MICROBIOLOGICAL REVIEWS, Mar. 1995, p. 94–123 Vol. 59, No. 1
0146-0749/95/$04.00+0

Protein-Protein Interactions: Methods for Detection and Analysis

ERIC M. PHIZICKY[1]* AND STANLEY FIELDS[2]

Department of Biochemistry, University of Rochester Medical School, Rochester, New York 14642,[1] and Department of Molecular Genetics and Microbiology, State University of New York at Stony Brook, Stony Brook, New York 11794[2]

* Corresponding author. Mailing address: Department of Biochemistry, University of Rochester Medical School, 601 Elmwood Ave., Box 607, Rochester, NY 14642. Phone: (716) 275-7268. Fax: (716) 271-2683.

INTRODUCTION

Protein-protein interactions are intrinsic to virtually every cellular process. Any listing of major research topics in biology—for example, DNA replication, transcription, translation, splicing, secretion, cell cycle control, signal transduction, and intermediary metabolism—is also a listing of processes in which protein complexes have been implicated as essential components. In consequence, the analysis of the proteins in these complexes is no longer the exclusive domain of biochemists; geneticists, cell biologists, developmental biologists, molecular biologists, and biophysicists have by necessity all gotten into the act. We attempt in this review to summarize both classical and recent methods to identify proteins that interact and to assess the strengths of these interactions.

Proteins that are composed of more than one subunit are found in many different classes of proteins. Some of the best-characterized multisubunit proteins are those that, as originally purified, contained two or more different components. These include classical proteins such as hemoglobin, tryptophan synthetase, aspartate transcarbamylase, core RNA polymerase, Qβ-replicase, and glycyl-tRNA synthetase. Since these proteins purified as multisubunit complexes, their protein-protein interactions were self-evident.

Other well-known examples of multisubunit proteins include much more complicated assemblies of polypeptides. These include metabolic enzymes such as the pyruvate dehydrogenase and α-ketoglutarate dehydrogenase complexes, the DNA replication complex of *Escherichia coli* and other organisms, the bacterial flagellar apparatus, the nuclear pore complex, and the tail assembly of bacteriophage T4. Also included in this group are ribonucleoprotein complexes, such as the signal recognition particle of the glycosylation pathway, small nuclear ribonucleoproteins of the spliceosome, and the ribosome itself. Although some of the subunits of these protein complexes are not tightly bound, activity is associated with a large structure that in many cases is called a protein machine (5).

There are also a large number of transient protein-protein interactions, which in turn control a large number of cellular processes. All modifications of proteins necessarily involve such transient protein-protein interactions. These include the interactions of protein kinases, protein phosphatases, glycosyl transferases, acyl transferases, proteases, etc., with their substrate proteins. Such protein-modifying enzymes encompass a large number of protein-protein interactions in the cell and regulate all manner of fundamental processes such as cell growth, cell cycle, metabolic pathways, and signal transduction. Transient protein-protein interactions are also involved in the recruitment and assembly of the transcription complex to specific promoters, the transport of proteins across membranes, the folding of native proteins catalyzed by chaperonins, individual steps of the translation cycle, and the breakdown and re-formation of subcellular structures during the cell cycle (such as the cytoplasmic microtubules, the spindle apparatus, nuclear lamina, and the nuclear pore complex). Transient complexes are much more difficult to study, because the proteins or conditions responsible for the transient reaction have to be identified first. Part of the goal of this review is to describe recent methods and developments that have allowed their identification and characterization.

Protein-protein interactions can have a number of different measurable effects. First, they can alter the kinetic properties of proteins. This can be reflected in altered binding of substrates, altered catalysis, or (as first enunciated by Monod et al. [153]) altered allosteric properties of the complex. Thus, the interaction of proliferating-cell nuclear antigen with DNA polymerase δ alters the processivity of the polymerase (174), the interaction of succinate thiokinase and α-ketoglutarate dehydrogenase lowers the K_m for succinyl coenzyme A by 30-fold (171), and the cooperative binding of oxygen to hemoglobin and the allosteric regulation of aspartate transcarbamylase are regulated by interactions of the protomers. Second, protein-protein interactions are one common mechanism to allow for substrate channeling. The paradigm for this type of complex is tryptophan synthetase from *Neurospora crassa*. It is a complex of two subunits, each of which carries out one of the two steps of reaction (formation of indole from indole 3-glycerol phosphate, followed by conversion of indole to tryptophan). The intermediate indole is noncovalently bound, but it is preferentially channeled to form tryptophan (241). Many similar examples of metabolic channeling have been demonstrated, both between different subunits of a complex and between different domains of a single multifunctional polypeptide (see reference 208 for a review). Third, protein-protein interactions can result in the formation of a new binding site. Thus, an ADP site forms at the interface of the α and β subunits of *Escherichia coli* F_1-ATPase (228), yeast hexokinase binds one ATP molecule at the interface of the asymmetric homodimer (209), and phosphofructokinase from *Bacillus stearothermophilus* binds both fructose 6-phosphate and ADP at the interface between subunits (60). Fourth, protein-protein interactions can inactivate a protein; this is the case with the interaction of phage P22 repressor with its antirepressor (213), with the interaction of trypsin with trypsin inhibitor (221), and with the interaction of

phage T7 gene 1.2 protein with *E. coli* dGTP triphosphohydrolase (156). Fifth, protein-protein interactions can change the specificity of a protein for its substrate; thus, the interaction of lactalbumin with lactose synthase lowers the K_m for glucose by 1,000-fold (95), and the interaction of transcription factors with RNA polymerase directs the polymerase to different promoters.

