The yeast two-hybrid system for studying protein–protein interactions

Jeremy Luban and Stephen P Goff

Columbia University and Howard Hughes Medical Institute, New York, USA

A great number of investigators are currently employing the yeast two-hybrid system to study protein–protein interactions. Recent applications and newer configurations of the technique include yeast strains with improved selectivity and screens for false-positive clones.

Current Opinion in Biotechnology 1995, 6:59–64

Introduction

Over the past several years, the two-hybrid system, first formally described by Stan Fields [1], has proved to be an extremely effective method for studying protein–protein interactions. This powerful methodology employs a transcriptional 'read-out' in yeast and derives its broad applicability from the inherent modularity and flexibility of eukaryotic transcription factors. It has been used successfully to study an array of defined protein–protein interactions as well as to identify proteins encoded by selected members of a cDNA library that interact with a given protein of interest. Here, we review the development of the system, some of its uses and forms, and some commonly encountered difficulties.

Modularity and flexibility of eukaryotic transcription factors

The two-hybrid system developed from basic observations regarding the function of eukaryotic transcription factors. Many eukaryotic genes are expressed at low basal levels unless induced by transcriptional activators. Transcriptional activator proteins, such as GAL4, have two separable functions (Fig. 1a). The first function is to bind to a specific sequence on the DNA. The second function is to increase the frequency with which transcription is initiated on the adjacent gene, presumably through interactions with the RNA polymerase or associated proteins.

Critical to the development of the two-hybrid system were the discoveries that the DNA-binding and transcriptional activation functions are separable and that each functional domain acts independently. When the DNA-binding domains of the yeast GAL4 [2] and GCN4 [3] transcriptional activator proteins were identified, it was found that these domains were not sufficient to activate transcription: a separate acidic activation domain was also required. The truly modular nature of transcription factors was demonstrated by experiments in which a hybrid LexA–GAL4 fusion protein was found to activate transcription in yeast containing a reporter gene under the control of *LexA* operator sequences [4]. In the case of this fusion protein, the DNA-binding activity was provided by the LexA component and the activation function by the GAL4 component.

Not only can heterologous DNA-binding and activation domains function as a single fusion protein, but these two domains can also function when non-covalently joined via protein–protein interactions. GAL80 is a protein that normally binds to GAL4 and prevents it from activating transcription. If GAL80 is itself fused to an activation domain, it will bind to GAL4 and activate transcription from genes under the control of GAL4 operators [5]. This spatial flexibility is perhaps related to the ability of these activators to function at a distance from promoters, spanning intervening loops of DNA.

Development of the two-hybrid system

Following the lead of the above experiments, it was then formally demonstrated that virtually any pair of proteins that interact with each other may be used to bring separate activation and DNA-binding domains together to reconstitute a transcriptional activator [1]. In a typical application, protein X is expressed as a fusion with a specific DNA-binding domain, such as that of GAL4; this fusion protein is often termed the 'bait'. Protein Y is expressed as a fusion to an activation domain, such as that of GAL4; this is termed the 'prey'. If the bait and the prey are expressed simultaneously in the nucleus of a cell, transcription of genes under the control of GAL4

Abbreviation

GST—glutathione *S*-transferase.

