1152–1156 Nucleic Acids Research, 1995, Vol. 23, No. 7

Protein–peptide interactions analyzed with the yeast two-hybrid system

Meijia Yang, Zining Wu[1] and Stanley Fields*

Department of Molecular Genetics and Microbiology, State University of New York at Stony Brook, Stony Brook, NY 11794, USA and [1]Department of Pharmacology and Toxicology, Dartmouth Medical School, Hanover, NH 03755, USA

Received January 5, 1995; Revised and Accepted February 15, 1995

ABSTRACT

The yeast two-hybrid system was used to screen a library of random peptides fused to a transcriptional activation domain in order to identify peptides capable of binding to the retinoblastoma protein (Rb). Seven peptides were identified, all of which contain the Leu–X–Cys–X–Glu motif found in Rb-binding proteins, although their activity in the yeast assay varied over a 40-fold range. Mutagenesis of the DNA encoding two of these peptides followed by screening in the two-hybrid system allowed the delineation of residues apart from the invariant Leu, Cys and Glu that affect binding to Rb. Binding affinities of a peptide and one of its variants to Rb, determined by surface plasmon resonance, correlated with results from the two-hybrid assay. This method offers several advantageous features compared to existing technology for screening peptide libraries: *in vivo* detection of protein–peptide interactions, high sensitivity, the capacity for rapid genetic screening to identify stronger and weaker binding peptide variants, and the use of a simple assay (transcriptional activity) as a means to assess binding affinity.

INTRODUCTION

Combinatorial libraries of macromolecules are finding increasing use as a means to identify ligands for biologically important targets. In particular, diverse approaches have been developed for the generation of large arrays of random peptides. These include libraries synthesized *in vivo* and displayed on the surface of a filamentous phage (1–3) or at the carboxy terminus of the *lac* repressor protein (4), and others synthesized chemically and coupled to resin beads (5). Biologically-derived peptide libraries are screened by sequential rounds of affinity enrichment ('panning'). Phage display libraries allow very complex mixtures of peptides to be screened; peptides with the relevant binding activity initially present at low abundance are amplified in subsequent rounds of infection. The primary structure of the specific peptides are then identified by DNA sequence analysis of the phage DNA. By contrast, peptides identified from chemically-derived libraries are generally characterized by Edman degradation (5) or by the analysis of molecular tags that had been attached to the beads for recording steps of the synthesis (6,7). These libraries have the advantage of allowing the incorporation of unnatural amino acids into the peptides. Recently, an *in vitro* polysome display system has been designed to select peptide ligands from 10^{12} random peptides (8). All of these methods detect protein–peptide binding under *in vitro* conditions, and require further biochemical experiments to determine affinity values for peptides identified in the random screen. Additionally, they have limited capacity for rapid discrimination of tighter or weaker binding variants. We describe here the use of a peptide library in the yeast two-hybrid system (9,10) which has unique features relative to existing technology.

The two-hybrid system relies on the modular nature of many eukaryotic site-specific transcriptional activators to generate a transcriptional signal from the interaction of a protein fused to a DNA-binding domain with another protein fused to a transcription activation domain. Unlike other methods, the two-hybrid system detects protein–protein interactions occurring *in vivo*. This leads to a high degree of sensitivity [see for example (11,12)], presumably due to an amplification effect in which a transient interaction can lead to multiple rounds of transcription and translation and the production of stable reporter proteins such as *E.coli* β-galactosidase. Furthermore, as a microbially-based assay, this system allows both the direct selection of only those yeast transformants that express a desired protein–protein interaction as well as the use of simple genetic strategies to identify small protein domains or single amino acid residues critical for such an interaction (11,12).

MATERIALS AND METHODS

DNA constructions and mutagenesis

Plasmids encoding the activation domain hybrids with T antigen peptides were generated by PCR amplification of the indicated T antigen sequences. The LTP hybrid was generated in pGAD.GH (13) and the LTP–Leu hybrid in pGAD424 (14). The peptide library was generated using the oligonucleotide 5′-GAAC-TAGTGGATTCCC(NNK)$_{16}$TAGGAATTCGGCCGC-3′ (N is an equal mixture of all four deoxynucleotides; K is an equal mixture of G and T) to encode 16 random codons flanked by DNA

* To whom correspondence should be addressed

sequences containing *Bam*HI and *Eco*RI restriction sites. PCR amplification was performed using 26 pmol oligonucleotide and 6 nmol of the primers 5′-GAACTAGTGGATTCCC-3′ and 5′-GCGGCCGAATTCCTA-3′. After digestion with *Bam*HI and *Eco*RI, fragments were ligated into pGAD.GH and electroporated into the *E.coli* strain DH5α to generate a library with ~10^7 inserts. Fifty plasmids were randomly chosen and sequenced. All contained inserts; ~15% of the inserts encoded peptides of <16 residues, due to the presence of termination codons generally resulting from a missing nucleotide in the oligonucleotide synthesis.

