Using peptide aptamers to analyse the proteome

Alejandro Colman-Lerner and Roger Brent

During the 20th century, a great deal of biological knowledge came from the perfection of methods to perform manipulative transmission genetics. During the early 21st century, such methods will be supplemented, and perhaps sometimes supplanted, by dominant 'genetic' methods to probe the function of the proteome.

trends in CELL BIOLOGY

Alejandro Colman-Lerner
The Molecular Sciences Institute, 2168 Shattuck Avenue, Berkeley, CA 94704, USA.
tel: +1 510 647 0690
fax: +1 510 647 0699
e-mail: colman-lerner@molsci.org

Roger Brent
The Molecular Sciences Institute, 2168 Shattuck Avenue, Berkeley, CA 94704, USA.

At the start of the 20th century, the 'rediscovery of Mendel'[1] allowed the consideration of inheritance on a rational, scientific (albeit abstract) basis. During the period 1910–1920, T.H. Morgan and co-workers[2] began to lay the groundwork for the practice of manipulative transmission genetics. This allowed researchers to collect mutants affected in a given genetic process, map those mutations to loci (genes), and further study the effects of combinations of those genes. The eventual linkage of genes to protein products was of course one of the early triumphs of molecular biology. Continuing improvements in transmission genetic methods, and their extension to more and more organisms, arguably accounted for the bulk of the increase in biological understanding during the latter half of the 20th century.

In search of altered phenotypes

Canonically, many classical genetic studies begin with the isolation of numerous mutants that affect a given trait (a 'mutant hunt'), followed by the application of various techniques to give clues to the function of the mutated genes (Fig. 1a). The most informative mutations are typically held to be recessive mutations that result in loss of gene function ('hypomorphs'), with the best of these resulting in complete loss of function ('nulls'). In addition, the ideal mutations produce strong effects on the phenotype and are highly penetrant. In diploid organisms, this sort of analysis works best in organisms when it is easy to generate and examine large numbers of F2 offspring, in which the recessive mutations can be homozygous and their effects made manifest.

Of course, dominant mutations were always isolated too and, although their analysis has never been as easy (and, more importantly, general doctrines for working with them were slow to emerge), these mutations have always supplemented the loss-of-function approach to understanding genetic pathways. Genetic analysis has also, from time to time, been supplemented by the use of chemical inhibitors: in particular, by cell-permeant, small-molecule drugs. As the mechanisms of action of more and more drugs have become known, drugs have come to be perceived as agents that inactivate specific protein targets. Such pharmacological inhibitors are dominant in that they act on the protein even if it is encoded by, for example, both copies of a gene in a diploid organism. More recently, during the past 20 years, antisense RNA and RNA derivatives have been developed[3,4]; these are also dominant.

During the past ten years, several lines of research have led to the development of inactivating proteins expressed from members of combinatorial libraries. It is likely that further developments along these lines will contribute to the rise of a kind of dominant 'genetics' that operates at the level of the expressed proteins of the genome (hereafter, the proteome). During the early 21st century, dominant analysis of the proteome might contribute to biological understanding to the same extent as loss-of-function transmission genetics contributed to biological understanding in the 20th century.

Selection of nucleic acid and protein affinity reagents from combinatorial libraries

One thread leading to dominant protein 'genetics' comes from work by Szostak *et al.*[5] and Gold *et al.*[6] These researchers described the selection of RNAs and, later, DNAs, from pools of random sequences that bound desired ligands. Successive rounds of selection, interspersed with rounds of mutation, resulted in the selection of agents with affinities in the micromolar–nanomolar range. Ellington and Szostak termed these molecules aptamers, from the Greek 'to fit' (the same root as hapten), a term that has entered general use. Nucleic acid aptamers are useful as research reagents that can recognize target proteins *in vitro*, and have the potential to be useful as drugs.

Another thread came from methods that selected from combinatorial libraries proteins that bind targets. The best known examples of these methods are referred to as 'phage display'. In these methods, libraries of *Escherichia coli* bacteriophage display different peptides fused to a viral surface protein. After selecting phages that

have peptides with the desired characteristics, the genetic material of the phage can be amplified and sequenced to determine the peptide sequence. Smith *et al.*[7], Dower *et al.*[8] and Devlin *et al.*[9] described libraries of peptides fused to the gene III protein of phage fd. Later, others described similar peptide libraries fused to the filamentous phage proteins VII, VIII and IX (Refs 10,11), and to surface proteins of phage T7 (Refs 12,13).