 ISSN 0958-1669

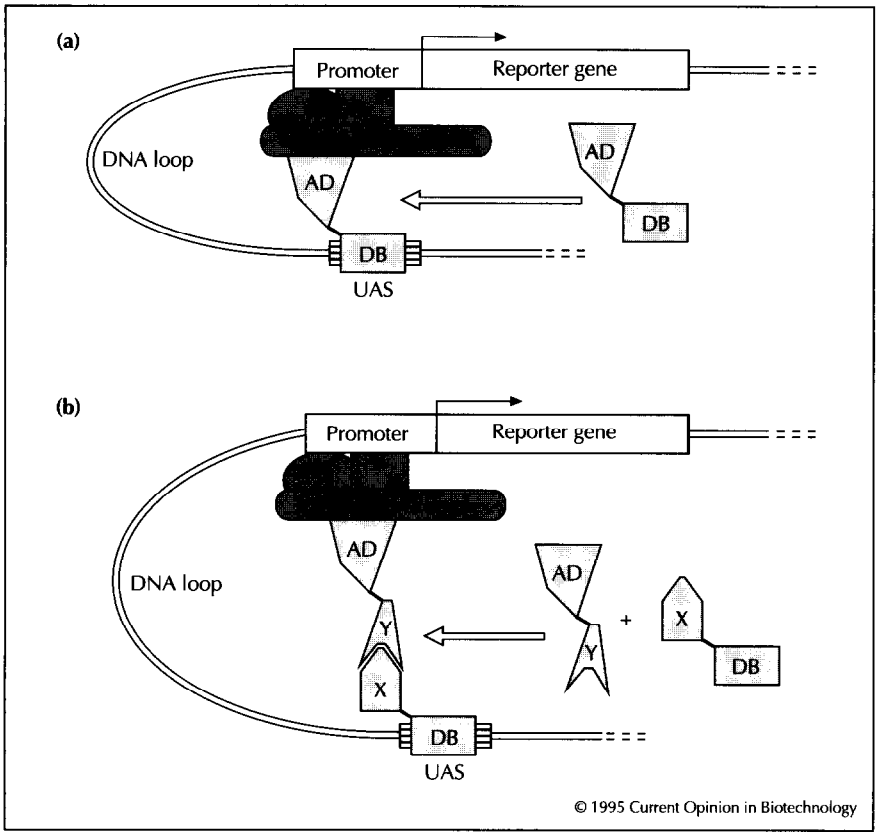

Fig. 1. Schematic diagram depicting the activation of transcription in a typical eukaryotic system and in the two-hybrid system. **(a)** Activation of transcription by a typical eukaryotic transcriptional activator with a separable DNA-binding domain (DB) and activation domain (AD). The operator or upstream activation sequences (UAS) specifically recognized by the DNA-binding domain are represented by the striped boxes. **(b)** Activation of transcription by chimeric proteins in the two-hybrid system. X represents a given protein fused to a specific DB. In library screens, this protein is termed the 'bait'. Y represents a given protein, or a pool of proteins encoded by a cDNA library, fused to a transcriptional activation domain. This fusion protein is often termed the 'prey'. If X and Y bind to each other, AD is brought to the vicinity of the DNA-bound DB and transcription is activated from the adjacent promoter.

will be activated (Fig. 1b). In this way, the proteins X and Y serve as molecular 'glue' bringing together the separately expressed activation and DNA-binding domains, and thereby reconstituting the activity of the transcriptional activator.

When compared with more traditional methods for studying protein–protein interactions, the two-hybrid system offers several advantages. First, as the interaction occurs *in vivo*, the need for the detailed and laborious manipulation of conditions necessary for *in vitro* biochemical binding assays is abrogated. Second, in at least some cases, the two-hybrid system is more sensitive than co-immunoprecipitation in detecting weak interactions [6]. Lastly, the biggest advantage of the two-hybrid system is that it readily lends itself to the design of genetic experiments (see below).

Configurations of the two-hybrid system

A variety of DNA-binding proteins can be used in the two-hybrid system, including the DNA-binding domains of prokaryotic proteins [4]. The reporter gene simply contains the operator or upstream activation sequences recognized by the particular DNA-binding protein. GAL4 [1,7,8••] and LexA [9••,10••] are the two DNA-binding proteins most commonly used in the two-hybrid system. A less commonly used system is based on the serum response factor [11,12]. A given protein may function better when fused to one or the other of these domains. The transcriptional activator function is most commonly provided by the GAL4 protein or by herpes simplex virus VP16 [10••,11,13].