To generate hybrids of a peptide with the Gal4 DNA-binding domain, the multiple cloning site in the vector pGBT9 (14) was modified to match to that of the pGAD.GH. The *Eco*RI of pGBT9 was changed to a *Spe*I site using the adaptor 5′-AATACTAGT-3′, and then the *Spe*I–*Sal*I fragment of the plasmid pBS SK ± (Stratagene Co., San Diego, USA) containing the multiple cloning site was inserted into the *Spe*I/*Sal*I sites to generate pGBT9BS. DNA sequences encoding peptides P1, P1-a1, P7, and P7-14 in the vector pGAD.GH were isolated as *Bam*HI–*Eco*RI fragments and ligated into pGBT9BS.

To create a fusion of glutathione S-transferase to Rb residues 301–928, the *Eco*RI site of pGEX-3X (Pharmacia, Uppsala, Sweden) was changed to a *Sal*I site using the adaptor 5′-AATTGTCGAC-3′, and the *Bam*HI–*Sal*I fragment of pASRb2 (15) was inserted into this vector.

Mutations were generated by using PCR under suboptimal conditions as described (16).

Yeast methods

The yeast reporter strain Y153 (14) was transformed and screened for His^+ $lacZ^+$ colonies, plasmids were recovered and analyzed, and β-galactosidase activity was determined as described (17). Values for liquid β-galactosidase assays represent the mean of at least four assays, and the standard errors were generally 10–20% of the mean.

Surface plasmon resonance analysis

The GST–Rb fusion was purified as described (18). Peptides were purchased from Quality Controlled Biochemicals, Inc. (Hopkinton, MA). Binding kinetics were determined by surface plasmon resonance (BIAcore, Pharmacia Biosensor). Purified GST–Rb was diluted to 15 μg/ml in sodium acetate buffer (10 mM, pH 4.5) and was coupled to the dextran-modified gold surface of a CM5 sensor chip (Pharmacia Biosensor) using amine coupling chemistry [*N*-hydroxysuccinimide/*N*-ethyl-*N*′-(3-dimethylaminopropyl)-carbodiimide] as described in the system manual. Peptides were dissolved in 10 mM NaOH and the pH was adjusted to 7.4 prior to being diluted to the desired concentrations in HBS buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3.4 mM EDTA, 0.005% Surfactant P-20) containing 100 μg/ml BSA. Serial dilutions of each peptide (1–10 μM of P1 and 5–50 μM of P1-a1) were injected over the GST–Rb coupled surface at 8 μl/min. Sensorgrams were recorded and normalized to a baseline of 0 resonance units (RU). Data were fitted to a single site binding model of A + B = AB by non-linear least squares curve fitting using BIAevaluation 2.0 software (Pharmacia Biosensor). The equation $R_t = R_o \times \exp[-k_{off} \times (t-t_o)]$ was used for the dissociation phase, where R_t was the amount bound in RU at time t, and t_o was the beginning of the dissociation phase. In one experiment, the dissociation rate constant (k_{off}) of P1 binding was $(5.03 \pm 0.35) \times 10^{-3}\ s^{-1}$, and of P1-a1 binding was $(1.13 \pm 0.05) \times 10^{-2}\ s^{-1}$). The equation $R_t = R_{eq} \times \{1 - \exp[-k_s \times (t-t_o)]\}$ was used for the association phase, where $k_s = k_{on} \times C + k_{off}$ and the association rate constant k_{on} was calculated by the linear curve fitting of a plot of k_s versus C. R_{eq} was the amount bound in RU at equilibrium, t_o was the time that injection started, and C was the concentration of peptide injected. In one experiment, for P1, $k_{on} = (3.77 \pm 0.20) \times 10^2\ M^{-1}\ s^{-1}$ (correlation coefficient, $R = 0.983$) and for P1-a1, $k_{on} = (1.48 \pm 0.12) \times 10^2\ M^{-1}\ s^{-1}$ ($R = 0.974$). For this experiment, the equilibrium constants K_D (calculated from the ratio k_{off}/k_{on}) are 13.3 μM for P1 and 76.4 μM for P1-a1. K_D was also calculated by the linear curve fitting of a plot of R_{eq} versus R_{eq}/C using the equation $R_{eq}/C = 1/K_D \times R_{max} - 1/K_D \times R_{eq}$, where R_{max} is the maximum binding in RU when the surface is saturated and R_{eq} was calculated as above.