There are at least three other schemes to select affinity peptides *in vitro* linked to the DNA that encodes it: (1) Random peptides are displayed fused to lac repressor[14]; in this scheme, the plasmid that encodes the repressor–peptide fusion also contains a lac operator, so that the repressor–peptide fusion binds to the plasmid that encodes it. (2) The peptide is linked to the yeast Gal4 activation region II, allowing selection of yeast that bear the activator–peptide chimera in a two-hybrid assay[15]. (3) Constrained peptides are displayed on *E. coli* Thioredoxin or other carrier proteins (TrxA), as further described below.

One of the questions most frequently asked about peptides selected by these methods is: what are they good for? Except for the special case in which selected peptides bind well to particular anti-protein antibodies (and whose sequence thus defines linear epitopes recognized by those antibodies), peptides selected by these schemes typically bind their targets with low affinity, too low to function in lieu of antibodies in conventional research protocols. The hope that the selected peptides might be useful as starting points for the design of small-molecule drugs that mimicked their structure has faded, or at least proved to be difficult. However, other uses have either materialized or show great promise: these include affinity reagents in purification schemes[16], and as signals that direct viral or cellular vectors to particular tissues and cell types[17,18] for therapies under development[19].

Development of intracellular protein aptamers

In 1993, John McCoy and co-workers[16] described another display scheme. The important feature of this approach is that the variable sequences are conformationally constrained by fusion at both ends to a platform protein. In this scheme, the variable region is displayed in lieu of the active site loop of *E. coli* TrxA, which is in turn displayed on the surface of the major *E. coli* flagellar protein, FliC. Discussions between McCoy and Brent led to the proposal that conformationally constrained peptides of variable sequence could be displayed by TrxA inside cells, and selected either to bind protein targets in two-hybrid interactor hunts, or to cause phenotypes directly by disrupting protein–protein interactions. Accordingly, we constructed three libraries that directed the synthesis in *Saccharomyces cerevisiae* of TrxA derivatives that contained constrained 20-mer variable regions and that did or did not carry the other moieties needed to work as 'preys' in the two-hybrid system.

In the first use of these libraries, Colas *et al.*[20] began with human Cdk2 as bait in a two-hybrid hunt to select