Klotz et al. (116) enumerated four advantages of multisubunit proteins relative to a single large protein with multiple sites. First, it is much more economical to build proteins from simpler subunits than to require multiple copies of the coding information to synthesize oligomers. Thus, for example, actin filaments and virus coats are much more simply assembled from monomers than by translation of a large polyprotein of repeated domains. Similarly, it is much more convenient to have one gene encoding a protein with different interacting partners, such as some of the eukaryotic RNA polymerase subunits, than to have the gene for that subunit reiterated for each different polymerase. Second, translation of large proteins can cause a significant increase in errors in translation; if such errors cause a lack of activity, they are much more economically eliminated by preventing assembly of that subunit into the complex than by eliminating the whole protein. Third, multisubunit assemblies allow for synthesis at one locale, followed by diffusion and assembly at another locale; this allows for both faster diffusion (since the monomers are smaller) and compartmentalization of activity (if assembly is required for activity). Fourth, homooligomeric proteins, if they have an advantage over monomers, are easily selected in evolution if the oligomers interact in an antiparallel arrangement; in this case, a single-amino-acid change that increases interaction potential has effects at two such sites.

Another advantage of multisubunit complexes is the ability to use different combinations of subunits to alter the magnitude or type of response. Thus, for example, adult hemoglobin ($\alpha 2\beta 2$) and fetal hemoglobin ($\alpha 2\gamma 2$) are each composed of heterooligomers with a common α subunit; differences in the binding of oxygen in these hemoglobins allow oxygen to be readily passed from mother to fetus. Other examples include the oligomerization of Jun with Fos or with itself, which results in distinct activities in transcription because the different dimers bend DNA in opposite directions (114); the interaction of TATA-binding protein with the transcription apparatus of RNA polymerase I, II, or III, in which TATA-binding protein plays different roles (235); the interactions of microtubules with the large set of proteins to which they bind (113), not all of which bind at the same time; the interaction of different transcription factors with core RNA polymerase in both eukaryotes and prokaryotes to direct transcription of different genes; and the interaction of retinoblastoma (Rb) protein with viral oncoproteins and other cellular proteins (31, 32).

Protein-protein interactions may be mediated at one extreme by a small region of one protein fitting into a cleft in another protein and at another extreme by two surfaces interacting over a large area. Examples of the first case include the large class of protein-protein interactions that involve a domain of a protein interacting tightly with a small peptide. The paradigm for this type of interaction is that of specific Src homology 2 (SH2) domains with specific small peptides containing a phosphotyrosyl residue. This interaction occurs with a dissociation constant as low as nM and is due to a specific binding pocket in SH2 domains not unlike a classical substrate-binding pocket (64, 205, 224, 225). Many other examples of domains that bind small peptides with affinities in the nanomolar to molar range have been described. The paradigm for the second case, i.e., surfaces that interact with each other over large areas, is that of the leucine zipper, in which a stretch of α-helix forms a surface that fits almost perfectly with another α-helix from another subunit protein (59, 161; also see reference 4). Binding also occurs in the nanomolar range for such interactions (196). Other interactions may occur through intermediate-sized complementary surfaces.

It is evident that protein-protein interactions are much more widespread than once suspected, and the degree of regulation that they confer is large. To properly understand their significance in the cell, one needs to identify the different interactions, understand the extent to which they take place in the cell, and determine the consequences of the interaction. This review is intended to supply an overview of three aspects of protein-protein interactions. First, we briefly describe a number of physical, molecular biological, and genetic approaches that have been used to detect protein-protein interactions. Second, we describe several experimental approaches that have been used to evaluate the strength of protein-protein interactions. Third, we describe three well-characterized domains that are responsible for protein-protein interactions in a number of different proteins. As the literature on this topic is vast, we have not attempted to conduct an exhaustive review. Rather, we hope that this article serves as a journeyman's guide to protein-protein interactions.

The first and still the most comprehensive review on protein-protein interactions is that of Klotz et al. (116). This review contains a survey of the subunit composition and binding energies of all oligomeric proteins that had been identified at the time, as well as a discussion of the geometry of interactions and an excellent discussion of the influence of binding constants, concentrations, and cooperativity parameters on the population of oligomers. A good discussion of channeling and compartmentation is found in the monograph by Friedrich on quaternary structure (70) and the article by Srere (208). The review by Eisenstein and Schachman (57) contains an interesting discussion of the functional roles of subunits of oligomeric proteins and of approaches used to determine whether the monomers of oligomeric proteins are active. Also of interest is the discussion of proteins as machines (5) and a discussion of protein size and composition (78).

PHYSICAL METHODS TO SELECT AND DETECT PROTEINS THAT BIND ANOTHER PROTEIN

Protein Affinity Chromatography

A protein can be covalently coupled to a matrix such as Sepharose under controlled conditions and used to select ligand proteins that bind and are retained from an appropriate extract. Most proteins pass through such columns or are readily washed off under low-salt conditions; proteins that are retained can then be eluted by high-salt solutions, cofactors, chaotropic solvents, or sodium dodecyl sulfate (SDS) (Fig. 1). If the extract is labeled in vivo before the experiment, there are two distinct advantages: labeled proteins can be detected with high sensitivity, and unlabeled polypeptides derived from the covalently bound protein can be ignored (these might be either proteolytic fragments of the covalently bound protein or subunits of the protein which are not themselves covalently bound). This method was first used 20 years ago to detect phage and host proteins that interacted with different forms of *E. coli* RNA polymerase (177). Proteins that were retained by an RNA polymerase-agarose column (which was shown to be enzymatically active) but not by a control column coupled with bovine serum albumin were judged as interacting candidates. The interactions were substantiated in two ways. First, the