RESULTS AND DISCUSSION

Identification of Rb-binding peptides

To test the feasibility of the two-hybrid system to detect protein–peptide interactions, we chose to analyze the binding of small peptides to the human retinoblastoma (Rb) protein. Rb binds to viral oncogene products such as SV40 large T antigen, adenovirus E1A and human papilloma virus E7 and to cellular proteins such as D-type cyclins (19). Rb-binding in these cases is mediated by the amino acid motif Leu–X–Cys–X–Glu (20–23). We generated Gal4 activation domain hybrids that contain peptides from SV40 large T antigen: the LTP (for Large T antigen Peptide) hybrid contains residues 103–115, which begins with the Leu of the conserved motif, and the LTP–Leu hybrid contains residues 104–115 (Table 1). The plasmid pASRb2 (15) encodes residues 301–918 of Rb fused to the Gal4 DNA-binding domain (BD–Rb). When the hybrid proteins LTP and BD–Rb were co-expressed in a yeast reporter strain, binding of the T antigen peptide to Rb was observed as substantial β-galactosidase activity (Table 1). By contrast, the LTP–Leu construct resulted in only background transcriptional activity (Table 1), indicating that protein–peptide interactions can be specifically detected in two-hybrid assays.

Table 1. A peptide from SV40 large T antigen and peptides identified from a library screen interact with the Rb protein in the two-hybrid assay

Peptide	Sequence[a]	β-galactosidase activity[b]
LTP	**L**FC**S**E**E**MPSSDDE	53
LTP–Leu	FC**S**E**E**MPSSDDE	0.2
P1	YGLWI**L**W**C**D**E**EGLDLG	45
P2	NQLLGDV**L**A**C**Y**E**QEVE	6.9
P3	WTEL**L**F**C**F**E**QVYGDPF	6.8
P4	EGGD**L**G**C**D**E**SWSEGYT	6.0
P5	CDG**L**L**C**T**E**TLL	5.4
P6	GGCPGAN**L**C**C**F**E**KSLD	2.5
P7	TTWRER**L**R**C**E**E**NGLGV	1.1

[a]Peptide sequences are aligned by the conserved residues (indicated in bold); LTP is derived from SV40 large T antigen.
[b]Activity was determined as in (17).

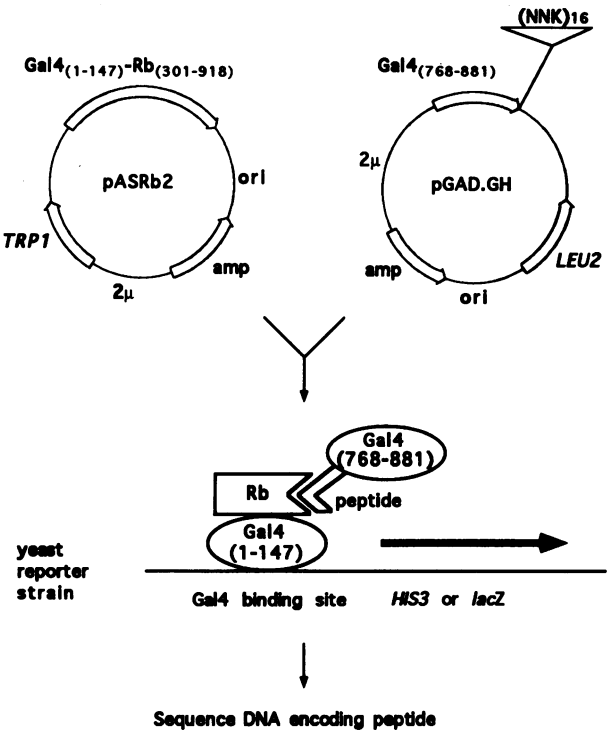

Figure 1. Screen for Rb-binding peptides. One plasmid encodes a fusion of the Gal4 DNA-binding domain (residues 1–147) to a large domain of Rb; the other plasmid is used to generate a library of 16-long peptides (encoded by NNK, N is all four nucleotides and K is G and T) fused to the Gal4 activation domain (residues 768–881). The Rb plasmid and the library are introduced into a yeast strain containing *HIS3* and *lacZ* genes under the regulation of Gal4 binding sites. Positive yeast transformants are identified and for each, the DNA sequence encoding the peptide is determined.