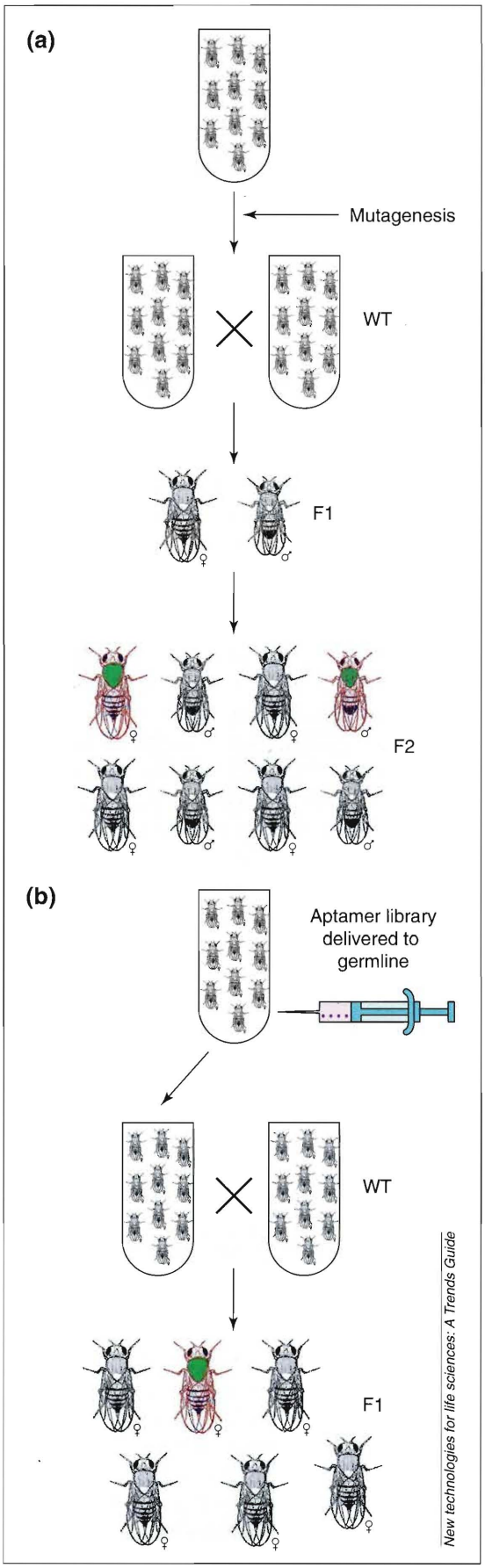

Figure 1. Use of protein aptamers for gene and protein identification

(a) Gene identifcation by F2 loss-of-function experiments. A diploid organism is mutagenized, in this case *Drosophila*, and F1 progeny are generated after a cross with a non-mutagenized organism (WT). F1 progeny are crossed with one another, or with members of stocks of organisms that contain haploid deletions, to create the F2. F2 progeny are surveyed to reveal those that exhibit a particular mutant phenotype (green organisms). Mutations are characterized further (e.g. by mapping, epistasis, gene cloning and sequencing) to elucidate their function. **(b)** Whole organism protein identification by dominant protein 'genetic' experiments. Cells or organisms (of any ploidy) are 'mutagenized' by introducing into individuals, constructs that direct the synthesis of individual protein aptamers. Synthesis of the aptamers is induced, and the cells or organisms (in this case in the F1) that expressed aptamers that cause a particular phenotype (green organisms) are identified. Constructs are rescued from these 'mutant' cells or organisms, and the proteins and/or protein interactions that they target are identified by two-hybrid or other affinity techniques. These techniques have been demonstrated to work in cells and unicellular organisms. Although the figure shows how they might be applied to *Drosophila*, the first applications of these methods to multicellular organisms will probably be for those of higher ploidy (for example maize) and/or for organisms that have not been subjected to extensive previous genetic analysis (including organisms not of great economic importance to developed countries), in which loss-of-function genetics is unlikely to be developed.

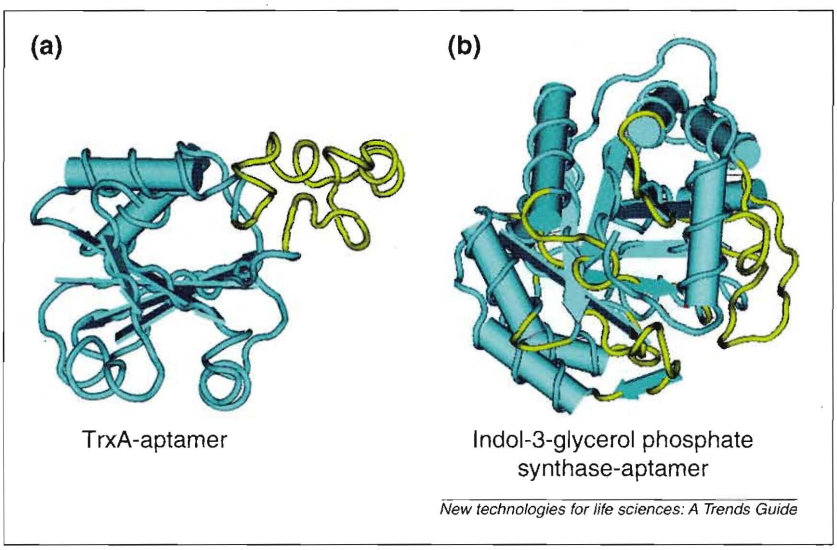

Figure 2. Protein aptamer structure

Protein aptamers consist of a constant platform (constant region) that also carries one or more conformationally constrained peptides of variable sequence. The figure shows modified ribbon diagrams of two platforms (light green) and variable regions (yellow). The structure of the V regions is arbitrary; in no case is an aptamer V-region structure known and, even for the best studied cases (anti-E2F TrxA-aptamers[23]), there is only indirect evidence that any V region is structured before the aptamer binds the target. The aptamer **(a)** has a single, continuous V region inserted in the active site of TrxA (Ref. 16), whereas the aptamer **(b)** carries eight V regions inserted in corresponding eight exposed loops of the enzyme indol-3-phosphate synthase[40]. Currently, TrxA aptamers have the highest reported affinities, with equilibrium dissociation constants in the nanomolar range. Abbreviation: TrxA, thioredoxin.