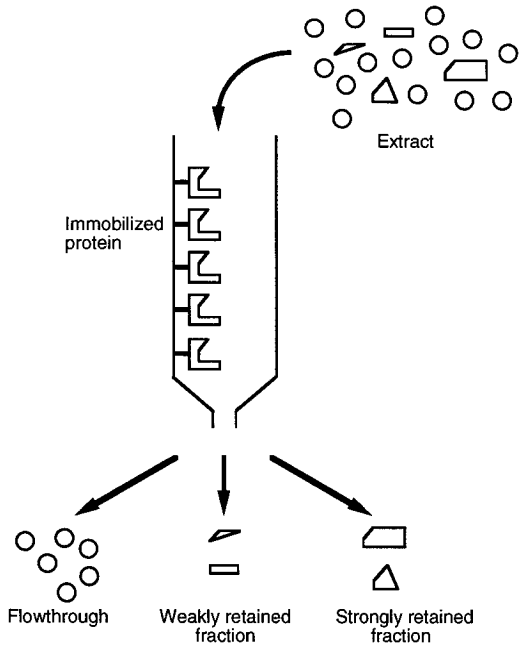

FIG. 1. Protein affinity chromatography. Extract proteins are passed over a column containing immobilized protein. Proteins that do not bind flow through the column, and ligand proteins that bind are retained. Strongly retained proteins have more contacts with the immobilized protein than do those that are weakly retained.

interaction of T7 0.3 protein with RNA polymerase was confirmed by coimmunoprecipitation of the 0.3 protein with RNA polymerase antibody. Second, the interaction of T4 proteins with RNA polymerase was shown to depend on the form of RNA polymerase on the column: one T4 protein interacted with core RNA polymerase and T4-modified RNA polymerase but not with RNA polymerase holoenzyme, and another interacted only with the T4-modified polymerase. The phage proteins that bound RNA polymerase were identified by their absence in appropriate T4 and T7 mutants.

Similar methods have been used, particularly by the laboratories of J. Greenblatt and B. Alberts, to identify many other protein-protein interactions. Two excellent reviews on the topic, which cover many of the details of coupling and a number of strategic considerations, have been published (69, 145).

Candidate proteins can be coupled directly to commercially available preactivated resins as described by Formosa et al. (69). Alternatively, they can be tethered noncovalently through high-affinity binding interactions. Thus, Beeckmans and Kanarek (14) demonstrated an interaction between fumarase and malate dehydrogenase by immobilizing the test enzyme with antibody bound to protein A-Sepharose, as well as by direct covalent coupling of the test enzyme to Sepharose. Some of the important considerations of a successful binding experiment are elaborated below.

Purity of the coupled protein and use of protein fusions. An essential requirement for a successful protein affinity chromatography experiment is pure protein; otherwise, any interacting protein that is detected might be binding to a contaminant in the preparation. Greenblatt and Li (80) did two experiments to establish that core RNA polymerase bound to NusA on the column rather than to a contaminant in the NusA preparation. First, they demonstrated that a fully active NusA variant protein, which presumably contained different amounts of various contaminants (since it eluted at different positions in columns used to purify it), still bound core RNA polymerase; second, they demonstrated by independent experiments that the complex contained equimolar amounts of NusA protein and core RNA polymerase.

The easiest way to obtain pure protein, if the gene is available, is through the use of protein fusions. Several such systems have been described; in each case, the protein of interest (or a domain of the protein) is fused to a protein or a domain that can be rapidly purified on the appropriate affinity resin. The most common such fusion contains glutathione *S*-transferase (GST), which can be purified on glutathione-agarose columns (202). Other fusions in common use include *Staphylococcus* protein A, which can be purified on columns bearing immunoglobulin G; oligohistidine-containing peptides, which can be purified on columns bearing Ni^{2+}; the maltose-binding protein, which can be purified on resins containing amylose; and dihydrofolate reductase, which can be purified on methotrexate columns. (Other common protein fusions which add an epitope for the influenza virus hemagglutinin [12CA5] or c-Myc are also in common use and are used most often for coimmunoprecipitation [see the section on immunoprecipitation, below].)

Purified fusion proteins are used in two ways to detect interactions on affinity columns. First, the protein is covalently coupled to the resins in the usual way, as was done by Mayer et al. (139) to detect a tyrosine-phosphorylated protein that bound to the SH2 domain of Abl tyrosine kinase and by Weng et al. (232) to demonstrate that the SH3 domain of c-Src binds paxillin. Second, the purified fusion proteins can be noncovalently bound to the beads and then mixed with an appropriate extract or protein. This was done by Zhang et al. (248) to demonstrate an interaction of the N-terminal portion of c-Raf with Ras, by Flynn et al. (68) to detect the binding of an actin filament-associated protein to Src-SH3/SH2, and by Hu et al. (99) to demonstrate the binding of the SH2 domain of the p85 subunit of phosphatidylinositol 3-kinase to two different growth factor receptors.