We constructed a library of 16-residue long peptides fused to the Gal4 activation domain and screened it to identify Rb-binding peptides (Fig. 1). The library was generated by cloning a pool of random oligonucleotides into the Gal4 activation domain vector pGAD.GH (13); the total complexity was 10^7. Statistically, the library should encompass at least all possible pentamer peptides. We cotransformed this library with pASRb2, encoding the Gal4 DNA-binding domain fusion to Rb, into a yeast reporter strain; selection was for the ability of transformants to grow in the absence of histidine by virtue of the expression of the Gal4-dependent gene *GAL1–HIS3*. His^+ colonies were also screened by a filter lift assay for expression of another Gal4-dependent gene *GAL1–lacZ*. Of three million transformants, seven peptides that bind to Rb were identified (Table 1).

All seven peptides contain the conserved Leu–X–Cys–X–Glu motif. However, transcriptional activity in the two-hybrid assay varied more than 40-fold, indicating that residues other than the three invariant ones play a role in determining this activity. Notably, the five peptides producing the most β-galactosidase activity have a high net negative charge (average –2.6) and contain no Arg or Lys residues. In contrast, P6 has an overall charge of –1 and contains a single Lys residue, and P7 is overall neutral and contains three Arg residues (Table 1). Additionally, the peptides producing the most activity tend to be hydrophobic in nature. Thus P1, with a net charge of –4 and six hydrophobic residues (apart from the invariant Leu and Cys), produced more than 7-fold greater β-galactosidase activity than did P4, with a net charge of –5 but only a single hydrophobic residue outside the invariant motif. The similar activity of P1 and the large T antigen peptide may be due to the presence of a negatively charged residue immediately C-terminal to the conserved motif. Figge *et al.* (22) analyzed the sequences of the Rb-binding domain from 21 proteins and derived a consensus that includes a net negative charge flanking the conserved motif, a bias against Lys and Arg residues, and an overall hydrophobic nature. However, their consensus includes such features as an Asp, Asn or Thr immediately N-terminal to the conserved Leu, as well as a Glu or Asp one to three residues N-terminal to the conserved Leu; these features are not found in the P1 peptide.

Affinity analysis

To analyze the role of residues other than the conserved motif, we took advantage of the ability in the two-hybrid system to readily detect changes in transcriptional activity as colonies that are darker or lighter blue in the Xgal filter assay. We randomly mutagenized the peptide P1, displaying greatest activity, and P7, with least activity, by amplifying their corresponding oligonucleotides using suboptimal PCR (16) and cloning of the PCR products into the activation domain vector. We obtained six peptides derived from P1 that produced less β-galactosidase activity than P1, and two that produced similar activity (Table 2). All six with decreased activity have a mutation that leads to a +1 change in charge, either by loss of an Asp or Glu or gain of an Arg. The two peptides with similar activity to P1 have no net change in charge. For P7, the two mutants with increased activity have additional negative charge; change of an Arg residue to a neutral residue (Gly) doubled the activity whereas change of this residue to an acidic residue (Glu) led to more than a 3-fold increase (Table 2).

Table 2. P1 and P7 peptide derivatives and their binding to the Rb protein in the two-hybrid assay

Peptide	Sequence	β-galactosidase activity
P1	Y G L W I **L** W **C** D **E** E G L D L G	45
P1-r	- - - - T - - - - - - - - - - -	45
P1-a	- - - - - - - - - E - - - - - - -	42
P1-h	- - - R - - - - - - - - - - - -	27
P1-l	- - - - - - R - - - - - - - - -	22
P1-i	- - - - - - - - - - - - - - G - -	18
P1-20	- - - - - - - - G - - - - - - -	16
P1-K	- - - - - - - - - - G - - - - -	12
P1-a1	- - R - - - - - - - - - - - - -	11
P7	T T W R E R **L** R **C** E **E** N G L G V	2.2
P7-7	- - - G - - - - - - - - - - - -	4.4
P7-14	- - - E - - - - - - - - - - - -	7.3