In theory, one should be able to utilize any reporter gene that functions in yeast to monitor transcriptional activation. Most commonly, *lacZ* has been utilized. The enzymatic activity of an individual colony grown in liquid media may be measured by standard methods [14]. More commonly, the activity of given colonies can be assessed on nitrocellulose filter replicas soaked in X-gal [15]. This method is clearly preferred for screening libraries because the blue color of an individual colony is easily seen amid a 'sea' of white colonies. Also, live yeast can be recovered directly from the filters, even after fixing and staining.

Prototrophic markers are the most commonly used reporter genes other than *lacZ*. These reporters have the advantage that they allow selection for colonies with transcriptional activation. Some researchers have had difficulty using this selection method, presumably as a result of the fact that the conditions of the assay were too stringent to detect the particular interaction. *HIS3* is an example of a reporter that is commonly used. Expression from one such *HIS3* reporter using GAL4 upstream activation sequences is leaky, but background cell growth is reduced by growing the cells in the presence of 3-aminotriazole, an inhibitor of the encoded enzyme [8••]. Other yeast strains have been developed in which the *HIS3* reporter is more tightly regulated; these strains, therefore,

do not require 3-aminotriazole [10••]. Another system utilizes the *LEU2* gene as a selectable marker [9••]. In general, we advise using a system that allows both screening and selection in the same yeast strain. Even if one uses *HIS3* or *LEU2* as reporters, it is still advisable to use the *lacZ* screen secondarily to eliminate any false positives [8••].

Applications of the two-hybrid system

The two-hybrid system is an effective means of studying a protein–protein interaction between known partners and for determining the primary sequence requirements for the given interaction by testing the activity of engineered mutant proteins [13,16–24]. One may also use the two-hybrid system to screen libraries of random mutant proteins for the ability to interact with a given protein. For example, libraries of random-deletion or point mutants of a given protein may be created, by any of a number of methods. Pools of these mutants may then be fused to an activation domain and screened using the two-hybrid system for the ability to bind to a bait [6,18]. In this way, one may determine the primary sequence requirements for a given interaction without preconceived notions concerning the importance of given residues.

The full potential of the two-hybrid system is realized when it is used to screen cDNA libraries for encoded proteins that interact with a given protein of interest [25]. These screens have revealed many unexpected interactions between known proteins, and served to identify many previously unknown partners of various baits [8••,26–28,29•,30–37]. In this approach, the cDNA library is fused to sequences encoding the activation domain because many proteins will contain acidic blobs or other activation domains which activate the reporter gene when fused to the DNA-binding domain. A variety of libraries have been subcloned into yeast fusion-protein expression vectors and used successfully in the two-hybrid system. A phage vector designed specifically for the two-hybrid system employs Cre-*lox* site-specific recombination to convert the phage to a plasmid [8••].

Problems and trouble-shooting

One may screen a library using the full protein, or just a domain of a protein, fused to the DNA-binding domain. Baits as large as the complete retinoblastoma tumor suppresser protein [38] or as small as 22 amino acids (MB Carlson, personal communication) have been used successfully. In some cases, such as screens for proteins that interact with the cytoplasmic tails of transmembrane receptors, it makes sense to use only the cytoplasmic domain. Whatever fragment is used must be expressed as a stable fusion protein, which can usually be monitored by western blotting of yeast lysates. If antibodies are not available for the protein of interest, one may use commercially available anti-GAL4 antibodies. Alternatively, some vectors are designed with epitope tags for monitoring expression [8••]. In addition to stable expression, the fusion protein must be capable of transport to the nucleus and must be able to bind to its partner in solution. If possible, before embarking on a library screen, we prefer to perform control experiments using proteins that are known to interact with the protein of interest.

One problem that is commonly encountered with the two-hybrid system is transcriptional activation in yeast solely transformed with a plasmid expressing the protein of interest fused to a DNA-binding domain. These proteins cannot be used as bait in two-hybrid screens. The problem may be remedied either by constructing fusions with different DNA-binding domains or by the use of smaller individual domains of the protein of interest. This issue is relevant to the analysis of any constructs tested in the two-hybrid system because in certain cases, we have observed mutants that are 'self-activating', even when the wild-type parent is not [22].