TrxA-variable region derivatives fused to activation domains ('preys') that specifically interacted with CDK2. These occurred in the library at a frequency of about 1 in every 10^5 expressed proteins. Some, but not most, crossreacted with related cyclin-dependent kinases. Most inhibited CDK2 activity, apparently by competing for its ability to bind the substrate. As measured directly (by binding) and indirectly (by inhibition of activity), all six aptamers tested had equilibrium dissociation constants below 100 nM, and, for one of them, it was below 2 nM. The relatively tight binding of these agents was attributed to the fact that the 20-mer variable region (V) was constrained at both ends, thus reducing the number of conformations assumed by the unbound variable region and lowering the entropic cost of assuming a single conformation upon binding (see also Ref. 21). These agents, proteins that express variable regions constrained at both ends by a platform protein, were termed peptide aptamers. Nowadays, these molecules are also called protein aptamers (Fig. 2).

More than 20 TrxA peptide aptamers have since been identified in single-step selections against numerous protein targets; one aptamer resulted from a round of mutagenesis of a successful V region followed by a second two-hybrid step using a reporter gene that demanded higher-affinity binding[22]. In cases where affinity was measured, first round aptamers had K_d values tighter than 100 nM (and frequently in the nM range); for the single second round aptamer, the K_d value increased from 100 nM to 5 nM (Ref. 22).

The selected aptamers function *in vivo*. Anti-Cdk2 aptamers inhibit the activity of their target proteins in human cells[21] and in developing *Drosophila*[23]. Anti-E2F aptamers affect the ability of E2F to form heterodimers with DP1 and to stimulate cell division in mammalian cells[24]. Anti-HPV16 E6 aptamers block the anti-apoptotic activity of E6 in HPV16$^+$ cancer cells, causing cell death[25]. Use of 'two-bait' two-hybrid systems allowed the selection of aptamers that interacted specifically with one allelic form of the Harvey Ras oncoprotein[26] and inhibited the activity of that form in transformed cells (C.W. Xu *et al.*, unpublished). In each case, the effect of the aptamers is likely to be due to binding their target *in vivo*, thereby keeping it from interacting with one or more partner proteins or other molecules. Although the selection is by two-hybrid, there is no reason why any of the surface display methods described above should not allow selection against larger numbers of baits *in vitro*. It thus seems likely that selection can be scaled up to isolate aptamers that interact with every protein encoded by a genome.

In the above cases, the approach is a reverse 'genetic' one. The word 'genetic' is in quotes as expression of peptide aptamers does not alter the cell's genetic material. However, by analogy with reverse genetic approaches where the word genetic is not in quotes, one begins with a target protein (or allelic form of a protein, or protein–protein interaction) and one inactivates it to determine its normal function. For example, the experiments of Butz *et al.*[25] demonstrate that one function of HPV E6 is to inhibit apoptosis. This reverse genetic approach does not need to be restricted to protein aptamers. Recently, Shi *et al.*[27] described a pentavalent RNA aptamer directed against the *Drosophila* splicing protein B52 that, when expressed in flies, inhibited B52 activity. RNA aptamers can be selected against many protein targets, can bind them tightly (many had K_d values below 1 nM) and can be stable *in vivo*[28]. Moreover, when aptamers come to be used as 'reverse genetic' reagents in vertebrate organisms, it is possible that nucleic-acid derived aptamers might be only weakly immunogenic, whereas animals expressing protein aptamers might need to be made tolerant to them.

Intracellular protein aptamers as forward 'genetic' reagents

More recently, aptamers have been used in forward 'genetic' experiments (Fig. 1b) to identify proteins and protein interactions that affect biological pathways. The first published uses of this approach were in the yeast *S. cerevisiae* to analyse the response to mating pheromone[29–31] and the spindle checkpoint[31]. These groups isolated aptamers [displayed, respectively, by GFP (Ref. 32), TrxA and *Staphylococcus aureus* nuclease] that blocked cell-cycle arrest caused either by mating pheromone or by the spindle checkpoint, and then used the aptamers as baits and/or preys in two-hybrid experiments to find the proteins targeted by those aptamers. This approach has been used in *E. coli* by Blum *et al.*[33] and extended to mammalian cells by G. Nolan and co-workers at Stanford and at Rigel, Inc. (pers. commun.) and by ourselves.