Influence of modification state. The interactions of many proteins with their target proteins often depends on the modification state of one or both of the proteins (mostly by phosphorylation). Thus, the recognition of Rb protein by the transcription factor E2F and by the transforming proteins simian virus 40 large T antigen, human papillomavirus-16 E7, and adenovirus E1A is more efficient with underphosphorylated than phosphorylated Rb (132, 133, 240). Conversely, SH2 domains of proteins, for example, recognize tyrosine phosphorylated substrates several orders of magnitude more efficiently than they do their nonphosphorylated counterparts (64). Protein-protein interactions that require a posttranslationally modified protein for interaction are not detected if the protein is purified by the use of expression vectors in cells in which the protein is not properly modified. A means to circumvent this problem is to use GST fusion vectors to express proteins in host cells more related to their origin. Thus, the interaction of bovine papillomavirus E5 oncoprotein with an α-adaptin-like molecule was confirmed by addition of beads to extracts of NIH 3T3 cells that were expressing the GST-E5 fusion (38). Similarly, a yeast GST vector that allows regulated expression of yeast GST fusion proteins has been described (148).

Retention of native structure of the coupled protein. Failure to detect an interacting protein can result from inactivation of the protein during coupling. Ideally, coupling would immobilize a protein or a complex by randomly tethering it to the matrix through one covalent bond. For example, binding of *E. coli* proteins to immobilized λ N protein occurred only when

the cyanogen bromide (CNBr)-activated residues on the matrix were partially inactivated before coupling; this was attributed to the large number of lysine residues in λ N protein and the generation of multiple (and denaturing) covalent bonds between λ N protein and the matrix if the concentration of CNBr-activated matrix sites was too high (80). Therefore, determining that the coupled protein has retained its native structure is an important control, when possible. With some proteins, such as RNA polymerase from *E. coli*, activity could be detected when the coupled protein was assayed on the matrix (177). With others, such as filamentous actin (F-actin) columns, the desired polymerized form was stabilized with phalloidin (or by chemical cross-linking), and the proteins that bound F-actin were shown not to bind monomeric actin (14). Similarly, microtubule columns were stabilized with taxol (113).

Native protein structure also depends on all subunits of a complex being present in the coupled resin. This can be assessed by SDS elution of a sample of the resin and comparison of the subunit composition of the eluted material with that of the starting material. In the case of *E. coli* RNA polymerase, all the components of the enzyme were still present (177). In the case of mammalian RNA polymerase II, one of the subunits did not reproducibly remain after coupling (206).

Concentration of the coupled protein. To detect interactions efficiently, the concentration of protein covalently bound to the column has to be well above the K_d of the interaction. Thus, for the detection of weak protein-protein interactions, the concentration of bound protein should be as high as possible. Weak interactions can be completely missed on columns with lower concentrations of coupled protein, even if they contain correspondingly larger amounts of resin to maintain the same total amount of bound protein (see the sections on importance of characterization of the binding interaction and on binding to immobilized proteins, below, for a discussion of this point).

Amount of extract applied. The amount of extract applied to the column can be critical for two opposing reasons. If too little extract is applied and the protein that binds is present at low concentration, too little protein will be retained to be detected, even if it binds with high affinity and is labeled with ^{35}S (see, for example, reference 206). Conversely, if too much protein is applied, competition among potential ligands may result in failure to detect minor species. This was observed by Miller and Alberts (144) in looking for minor protein species that interact with F-actin.

Other considerations. There are four distinct advantages of protein affinity chromatography as a technique for detecting protein-protein interactions. First, and most important, protein affinity chromatography is incredibly sensitive. With appropriate use (high concentrations of immobilized test protein), it can detect interactions with a binding constant as weak as 10^{-5} M (69) (see the section on binding to immobilized protein, below). This limit is within range of the weakest interaction likely to be physiologically relevant, which we estimate to be in the range of 10^{-3} M (see the section on limits of binding-constant considerations, below). Second, this technique tests all proteins in an extract equally; thus, extract proteins that are detected have successfully competed for the test protein with the rest of the population of proteins. Third, it is easy to examine both the domains of a protein and the critical residues within it that are responsible for a specific interaction, by preparing mutant derivatives (38, 216). Fourth, interactions that depend on a multisubunit tethered protein can be detected, unlike the case with protein blotting.

One potential problem derives from the very sensitivity of the technique. Since it detects interactions that are so weak, independent criteria must be used to establish that the interaction is physiologically relevant. Detection of a false-positive signal can arise for a number of other reasons. First, the protein may bind the test protein because of charge interactions; for this reason, it is desirable to use a control column with approximately the same ionic charges. Second, the proteins may interact through a second protein that interacts with the test protein; although interesting in itself, the interaction may not be direct. Third, the proteins may interact with high specificity even though they never encounter one another in the cell. The most famous example of this type is the high affinity of actin for DNase I (125).

For all of these reasons, the prudent course is to independently demonstrate the interaction in vitro or, if possible, in vivo. Cosedimentation was used to confirm the interaction of RAP 72 (now known as RAP 74) and RAP 30 with RNA polymerase II (206), NusA protein with core RNA polymerase (80), and NusB protein with ribosomal protein S10 (138). In other cases, more biological criteria were used. For example, antibodies were generated against many of the proteins that interacted with F-actin (but not monomeric G-actin) on columns, and these were used to demonstrate that more than 90% of the corresponding proteins were localized with an actin-like distribution during mitosis of *Drosophila* embryos at the syncytial blastoderm stage of development (144). The identification of three yeast actin-binding proteins was confirmed in three separate ways: one of the proteins was shown to correspond to the yeast analog of myosin by virtue of a shared epitope; another protein colocalized with actin cables and cortical actin patches, and overproduction of the third protein caused a reorganization of the actin cytoskeleton (53). In the identification of microtubule-associated proteins, two criteria were used to demonstrate the authenticity of the results (113). First, antibodies for 20 of the 24 candidate microtubule-associated proteins stain various parts of microtubule structures of *Drosophila* embryos during the cell cycle. Second, many (but not all) of the microtubule-associated proteins isolated on microtubule affinity columns are the same as those isolated by traditional cosedimentation methods of Vallee and Collins (219).