An additional problem that is commonly encountered with library screens is achieving optimal transformation efficiency. Most groups use some variation of standard lithium acetate transformation protocols (e.g. see [39]). By paying particular attention to the freshness of the plated yeast streak and the quality of the carrier DNA, one may obtain yeast transformation efficiencies sufficient to screen several million colonies. To improve the efficiency, rather than performing a co-transfection, some researchers transform the yeast with the 'bait' expression plasmid first, and then subsequently transform with the cDNA library plasmids. Other workers have transformed the bait into yeast of one mating type and the library into yeast of the opposite mating type [39]. The two strains are then mated, and the bait and library plasmids are brought together in the diploid cells. This method has the advantage that each culture is transformed only once, rather than twice. Also, the library DNA can be introduced into yeast once and the recipients stored frozen.

Another major problem with two-hybrid screens is the appearance of false-positives inherent in any transcriptional readout. Any candidates from the primary screen must be retested in several configurations before they can be considered 'true' positives. First, the library plasmid must be separated from the bait plasmid. This is usually done by recovering the plasmid from yeast by transforming a bacterial strain that is an auxotroph for the marker on the plasmid encoding the candidate gene. A recently developed system precludes the requirement for recovery of the plasmids through yeast [29•]. This system uses a yeast strain that is cycloheximide resistant. The plasmid encoding the DNA-binding domain fusion protein has the *CYH2* gene conferring cycloheximide sensitivity. If the yeast are grown in cycloheximide, they are cured of the DNA-binding domain fusion protein expression plasmid. The resulting strain can then be mated to a

strain that has been transformed with all the appropriate control plasmids.

Once a plasmid has been recovered, it should be tested in the absence of the bait. A given protein may activate transcription if it binds to the promoter or to another protein that is part of the transcriptional machinery. Alternatively, a given protein may bind to the DNA-binding domain of the bait fusion and should be tested against the parental plasmid of the bait. Often, a false positive does not give activity against the DNA-binding domain alone, but does give activity when unrelated proteins are fused to it. Some researchers have found the lamin C fusion to be a useful control for this purpose [40]. Additionally, some false positives only exhibit activity with particular promoters or with promoters in the context of particular operator sequences. By using a selection and a screen with genes driven by different promoters, one can eliminate many of these false positives [8••]. Strong evidence for a direct protein–protein interaction is also provided if transcription is activated when the bait is expressed as fusion with the activation domain, and the candidate cDNA is expressed as a fusion with the DNA-binding domain.

Finally, even if a given clone passes through all the control experiments in yeast, we prefer to see proof of a direct interaction between the proteins in a biochemical assay. One common approach is to express the candidate cDNA as a fusion protein with glutathione *S*-transferase (GST) in bacteria [41]. One then demonstrates that the bait protein is precipitated with the GST-fusion protein bound to glutathione-beads [8••,34]. The two-hybrid fusion proteins may subsequently be co-immunoprecipitated in the yeast reporter strain, but ideally, one should show that the two proteins co-precipitate in their native context [35,37].

The last issue, which is usually the most difficult aspect of working with the two-hybrid system, is that one must demonstrate the functional significance of the protein–protein interaction that one has discovered. In fact, before embarking on a library screen with the two-hybrid system, we advise careful consideration of this issue. How one goes about this depends upon the nature of the functions and assays of the particular protein partners. The only general advice we can offer on this point is that mutants with critical functions disrupted are often valuable tools at this stage.

Future prospects: two-hybrid systems beyond yeast

The flexibility of the two-hybrid system has not been observed with prokaryotic systems, where in general, activators require exact positioning with respect to the promoter. Nonetheless, genetic assays in bacteria of more restricted utility have now been developed [42]. Perhaps one day, truly flexible prokaryotic systems will be developed that exploit the multifunctional nature of bacterial enhancer-binding proteins, such as NtrC or NifA [43,44].

