Chan, Y. Luo, E. Masuda, D.A. Ferrick, D.C. Anderson, D.G. Payan & G.P. Nolan: Retroviral delivery of peptide modulators of cellular functions. *Mol Ther* 5 Pt 1, 438-47 (2000)
66. Zerangue, N., M.J. Malan, S.R. Fried, P.F. Dazin, Y.N. Jan, L.Y. Jan & S. B.: Analysis of endoplasmic reticulum trafficking signals by combinatorial screening in mammalian cells. *Proc Natl Acad Sci U S A* 98(5), 2431-6 (2001)
67. Pasqualini, R. & E. Ruoslahti: Organ targeting *in vivo* using phage display peptide libraries. *Nature* 380(7572), 364-6 (1996)
68. Geyer, C.R., A. Colman-Lerner & R. Brent: "Mutagenesis" by peptide aptamers identifies genetic network members and pathway connections. *Proc Natl Acad Sci USA* 96(15), 8567-8572 (1999)
69. Cohen, B.A., P. Colas & R. Brent: An artificial cell-cycle inhibitor isolated from a combinatorial library. *Proc Natl Acad Sci USA* 95(24), 14272-7 (1998)
70. Souroujon, M.C. & D. Mochly-Rosen: Peptide modulators of protein-protein interactions in intracellular signaling. *Nat Biotechnol* 16(10), 919-24 (1998)
71. Alam, M.R., M. Maeda & S. Sasaki: DNA-binding peptides searched from the solid-phase combinatorial library with the use of the magnetic beads attaching the target duplex DNA. *Bioorg Med Chem* 8(2), 465-73 (2000)
72. Zucconi, A., S. Panni, S. Paoluzi, L. Castagnoli, L. Dente & G. Cesareni: Domain repertoires as a tool to derive protein recognition rules. *FEBS Lett* 480(1), 49-54 (2000)
73. Hu, J., E.K. O'Shea, P.S. Kim & R.T. Sauer: Sequence requirements for coiled-colis: analysis with lambda repressor-GCN4 leucine zipper fusions. *Science* 250(4986), 1400-3 (1990)
74. Hu, J.: Repressor fusions as a tool to study protein-protein interactions. *Structure* 15(3), 431-3 (1995)
75. Zhang, Z., W. Zhu & T. Kodadek: Selection and application of peptide-binding peptides. *Nat Biotechnol* 18(1), 71-4 (2000)
76. Goodman, M., S. Ro, T. Yamazaki, J.R. Spencer, A. Toy, Z. Huang, Y. He & T. Reisine: Topochemical design of bioactive peptides and peptidomimetics. *Bioorg Khim* 18(10-11), 1375-93 (1992)
77. Till, J.H., R.S. Annan, S.A. Carr & W.T. Miller: Use of synthetic peptide libraries and phosphopeptide-selective mass spectrometry to probe protein kinase substrate specificity. *J Biol Chem* 269(10), 7423-8 (1994)
78. Yan, X., K. Curley & D.S. Lawrence: The specificity of the protein kinase C alpha, betaII and gamma isoforms as assessed by an unnatural alcohol-appended peptide library. *Biochem J* 349, 709-15 (2000)
79. Demangel, C., R.C. Maroun, S. Rouyre, C. Bon, J.C. Mazie & V. Choumet: Combining phage display and molecular modeling to map the epitope of a neutralizing antitoxin antibody. *Eur J Biochem* 267(8), 2345-53 (2000)

80. Dwyer, M.A., W. Lu, J.J. Dwyer & A.A. Kossiakoff: Biosynthetic phage display: a novel protein engineering tool combining chemical and genetic diversity. *Chem Biol* 7(4), 263-74 (2000)
81. Young, R.A.: Biomedical discovery with DNA arrays. *Cell* 102(1), 9-15 (2000)
82. Emili, A.Q. & G. Cagney: Large-scale functional analysis using peptide or protein arrays. *Nat Biotechnol* 18(4), 393-7 (2000)
83. Floyd, C.D., C. Leblanc & M. Whittaker: Combinatorial chemistry as a tool for drug discovery. *Prog Med Chem* 36, 91-168 (1999)

Key Words: Genomic Libraries, Peptide, Retrovirus, Screening, Selection, Review

Send correspondence to: Dr. Garry Nolan, Department of Microbiology and Immunology, Baxter Laboratory in Genetic Pharmacology, Stanford University Medical Center, 269 Campus Drive, CCSR3205, Stanford, CA 94305-5175, Tel: 650-725-7002 or 650-723-7875, Fax: 650-723-2383, E-mail: gnolan@stanford.edu