Failure to detect an interaction can occur for a number of technical reasons, described above. A false-negative result can arise for two additional reasons: the interacting protein may not be able to exchange with another protein to which it is binding, or the two proteins may not be able to interact both with each other and with the resin.

Protein affinity chromatography does not always yield answers corresponding to other approaches. For reasons that are unclear, a large number of proteins were detected by probing SDS-polyacrylamide gel electrophoresis (PAGE) gels with a GST fusion of the SH2 domain of Abl tyrosine kinase, but only a couple of proteins were detected on columns coupled with this protein (139). Similarly, a specific protein was detected on F-actin columns stabilized by suberimidate cross-linking but not with phalloidin (144). Finally, G-actin interacting proteins are very difficult to detect with columns of G-actin, although such columns bind DNase I; by contrast, DNase I columns can be used to detect such G-actin interactions (24).

Affinity Blotting

In a procedure analogous to the use of affinity columns, proteins can be fractionated by PAGE transferred to a nitrocellulose membrane, and identified by their ability to bind a protein, peptide, or other ligand. This method is similar to immunoblotting (Western blotting), which uses an antibody as

the probe. Complex mixtures of proteins, such as total-cell lysates, can be analyzed without any purification. Therefore, this method has been particularly useful for membrane proteins, such as cell surface receptors (see reference 207 for a discussion). Cell lysates can also be fractionated before gel electrophoresis to increase the sensitivity of the method for detecting interaction with rare proteins.

Considerations in affinity blotting include the biological activity of the proteins on the membrane, the preparation of the protein probe, and the method of detection. Denaturing gels, which are run in the presence of SDS and sulfhydryl reducing agents, will inactivate most proteins and separate subunits of a complex. These denaturants are removed during the blotting procedure, which allows many proteins to recover (or partially recover) activity. However, if biological activity is not recoverable, the proteins can be fractionated by a nondenaturing gel system. This method eliminates the problem of regeneration of activity and allows the detection of binding in cases when binding requires the presence of a protein complex.

The protein probe can be prepared by any one of several procedures, and, as with affinity columns, the recent advent of fusing tags to the protein has greatly facilitated this purification. Synthesis in *E. coli* with a GST fusion, epitope tag, or other affinity tag is most commonly used. The protein of interest can then be radioactively labeled, biotinylated, or used in the blotting procedure as an unlabeled probe that is detected by a specific antibody. Vectors that incorporate into the protein a short amino acid sequence recognized by the heart muscle cyclic AMP (cAMP)-dependent protein kinase provide another convenient means for in vitro labeling (18).

One example of affinity blotting is the study of calmodulin-binding proteins (77). Calmodulin can be ^{125}I labeled and used either to probe a gel strip directly or to probe a nitrocellulose membrane after transfer of fractionated proteins. Because the extent of renaturation of calmodulin-binding proteins is variable, the assay is not quantitative. False-positive results have been detected in which a basic sequence binds calmodulin, although generally this binding is Ca^{2+} independent. A major advantage of this technique is that in the analysis of a multimeric protein that binds calmodulin, the precise binding polypeptide can be readily identified by affinity blotting with calmodulin. Using a combination of genetic approaches, Geiser et al. (73) identified the spindle pole body component Spc110 (Nuf1) as interacting with yeast calmodulin and then used affinity blotting to demonstrate that labeled calmodulin could directly detect a GST-Spc110 fusion transferred to a blot after fractionation by SDS-PAGE.

Affinity blotting has been widely used in studies of the association of the regulatory subunit of the type II cAMP-dependent protein kinase with numerous specific anchoring proteins (reviewed in reference 29). Two-dimensional procedures of isoelectric focusing followed by SDS-PAGE have been used to increase the separation of these anchoring proteins. As a control in some of these experiments, a mutant of the regulatory subunit that is deleted for the first 23 residues did not detect any anchoring proteins.

Immunoprecipitation

Coimmunoprecipitation is a classical method of detecting protein-protein interactions and has been used in literally thousands of experiments. The basic experiment is simple. Cell lysates are generated, antibody is added, the antigen is precipitated and washed, and bound proteins are eluted and analyzed. Several sources of material are in wide use. The antigen used to make the antibody can be purified protein (either from the natural tissue or organism or purified after expression in another organism) or synthetic peptide coupled to carrier, and the antibody can be polyclonal or monoclonal. Alternatively, the protein can carry an epitope tag for which commercially available antibodies are available (12CA5 and c-Myc are in common use) or a protein tag (such as GST) for which beads are available to rapidly purify the GST fusion protein and any copurifying proteins. Glutathione-agarose beads were used, for example, to detect and characterize a GTP-dependent interaction of Ras and Raf (227) and to demonstrate that the v-Crk SH2 domain binds the phosphorylated form of paxillin (16). The GST fusion immunoprecipitates a 70-kDa protein that reacts with anti-paxillin antibody and with anti-phosphotyrosine antibody; moreover, anti-paxillin immunoprecipitates a protein that reacts with anti-Crk antibody but only under conditions when the paxillin is phosphorylated.