It may be that certain protein–protein interactions will be best studied in mammalian cells. Several groups have established two-hybrid systems in mammalian cells in which various indicator genes, such as the chloramphenicol acetyl transferase gene, or selectable markers are activated on interaction of the two proteins [45–49]. The major limitation with these systems is that the efficiencies are not adequate for screening complete mammalian libraries.

Other systems, such as cytoplasmic or extracellular read-outs for protein–protein interactions, are currently under development, or under consideration, as extensions of the two-hybrid technique. With such systems, one can imagine the analysis of proteins that are incapable of nuclear transport or that are membrane associated. Non-transcriptional read-outs would allow one to study proteins with intrinsic transcriptional activation activity, which cannot be studied using the traditional two-hybrid system.

Acknowledgements

We thank Marian Carlson for critical reading of the manuscript and David Figurski for thoughtful discussion. The authors were supported by grants AI 36199 and CA 30488 from the National Institutes of Health. JL is also supported by the James S McDonnell Foundation and the Irma T Hirschl Trust. SPG is an Investigator of the Howard Hughes Medical Institute.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Fields S, Song O: **A novel genetic system to detect protein–protein interactions.** *Nature* 1989, **340**:245–246.

2. Keegan L, Gill G, Ptashne M: **Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein.** *Science* 1986, **231**:699–704.

3. Hope IA, Struhl K: **Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast.** *Cell* 1986, **46**:885–894.

4. Brent R, Ptashne M: **A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor.** *Cell* 1985, **43**:729–736.

5. Ma J, Ptashne M: **Converting a eukaryotic transcriptional inhibitor into an activator.** *Cell* 1988, **55**:443–446.

6. Li B, Fields S: **Identification of mutations in p53 that affect its binding to SV40 large T antigen by using the yeast two-hybrid system.** *FASEB J* 1993, **7**:957–963.

7. Chevray P, Nathans D: **Protein interaction cloning in yeast: identification of mammalian proteins that react with the leucine zipper of Jun.** *Proc Natl Acad Sci USA* 1992, **89**:5789–5793.

8. •• Durfee T, Becherer K, Chen P-L, Yeh S-H, Yang Y, Kilburn AE, Lee W-H, Elledge SJ: **The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit.** *Genes Dev* 1993, **7**:555–569.
See [10••] annotation.

9. •• Zervos A, Gyuris J, Brent R: **Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites.** *Cell* 1993, **72**:223–232.
See [10••] annotation.

10. •• Vojtek AB, Hollenberg SM, Cooper JA: **Mammalian Ras interacts directly with the serine/threonine kinase Raf.** *Cell* 1993, **74**:205–214.
This paper and [8••,9••] describe some interesting applications of the two-hybrid system. In addition, they discuss some of the more popular systems currently in use. Each of these systems has particular strengths, which dictate suitability to the particular application (see text for further details).

11. Dalton S, Treisman R: **Characterization of SAP-1, a protein recruited by serum resonse factor to the c-*fos* serum response element.** *Cell* 1992, **68**:597–612.

12. Fagan R, Flint KJ, Jones N: **Phosphorylation of E2F-1 modulates its interaction with the retinoblastoma gene product and the adenoviral E4 19 kDa protein.** *Cell* 1994, **78**:799–811.

13. Bogerd HP, Fridell RA, Blair WS, Cullen BR: **Genetic evidence that the Tat proteins of human immunodeficiency virus types 1 and 2 can multimerize in the eukaryotic cell nucleus.** *J Virol* 1993, **67**:5030–5034.

14. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K: *Current protocols in molecular biology.* New York: John Wiley and Sons; 1987.