Several lines of evidence demonstrate the power of the combination of two-hybrid and aptamer technology. First, its use resulted in the identification of a protein, Cbk1, whose loss diminishes but does not abolish the mating response. Before this analysis, Cbk1 was not known to function in the pathway[30]. This finding illustrates that aptamer technology can be useful for the study of cellular processes that have been extensively explored

by conventional genetics. The authors showed that most of the aptamers targeted the same genetic step in the signal transduction pathway: the activation of Ste11, a prototype for the mitogen-activated protein kinase kinase kinase in mammalian cells. The finding that this step in this pathway is most sensitive to perturbation by diminution of function by aptamers suggests that the analogous step in mammalian pathways would be a good target for drug discovery efforts.

Second, as ordered panels of yeast bearing appropriately derivatized proteins for interaction mating two-hybrid experiments[34] are generated for more organisms[35] (R. Finley, pers. commun.), it will become easier to identify proteins targeted by the aptamers to produce their effects.

Third (as mentioned previously), the use of two-hybrid makes it easy to mutagenize the DNA encoding the aptamer variable region and select tighter-binding variants (B.A. Cohen, PhD thesis, Harvard University, 1998). The ability to generate such tighter-binding aptamers is helpful in making aptamer derivatives that can change the subcellular localization of their targets and/or decorate the targets with ubiquitin moieties[22].

Fourth, the use of appropriately configured two-hybrid constructs, in at least some cases, allows the investigator to define not just the protein, but the protein interaction that is targeted[30].

Finally, we imagine that the use of mixed pools of aptamers targeted to different subcellular compartments or derivatized with different effector functions will make it possible to generate some knowledge about the proteins that function in a given process, even before the specific targets of those aptamers are identified from ordered two-hybrid arrays or other interaction techniques.

Future developments and their impact on biological research

One reason that protein aptamer technologies are developing rapidly is that they make direct use of biology's means of inventing and perfecting designs: namely, generation of diversity followed by selection for fitness according to some criterion – followed by, if desired, improvement of the progeny by repeating this process. One foreseeable consequence of continued peptide aptamer development is their selection against a great many more targets, and their eventual adoption as affinity reagents in lieu of antibodies in 'immunological' techniques. Antibodies are wonderful affinity molecules, but the last significant improvement in this field, the invention of monoclonal antibodies, dates back to 1975 (Ref. 36). Since that time, there have been numerous schemes to generate antibody diversity and select antibodies *ex vivo*. Although space does not permit a comprehensive review of these methods, some of them (see, for example, the work of Knappik *et al.*[37]) have allowed the selection of single chain antibodies with picomolar dissociation constants. For several reasons, including the extent to which incremental advances have tended to become encrusted with patents, the defense of that intellectual property by undercapitalized commercial firms, and the fact that the business models of those firms normally require that they license and sell what they produce to a small number of large pharmaceutical companies, none of these approaches has gained wide adoption. For use in lieu of antibodies, aptamers must bind strongly and specifically and be able, in principle, to bind any protein surface. Compare a TrxA 20-mer V region aptamer with the two-chain, β-sheet-dominated, 6-variable regions (totalling more than 50 residues) of an immunoglobulin Fab. To gain the needed affinity, specificity and universality, researchers will need to explore different configurations of platforms and constrained variable regions (Fig. 2).

Peptide aptamers can be selected *in vitro*. The puromycin crosslinking *in vitro* peptide display scheme of Roberts *et al.*[38] allows selection of three variable region fibronectin-based aptamers from large (greater than 10^{12}) synthetic pools (B. Kreider, pers. commun.). Although that particular technology is largely proprietary to a single company, Phylos, the overall ideas are not, and in ten years time, it is likely that antibodies generated *in vivo* will be out-competed by these more modern (and in the case of aptamers, simpler) protein-based agents. Such development would free large numbers of mice and rabbits from their involuntary service to human research, diagnosis and therapy. It will also allow the generation of affinity reagents against the entire complement of genome-encoded proteins, specifically peptide aptamers directed against the products of different splice variants and against modification states of the proteins. Development of these affinity reagents will, in turn, enable the construction of new diagnostic devices and research instruments to survey entire proteomes.