Several criteria are used to substantiate the authenticity of a coimmunoprecipitation experiment. First, it has to be established that the coprecipitated protein is precipitated by the antibody itself and not by a contaminating antibody in the preparation. This problem is avoided by the use of monoclonal antibodies. Polyclonal antibodies are usually preadsorbed against extracts lacking the protein to remove contaminants or are prepurified with authentic antigen. Peptide-derived antisera (which are usually made by coupling of the peptide to a carrier protein) are usually preadsorbed against the protein that was coupled, to remove antibody against the carrier, in addition to the usual purification to remove contaminating antibody. Second, it has to be established that the antibody does not itself recognize the coprecipitated protein. This can be accomplished by demonstrating persistence of coprecipitation with independently derived antibodies, ideally with specificities toward different parts of the protein. Alternatively, it can sometimes be demonstrated that coprecipitation requires the presence of the antigen; cell lines, growth conditions, or strains that lack the protein cannot coprecipitate the protein unless the antigen is added. In certain cases, it can also be shown that antibody generated against the coprecipitated protein will coprecipitate the original antigen. Third, one would like to determine if the interaction is direct or proceds through another protein that contacts both the antigen and the coprecipitated protein. This is usually addressed with purified proteins, by immunological or other techniques. Fourth, and most difficult, is determining that the interaction takes place in the cell and not as a consequence of cell lysis. Such proteins ought to colocalize, or mutants ought to affect the same process.

A particularly good example of this technique is the demonstration that adenovirus E1A protein interacts with Rb protein. A mixture of monoclonal antibodies against E1A coimmunoprecipitated a discrete set of five polypeptides (and some smaller ones) from a cell line expressing E1A, including a particularly abundant one of 110 kDa (84). Four lines of evidence supported the claim that the 110-kDa polypeptide was forming a complex with E1A protein. First, coprecipitation was not specific to a single antibody; three independent monoclonal antibodies against E1A protein coimmunoprecipitated this protein. Second, these antibodies did not themselves recognize or immunoprecipitate the native or denatured 110-kDa protein, although they recognized and immunoprecipitated native and denatured E1A protein. Third, coprecipitation required E1A protein; the 110-kDa polypeptide could be immunoprecipitated from HeLa extracts (which do not contain E1A protein) only if a source of E1A protein was added. Fourth, the complex could be detected independently in crude lysates; a subpopulation of E1A protein in lysed cells sedimented at 10S rather than at 4S, and this subpopulation contained coimmu-

noprecipitable 110-kDa protein. A similar 110-kDa protein (as well as a similar set of other proteins) was also identified with antipeptide antisera against E1A protein (242). Two separate antisera (one against an amino-terminal peptide and one against a carboxyl-terminal peptide) each coprecipitated the 110-kDa polypeptide, and coprecipitation was prevented either with an excess of the corresponding E1A peptide antigen or in cell extracts lacking E1A protein.

Subsequent studies established that this 105- to 110-kDa polypeptide was the Rb gene product (236). To this end, monoclonal antibodies against the 110-kDa protein were prepared by immune purification of the 110-kDa protein. The resulting antibody coprecipitated E1A protein, just as anti-E1A coprecipitated the 110-kDa protein. Since the 110-kDa protein was the same size as Rb protein, and since it was present in a wide variety of cell lines but not in cell lines known to contain deletions of the Rb gene, it seemed likely that the 110-kDa protein was Rb protein. This was proved by using anti-Rb peptide antibodies against different regions of Rb in three experiments. First, 110-kDa protein precipitated with anti-110-kDa antibody comigrated and had the same partial peptide map as that precipitated with anti-Rb antibody. Second, 110-kDa protein precipitated with anti-E1A antibody could be detected in immunoblots with two different anti-Rb antibodies, and this detection was inhibited by the corresponding peptide antigen. Third, anti-110-kDa antibody could immunoprecipitate Rb protein synthesized in vitro.

When coimmunoprecipitation is performed with unsupplemented crude lysates, as is often the case, this technique has four distinct advantages. First, like protein affinity chromatography, it detects the interactions in the midst of all the competing proteins present in a crude lysate; therefore, the results from this sort of experiment have a built-in specificity control. Second, both the antigen and the interacting proteins are present in the same relative concentrations as found in the cell; therefore, any artificial effects of deliberate overproduction of the test protein are avoided. Third, elaborate complexes are already in their natural state and can be readily coprecipitated; such complexes might otherwise be difficult to assemble in vitro. Fourth, the proteins are present in their natural state of posttranslational modification; therefore, interactions that require phosphorylation (or lack of phosphorylation) are more realistically assessed. Two disadvantages are also apparent. First, coimmunoprecipitating proteins do not necessarily interact directly, since they can be part of larger complexes. For example, the coprecipitation of E1A and p60 (now known to be cyclin A) (84) occurs indirectly; E1A interacts with p107 (237), and p107 interacts with cyclin A (61, 62). Similarly, coprecipitation of Rb protein with E2F probably occurs through another protein (92, 179). Second, coprecipitation is not as sensitive as other methods, such as protein affinity chromatography, because the concentration of the antigen is lower than it is in protein affinity chromatography. This can be overcome by deliberately adding an excess of the antigen to the crude lysates to drive complex formation, as was done to detect a 46-kDa protein that competed with simian virus 40 T antigen for Rb protein (100). It can also be overcome by covalently cross-linking the proteins prior to immunoprecipitation (48) see the section on cross-linking, below). These alterations of course perturb the natural conditions that make immunoprecipitation an attractive method.

Cross-Linking

Cross-linking is used in two ways to deduce protein-protein interactions. First, it is used to deduce the architecture of proteins or assemblies that are readily isolated intact from the cell. Second, it is used to detect proteins that interact with a given test protein ligand by probing extracts, whole cells, or partially purified preparations.