15. Breedon L, Nasmyth K: **Regulation of the HO gene.** *Cold Spring Harbor Symp Quant Biol* 1985, **50**:643–650.

16. Feilotter HE, Hannon GJ, Ruddel CJ, Beach D: **Construction of an improved yeast strain for two hybrid screening.** *Nucleic Acids Res* 1994, **22**:1502–1503.

17. Cress WD, Nevins JR: **Interacting domains of E2F1, DP1, and the adenovirus E4 protein.** *J Virol* 1994, **68**:4213–4219.

18. Kalpana GV, Goff SP: **Genetic analysis of homomeric interactions of human immunodeficiency virus type 1 integrase using the yeast two-hybrid system.** *Proc Natl Acad Sci USA* 1993, **90**:10593–10597.

19. Iwabuchi K, Li B, Bartel P, Fields S: **Use of the two-hybrid system to identify the domain of p53 involved in oligomerization.** *Oncogene* 1993, **8**:1693–1696.

20. Li L, Elledge SJ, Peterson CA, Bales ES, Legerski RJ: **Specific association between the human DNA repair proteins XPA and ERCC1.** *Proc Natl Acad Sci USA* 1994, **91**:5012–5016.

21. Hensey CE, Hong F, Durfee T, Qian YW, Lee EY, Lee WH: **Identification of discrete structural domains in the retinoblastoma protein. Amino-terminal domain is required for its oligomerization.** *J Biol Chem* 1994, **269**:1380–1387.

22. Franke E, Yuan HEH, Bossolt KL, Goff SP, Luban J: **Specificity and sequence requirements for interactions between various retroviral Gag proteins.** *J Virol* 1994, **68**:5300–5305.

23. Tsui S, Schubach WH: **Epstein-Barr virus nuclear protein 2A forms oligomers *in vitro* and *in vivo* through a region required for B-cell transformation.** *J Virol* 1994, **68**:4287–4294.

24. Mosteller RD, Han J, Broek D: **Identification of residues of the H-ras protein critical for functional interaction with guanine nucleotide exchange factors.** *Mol Cell Biol* 1994, **14**:1104–1112.

25. Chien C-T, Bartel PL, Sternglanz R, Fields S: **The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest.** *Proc Natl Acad Sci USA* 1991, **88**:9578–9582.

26. Hardy CFJ, Sussel L, Shore D: **A RAP-1-interacting protein involved in transcriptional silencing and telomere length regulation.** *Genes Dev* 1992, **6**:801–814.

27. Luo Y, Yu H, Peterlin BM: **Cellular protein modulates effects of human immunodeficiency virus type 1 Rev.** *J Virol* 1994, **68**:3850–3856.

28. Iwabuchi K, Bartel PL, Li B, Marraccino R, Fields S: **Two cellular proteins that bind to wild-type but not mutant p53.** *Proc Natl Acad Sci USA* 1994, **91**:6098–6102.

29. • Harper JW, Adami GR, Wei N, Keyomarsi K, Elledge SJ: **The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases.** *Cell* 1993, **75**:805–816.
This paper describes the use of a cycloheximide-resistant yeast strain to facilitate separation of the bait plasmid from the library plasmid. Growth on cycloheximide allows one to conveniently cure the strain of a plasmid containing both the gene encoding the bait protein and the gene conferring cycloheximide sensitivity.

30. Hannon GJ, Demetrick D, Beach D: **Isolation of the Rb-related p130 through its interaction with CDK2 and cyclins.** *Genes Dev* 1993, **7**:2378–2391.

31. Holt KH, Olson L, Moye-Rowley WS, Pessin JE: **Phosphatylinositol 3-kinase activation is mediated by high-affinity interactions between distinct domains within the p110 and p85 subunits.** *Mol Cell Biol* 1994, **14**:42–49.

32. Cuomo CA, Kirch SA, Gyuris J, Brent R, Oettinger MA: **Rch1, a protein that specifically interacts with the RAG-1 recombination-activation protein.** *Proc Natl Acad Sci USA* 1994, **91**:6156–6160.