To compare the two-hybrid results with biochemical affinity determinations, we synthesized the peptides P1, P1-a1, P7 and P7-14 (Table 2) and purified Rb from *E.coli* as a glutathione S-transferase fusion protein. Association and dissociation kinetics were measured using surface plasmon resonance technology (24) with a BIAcore system (Pharmacia Biosensor). P1 bound to Rb with an apparent K_D of 13–23 μM, and P1-a1 with an apparent

K_D of 61–76 μM. The binding of neither P7 nor P7-14 could be detected by this assay. These results correlate well with our initial identification of P1 and P7 as peptides producing the greatest and least β-galactosidase activities, respectively. Additionally, P1 produced 4.2-fold more activity in the two-hybrid assay than the P1-a1 mutant identified as having lower activity; this result compares with the ~4-fold effect in the BIAcore binding assay. These binding data suggest that the two-hybrid assay can detect affinities well above ~70 μM, which compares with K_D values as high as 100–500 μM that can be detected with phage and LacI display systems (25).

Use of peptides in DNA-binding domain hybrids

We switched the orientation of the Rb domain and selected peptides such that the Rb domain was fused to the Gal4 activation domain (AD–Rb) and the peptide to the Gal4 DNA-binding domain. For the combinations that included P7 and its mutant P7-14, the two-hybrid transcriptional signals (Table 3) were similar to those with the original orientation (Table 2). These results indicate that small peptides could serve as potential targets in a two-hybrid approach to identify proteins that bind to a peptide of interest. For the P1 peptide, the DNA-binding domain alone produced substantial transcriptional activity, possibly due to the high negative charge of the peptide. We obtained much greater β-galactosidase activity when AD–Rb was co-expressed (Table 3), indicating that binding between Rb and the peptide occurred. However, the two P1 mutants produced reduced transcription both in the absence and presence of AD–Rb compared to the wild type P1 sequence (Table 3). Therefore, it is not possible to use the two-hybrid assay with peptides of this type to assess their binding affinities to some protein, as mutations can affect intrinsic transactivation capacity as well as protein–peptide interaction.

Table 3. Peptides can function in the two-hybrid assay as fusions with the Gal4 DNA-binding domain

Peptide hybrid	β-galactosidase activity	
	+vector	+AD–Rb
BD-P7	0.2	1.6
BD-P7-14	0.1	4.4
BD-P1	66	110
BD-P1-a1	17	66
BD-P1-20	27	58

CONCLUSION

The combinatorial peptide library we describe here allows the identification of specific peptides in a single round of screening. Furthermore, the peptides bind to their target proteins within the cellular milieu, an advantageous feature for peptides that are designed for *in vivo* uses. Biochemical analysis of protein–peptide binding correlates well with the two-hybrid results, indicating that rapid assessments of affinity can be performed using the yeast assay. These features suggest that the two-hybrid system may have relevance for drug discovery schemes based on the identification of peptides with affinity for selected proteins. Initially, peptides identified in the two-hybrid assay may have low binding affinity. However, the sequence of these peptides may suggest a consensus for binding. In a second stage, mutagenesis of individual peptides and rapid rescreening in yeast can be performed. This stage can determine which residues are critical for binding and which residues can be mutated without loss of binding. It may also identify new peptides with increased affinity. Ultimately, the peptides identified as having the highest affinity might form the basis for the synthesis of mimetics to be used as therapeutic agents.

It should also be feasible to design simple genetic assays to identify peptide variants capable of binding to one target protein but not to another. For example, it might be possible to identify peptides that discriminate between wild type and mutant versions of a protein. The two-hybrid assay might also be capable of identifying peptides that can inhibit specific protein–protein interactions; such peptides could be synthesized within cells to examine the phenotypic consequences.

While the peptides used in these applications have a free carboxy terminus, it is likely that the peptide moiety can also function when present at the amino terminus of the hybrid protein. In either of these cases, one end of the peptide would be unconstrained by other portions of the polypeptide. However, it should also be feasible to construct the peptide library such that the insertion is flanked at both ends by domains of the polypeptide such that the peptide is embedded in a rather rigid structure.