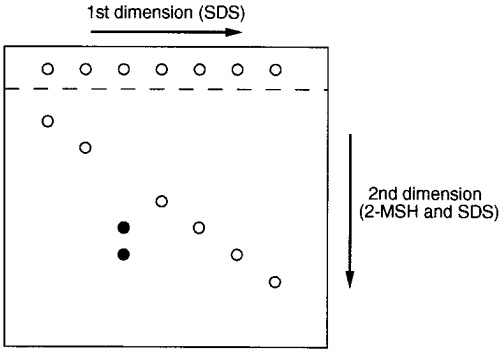

FIG. 2. Two-dimensional gels to identify cross-linked proteins in a complex. Proteins that are not cross-linked have the same mobility in both dimensions of the SDS gel and form a diagonal. Proteins that are cross-linked migrate slowly in the first dimension; after cleavage of the cross-link with mercaptoethanol (2-MSH), these proteins migrate at their native positions in the second dimension and are off the diagonal.

Determination of architecture. The classical method of identifying interacting partners in a purified protein complex involves analysis by two-dimensional gels (Fig. 2). The procedure involves three steps. First, the complex is reacted with a cleavable bifunctional reagent of the form RSSR′, and the R and R′ groups react with susceptible amino acid side chains in the protein complex. This reaction forms adducts of the form P-RSSR′-P′. Second, the proteins are fractionated on an SDS-gel in the absence of reducing agents. The gel separates the proteins based on molecular weight, and cross-linked proteins of the form P-RSSR′-P′ migrate as species of greater molecular weight. Third, a second dimension of the SDS-gel is run after treatment of the gel with a reducing agent to cleave the central S—S bond. Un-cross-linked species align along the diagonal, because their molecular weights do not change after reduction. Cross-linked proteins migrate off the diagonal because they migrated as P-RSSR′-P′ in the first dimension and as molecules of the form P-RSH and P′-R′SH in the second dimension. The cross-links are identified by their size, which corresponds to that of the un-cross-linked species P and P′. This method has been discussed at a practical step-by-step level by Traut et al. (215).

Cross-linking has been used to study the architecture of multienzyme complexes such as CF_1-ATPase (7) and *E. coli* F_1-ATPase (21). It has also been used to study the structure of much more complicated structures like the ribosome (41, 215). Since these structures are complex, the corresponding cross-linking pattern is necessarily complex. Furthermore, as might be expected, different patterns are sometimes obtained as the reactive group is changed and as the distance between the reactive groups is altered (41, 215). Several approaches have been taken to simplify the cross-linking patterns resulting from these experiments. In one approach, the proteins are prefractionated on urea-acrylamide gels or on CM-Sepharose before diagonal electrophoresis (41, 217). A second approach involves running two-dimensional gels without cleaving the cross-link, followed by elution of individual species, cleavage of the cross-link, and resolution of the resulting proteins on a third gel (22). A third approach involves the use of antibody to identify cross-linked partners after the use of appropriate gels (180, 212). Transfer of the gels followed by immunoblotting allows one to

unequivocally identify each member of a cross-linked pair. Since this method is so powerful, one-dimensional gels often suffice and noncleavable cross-linking reagents are easily used. Since immunoblotting is also very sensitive, one can take care to limit cross-linking to acceptably low levels.

Detection of interacting proteins. (i) Detection in vivo. Cross-linking in vivo can be accomplished with membrane-permeable cross-linking reagents followed by immunoprecipitation of the ligand protein. This method was used to detect a 60-kDa protein that interacts with Ras (48). Immunoprecipitation of this protein required both immune sera and cross-linking and was inhibited when excess Ras was added before immunoprecipitation. Since the cross-linked protein could be released from the immune complex by cleavage of the cross-link with dithiothreitol (but not by incubation of the immune complex in buffer), it was truly cross-linked. Since pretreatment of the cross-linking reagent with excess amino groups inhibited cross-linking but excess amino groups did not inhibit cross-linking if cells were lysed in their presence, cross-linking must have occurred in vivo. The complex was reproducibly increased after mitogenic stimulation and could be detected in cells producing normal amounts of Ras. This experiment makes another point: at least in these experiments, cross-linking before immunoprecipitation is a more sensitive technique than immunoprecipitation alone.

(ii) Detection in vitro. The addition of an isolated protein or a peptide to a complex system offers a huge potential for precise and powerful cross-linking methods. Several different such methods have been used to detect interacting proteins.

(a) Labeled peptide or protein. Detection of cross-linking partners is incomparably cleaner if the protein or peptide is labeled before cross-linking, because there is only one source of labeled material. For example, ^{125}I-labeled gamma interferon was used to detect receptors that were cross-linked (192), and in vivo labeled interleukin-5 was purified before cross-linking to detect interacting receptors (147).

Proteins are also routinely labeled in vitro with [^{35}S]methionine during translation, and this was followed by cross-linking and by immunoprecipitation to detect protein interactions. This has been done, for example, to detect interaction of preprolactin and pre-β-lactamase with signal sequence receptor and translocation chain-associating protein during glycosylation (79) and to detect mitochondrial import proteins in contact with translocation intermediates (195, 204).