33. Cortes P, Ye ZS, Baltimore D: **RAG-1 interacts with the repeated amino acid motif of the human homologue of the yeast protein SRP1.** *Proc Natl Acad Sci USA* 1994, **91**:7633–7637.

34. Luban J, Bossolt KA, Franke EK, Kalpana GV, Goff SP: **Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B.** *Cell* 1993, **73**:1067–1078.

35. Shrivastava A, Saleque S, Kalpana GV, Artandi S, Goff SP, Calame K: **Inhibition of transcriptional regulator Yin-Yang-1 by association with c-Myc.** *Science* 1993, **262**:1889–1892.

36. Yang X, Hubbard J, Carlson M: **A protein kinase substrate identified by the two-hybrid system.** *Science* 1992, **257**:680–682.

37. Staudinger J, Perry M, Elledge SJ, Olson EN: **Interactions among vertebrate helix-loop-helix proteins in yeast using the two-hybrid system.** *J Biol Chem.* 1993, **268**:4608–4611.

38. Dunaief JL, Strober BE, Guha S, Khavari PA, Alin K, Luban J, Begemann M, Crabtree GR, Goff SP: **Rb and BRG1 form a complex and cooperate to induce cell cycle arrest.** *Cell* 1995, **79**:in press.

39. Gietz D, St Jean A, Woods RA, Schiestl RH: **Improved method for high efficiency transformation of intact yeast cells.** *Nucleic Acids Res* 1992, **20**:1425.

40. Bendixen C, Gangloff S, Rothstein R: **A yeast mating-selection scheme for detection of protein–protein interactions.** *Nucleic Acids Res* 1994, **22**:1778–1779.

41. Bartel P, Chien C-T, Sternglanz R, Fields S: **Elimination of false positives that arise in using the two-hybrid system.** *Biotechniques* 1993, **14**:920–924.

42. Smith DB, Johnson KS: **Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase.** *Gene* 1988, **67**:31–40.

43. Hu JC, O'Shea EK, Kim PS, Sauer RT: **Sequence requirements for coiled-coils: analysis with l repressor–GCN4 leucine zipper fusion.** *Science* 1990, **250**:1400–1403.

44. Morett E, Segovia L: **The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains.** *J Bacteriol* 1993, **175**:6067–6074.

45. Kustu S, North AK, Weiss DS: **Prokaryotic transcriptional enhancers and enhancer-binding proteins.** *Trends Biochem Sci* 1991, **16**:397–402.

46. Vasavada HA, Ganguly S, Germino FJ, Wang ZX, Weissman SM: **A contingent replication assay for the detection of**

protein–protein interactions in animal cells. *Proc Natl Acad Sci USA* 1991, **88**:10686–10690.

47. Fearon ER, Finkel T, Gillison ML, Kennedy SP, Casella JF, Tomaselli GF, Morrow JS, Van Dang C: **Karyoplasmic interaction selection strategy: a general strategy to detect protein–protein interactions in mammalian cells.** *Proc Natl Acad Sci USA* 1992, **89**:7958–7962.

48. Chakraborty T, Martin JF, Olson EN: **Analysis of the oligomerization of myogenin and E2A products *in vivo* using a two-hybrid system.** *J Biol Chem.* 1992, **267**:17498–17501.

49. Takacs AM, Das T, Banerjee AK: **Mapping of interacting domains between the nucleocapsid protein and the phosphoprotein of vesicular stomatitis virus by using a two-hybrid system.** *Proc Natl Acad Sci USA* 1993, **90**:10375–10379.

J Luban, Department of Medicine, Columbia University, College of Physicians and Surgeons, 630 West 168th Street, New York, New York 10032, USA.

SP Goff, Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biophysics, Columbia University, College of Physicians and Surgeons, 630 West 168th Street, New York, New York 10032, USA.

Author for correspondence: SP Goff.