In contrast to some phage display libraries, there is no limit to the number of residues that can comprise the peptide in the two-hybrid assay. Finally, peptides found in screening an activation domain library can themselves be used as targets in a two-hybrid search to identify cellular proteins capable of binding these peptides.

ACKNOWLEDGEMENTS

We thank B. Gunter, X. Cao, P. Tegtmeyer, R. Sternglanz, P. Bartel and B. Aronson for discussions or comments on the manuscript, M. Botfield for comments and suggestions on the BIAcore analysis and T. Durfee and W. -H. Lee for providing unpublished results and plasmid pASRb2. This work was supported by grants from the National Institutes of Health, New York State Science and Technology Foundation, and Amgen, Inc.

REFERENCES

1 Scott, J.K. and Smith, G.P. (1990) *Science*, **249**, 386–390.
2 Cwirla, S.E., Peters, E.A., Barrett, R.W. and Dower, W.J. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 6378–6382.
3 Devlin, J.J., Panganiban, L.C. and Devlin, P.E. (1990) *Science*, **249**, 404–406.
4 Cull, M.G., Miller, J.F. and Schatz, P.J. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1865–1869.
5 Lam, K.S., Salmon, S.E., Hersh, E.M., Hruby, V.J., Kazmierski, W.M. and Knapp, R.J. (1991) *Nature*, **354**, 82–86.
6 Needels, M.C., Jones, D.G., Tate, E.H., Heinkel, G.L., Kochersperger, L.M., Dower, W.J., Barrett, R.W. and Gallop, M.A. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 10700–10704.
7 Ohlmeyer, M.H.J., Swanson, R.N., Dillard, L.W., Reader, J.C., Asouline, G., Kobayashi, R., Wigler, M. and Still, W.C. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 10922–10926.
8 Mattheakis, L.C., Bhatt, R.R. and Dower, W.J. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 9022–9026.
9 Fields, S. and Sternglanz, R. (1994) *Trends Genet.*, **10**, 286–292.
10 Fields, S. and Song, O.-K. (1989) *Nature*, **340**, 245–246.
11 Li, B. and Fields, S. (1993) *FASEB J.*, **7**, 957–963.
12 Van Aelst, L., Barr, M., Marcus, S., Polverino, A. and Wigler, M. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 6213–6217.

13 Hannon, G.J., Demetrick, D. and Beach, D. (1993) *Genes Dev.*, **7**, 2378–2391.
14 Bartel, P.L., Chien, C.-T., Sternglanz, R. and Fields, S. (1993) In Hartley, D.A. (ed.), *Cellular Interactions in Development: A Practical Approach*. Oxford University Press, Oxford, pp. 153–179.
15 Durfee, T., Becherer, K., Chen, R.-L., Yeh, S.H., Yang, Y., Kilburn, A.E., Lee, W.H. and Elledge, S.J. (1993) *Genes Dev.*, **7**, 555–569.
16 Leung, D.W., Chen, E. and Goeddel, D.V. (1989) *Technique*, **1**, 11–15.
17 Iwabuchi, K., Li, B., Bartel, P.L. and Fields, S. (1993) *Oncogene*, **8**, 1693–1696.
18 Iwabuchi, K., Bartel, P.L., Li, B., Marraccino, R. and Fields, S. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 6098–6102.
19 Wiman, K.G. (1993) *FASEB J.*, **7**, 841–845.
20 DeCaprio, J.A., Ludlow, J.W., Lynch, D., Furukawa, Y., Griffin, J., Piwnica-Worms, H., Huang, C.-M. and Livingston, D.M. (1989) *Cell*, **58**, 1085–1095.
21 Dyson, N., Howley, P.M., Munger, K. and Harlow, E. (1989) *Science*, **243**, 934–937.
22 Figge, J., Breese, K., Vajda, S., Zhu, Q.-L., Eisele, L., Andersen, T.T., MacColl, R., Friedrich, T. and Smith, T.F. (1993) *Prot. Sci.*, **2**, 155–164.
23 Moran, E. (1988) *Nature*, **334**, 168–170.
24 Fisher, R.J. and Fivash, M. (1994) *Curr. Opin. Biotech.*, **5**, 389–395.
25 Gordon, E.M., Barrett, R.W., Dower, W.J., Fodor, S.P.A. and Gallop, M.A. (1994) *J. Med. Chem.*, **37**, 1385–1401.