It is likely that aptamers will find use in other aspects of biological discovery and analysis. A particularly good example is the case of redundant gene function. As is now widely appreciated, vertebrate animals are, roughly speaking, products of an ancestral gene quadruplication, and the four (or three, or five) copies of each gene frequently have redundant functions[39]. Armed with the right aptamers and the right selection schemes, it should be possible to generate aptamers that specifically crossreact with and inactivate the function of all members of a given protein family. Such aptamers can then be used to determine the function of an entire family of proteins in a single study rather than serially inactivating the genes and performing the necessary crosses.

Another possible development is highly speculative, but worth mentioning. Genetic analysis is easiest when applied to strong, penetrant phenotypes. In fact, we can even posit a straw-man geneticist, who would say that no gene is weakly acting or incompletely penetrant if it is in the right background, and if the right phenotype can be found. In the near future, however, there are at least two reasons why biologists might need to grapple with phenotypes affected by genes whose action might be weak and incompletely penetrant. One reason is simply the fact that there are huge piles (soon to grow vastly larger) of quantitative genetic

data (consider SNPs) that correlate with quantitative traits of interest (consider the odds of developing a particular malady). For some of these traits, there is no relevant model organism tractable to loss-of-function genetics. Another reason is that, even if we could devise phenotypes that would allow selection of decisive mutants for all important processes in, for example, mice, the whole process would take a great deal of effort and a very long time. We have argued that protein-family-specific peptide aptamers might provide a useful shortcut to the analysis of groups of genes of redundant function. We also believe that they will play a part in a necessary task – the analysis of gene products that quantitatively affect phenotypes. Comprehensive, scalable methods to address such traits in important organisms will increasingly be needed. The early 20th century taught us how to identify some genes in some organisms. The end of the 20th century taught us how to identify all the genes in any organism. The corresponding spirit of the early 21st century may well teach us how, in any organism, to identify all the gene products that contribute to a given trait.

Acknowledgements

We are grateful to R. Davis, R. Finley, L. Gold, B. Kreider, G.P. Nolan and J. Szostak for useful discussions and personal communications. We also acknowledge the influence of S. Schreiber and T. Mitchison, whose expressed hopes to isolate small molecule drugs that inhibit every gene product may result in technologies complementary to the ones we have described. This work was supported by a grant to R.B. from the NIH National Institute of General Medical Sciences. Other work at TMSI is supported by grants from the NIH National Human Genome Research Institute, the US Office of Naval Research and the US Defense Advanced Research Projects Agency.

References

1 Mendel, G. (1866) Experiments in plant hybridization. *Verh. Naturforsch. Ver. Brünn* 4, 3–47
2 Morgan, T.H. (1919) *The Physical Basis of Heredity*, Lippincott
3 Izant, J.G. and Weintraub, H. (1984) Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. *Cell* 36, 1007–1015
4 Melton, D.A. (1985) Injected anti-sense RNAs specifically block messenger RNA translation *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* 82, 144–148
5 Ellington, A.D. and Szostak, J.W. (1990) *In vitro* selection of RNA molecules that bind specific ligands. *Nature* 346, 818–822
6 Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249, 505–510
7 Scott, J.K. and Smith, G.P. (1990) Searching for peptide ligands with an epitope library. *Science* 249, 386–390
8 Cwirla, S.E. *et al.* (1990) Peptides on phage: a vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. U. S. A.* 87, 6378–6382
9 Devlin, J.J. *et al.* (1990) Random peptide libraries: a source of specific protein binding molecules. *Science* 249, 404–406
10 Gallop, M.A. *et al.* (1994) Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries. *J. Med. Chem.* 37, 1233–1251
11 Gao, C. *et al.* (1999) Making artificial antibodies: a format for phage display of combinatorial heterodimeric arrays. *Proc. Natl. Acad. Sci. U. S. A.* 96, 6025–6030
12 Rosemberg, A. *et al.* (1996) T7Select phage display system: a powerful new protein display system based on bacteriophage T7. *Innovations* 6, 1–6
13 Mikawa, Y.G. *et al.* (1996) Surface display of proteins on bacteriophage lambda heads. *J. Mol. Biol.* 262, 21–30
14 Cull, M.G. *et al.* (1992) Screening for receptor ligands using large libraries of peptides linked to the C terminus of the lac repressor. *Proc. Natl. Acad. Sci. U. S. A.* 89, 1865–1869
15 Yang, M. *et al.* (1995) Protein–peptide interactions analysed with the yeast two-hybrid system. *Nucleic Acids Res.* 23, 1152–1156
16 Lu, Z. *et al.* (1995) 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* 13, 366–372
17 Pasqualini, R. and Ruoslahti, E. (1996) Organ targeting *in vivo* using phage display peptide libraries. *Nature* 380, 364–366
18 Barry, M.A. *et al.* (1996) Toward cell-targeting gene therapy vectors: selection of cell-binding peptides from random peptide-presenting phage libraries. *Nat. Med.* 2, 299–305
19 Arap, W. *et al.* (1998) Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. *Science* 279, 377–380
20 Colas, P. *et al.* (1996) Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature* 380, 548–550
21 Cohen, B.A. *et al.* (1998) An artificial cell-cycle inhibitor isolated from a combinatorial library. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14272–14277
22 Colas, P. *et al.* Targeted modification and transportation of cellular proteins. *Proc. Natl. Acad. Sci. U. S. A.* (in press)
23 Kolonin, M.G. and Finley, R.L., Jr (1998) Targeting cyclin-dependent kinases in *Drosophila* with peptide aptamers. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14266–14271
24 Fabbrizio, E. *et al.* (1999) Inhibition of mammalian cell proliferation by genetically selected peptide aptamers that functionally antagonize E2F activity. *Oncogene* 18, 4357–4363
25 Butz, K. *et al.* (2000) Induction of apoptosis in human papillomavirus-positive cancer cells by peptide aptamers targeting the viral E6 oncoprotein. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6693–6697
26 Xu, C.W. *et al.* (1997) Cells that register logical relationships among proteins. *Proc. Natl. Acad. Sci. U. S. A.* 94, 12473–12478
27 Shi, H. *et al.* (1999) RNA aptamers as effective protein antagonists in a multicellular organism. *Proc. Natl. Acad. Sci. U. S. A.* 96, 10033–10038
28 Brody, E.N. and Gold, L. (2000) Aptamers as therapeutic and diagnostic agents. *J. Biotechnol.* 74, 5–13
29 Caponigro, G. *et al.* (1998) Transdominant genetic analysis of a growth control pathway. *Proc. Natl. Acad. Sci. U. S. A.* 95, 7508–7513
30 Geyer, C.R. *et al.* (1999) 'Mutagenesis' by peptide aptamers identifies genetic network members and pathway connections. *Proc. Natl. Acad. Sci. U. S. A.* 96, 8567–8572
31 Norman, T.C. *et al.* (1999) Genetic selection of peptide inhibitors of biological pathways. *Science* 285, 591–595
32 Abedi, M.R. *et al.* (1998) Green fluorescent protein as a scaffold for intracellular presentation of peptides. *Nucleic Acids Res.* 26, 623–630
33 Blum, J.H. *et al.* (2000) Isolation of peptide aptamers that inhibit intracellular processes. *Proc. Natl. Acad. Sci. U. S. A.* 97, 2241–2246
34 Finley, R.L., Jr and Brent, R. (1994) Interaction mating reveals binary and ternary connections between *Drosophila* cell cycle regulators. *Proc. Natl. Acad. Sci. U. S. A.* 91, 12980–12984
35 Hudson, J.R., Jr *et al.* (1997) The complete set of predicted genes from *Saccharomyces cerevisiae* in a readily usable form. *Genome Res.* 7, 1169–1173
36 Kohler, G. and Milstein, C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256, 495–497
37 Knappik, A. *et al.* (2000) Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* 296, 57–86
38 Roberts, R.W. and Szostak, J.W. (1997) RNA–peptide fusions for the *in vitro* selection of peptides and proteins. *Proc. Natl. Acad. Sci. U. S. A.* 94, 12297–12302
39 Holland, P.W. *et al.* (1994) Gene duplications and the origins of vertebrate development. *Development* (Suppl.) 125–133
40 Cho, G. *et al.* (2000) Constructing high complexity synthetic libraries of long ORFs using *in vitro* selection. *J. Mol. Biol.* 297, 309–319