(b) Photoaffinity cross-linking with labeled cross-linking reagent. A particularly useful reagent is the Denny-Jaffee reagent, a cleavable heterobifunctional photoactivatable cross-linking reagent that is labeled on the photoactivated moiety (49). This reagent can be coupled to an isolated protein, which is then incubated in an appropriate extract and photoactivated to cross-link nearby proteins. Since the label is on the photoactivatable moiety of the cross-linking reagent, it is transferred to the cross-linked protein after cleavage of the cross-linking reagent (Fig. 3). This cross-linking reagent has been used to identify a specific 56-kDa ZP3-binding protein on acrosome-intact mouse sperm (19). As much as 90% of the label initially on ZP3 could be transferred to the 56-kDa protein, and cross-linking was inhibited by excess unlabeled ZP3 protein. Moreover, ZP3 affinity columns retained a protein with the same molecular mass. This reagent has also been used to demonstrate that phospholamban interacts with a specific site on the ATPase from sarcoplasmic reticulum only when it is nonphosphorylated and the ATPase is in the Ca^{2+}-free state (106).

Another useful reagent of this type is ^{125}I-{S-[N-(3-iodo-4-azidosalicyl)cysteaminyl]-2-thiopyridine}, also called IAC, a cysteine-specific modifying reagent. This reagent was used to

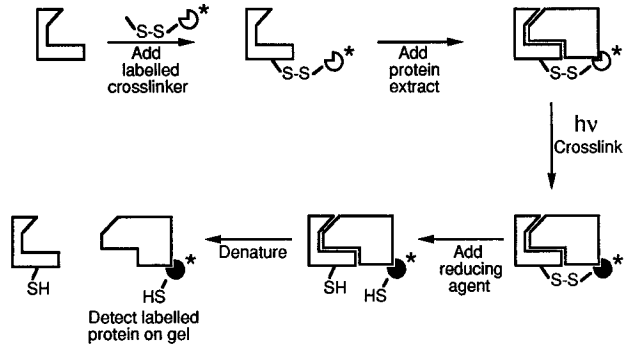

FIG. 3. Specific labeling of an interacting protein with a labeled photoactivatable cross-linking reagent.

demonstrate that the carboxy-terminal region of the subunit of *E. coli* RNA polymerase was adjacent to the activating domain of the catabolite activator protein (CAP) (33). To do this, a unique cysteine was introduced onto the surface of CAP, in a residue which tolerates a large number of mutations, and a preexisting surface cysteine was changed to serine. Subsequent reaction with labeled IAC resulted in quantitative incorporation of label and in protein with 70% of its transcription activation activity. Irradiation of the ternary complex of DNA, CAP, and RNA polymerase yielded 20% cross-linking, all of which was with a particular domain of the subunit of polymerase.

(c) Direct incorporation of photoreactive lysine derivative during translation. A photoactivatable group can be incorporated directly into the translation product by using a modified lysyl-tRNA. If translation is done in the presence of [^{35}S]methionine, the protein is simultaneously labeled and ready for photoactivated cross-linking. This approach has been particularly valuable in investigating the process by which proteins are inserted into the endoplasmic reticulum. During elongation, signal recognition particle (SRP) binds the nascent chain and halts translation until the arrested translation product is brought to the SRP receptor. This releases SRP, allowing translation to continue, coupled with translocation of the protein into the endoplasmic reticulum. With bovine preprolactin, there are two lysines at positions 4 and 9 of the signal sequence and no other lysine residues within the first 70 amino acids, after which translation is normally stopped by SRP. Thus, incorporation of lysine with a photoactivated group specifically probes interaction of the signal sequence with other interacting proteins. In this way, the nascent chain was specifically cross-linked with the 54-kDa protein of SRP and a 35-kDa microsomal membrane protein, called the signal sequence receptor (239). Subsequent experiments in the same system relied on translation of truncated mRNAs bearing lysine codons at different positions. These templates produce proteins that remain tethered to the ribosome through peptidyl-tRNA because of the lack of a termination codon. They therefore cannot complete translocation and are trapped, presumably as intermediates. In this way, it was shown that lysines in different positions also recognized the same 35-kDa membrane protein (121, 238). Moreover, this protein is probably required for translocation because antibodies against it inhibit translocation in vitro (87).

Investigation with the same system in *S. cerevisiae* demonstrated that prepro-α-factor is in contact with Sec61 protein (155). Antibody against either Sec61 or prepro-α-factor precipitated the same labeled cross-linked protein. Cross-linking

was observed only when prepro-α-factor was tethered; release of the protein with puromycin or a complete translation sequence abolished cross-linking. Moreover, the tethered prepro-α-factor was glycosylated while it was tethered, and cross-linking was ATP dependent for large tethered prepro-α-factor peptides; this indicated that prepro-α-factor had entered the normal glycosylation pathway. Sanders et al. (191) also demonstrated by conventional cross-linking followed by immunoprecipitation that Sec61 is in contact with tethered proteins being translocated (in this case by covalent coupling to avidin); the same experiments also demonstrated that BiP (Kar2) was cross-linked to the translocation intermediates and that *sec62* and *sec63* mutants modulate the process. The convergence of genetics and biochemical cross-linking studies further substantiates these interactions.

(d) Site-specific incorporation of photoreactive amino acid derivative during translation. Use of a suppressor tRNA to incorporate a photoactivatable amino acid derivative results in site-specific incorporation by use of a gene carrying a single stop codon. Two such reports have been described. High et al. (94) used a charged amber suppressor tRNA to insert a phenylalanine derivative into various regions of the signal sequence of preprolactin. Cross-linking experiments demonstrated that the amino-terminal end of the signal sequence is in proximity to the translocating chain-associating protein, whereas the hydrophobic core of the sequence contacts Sec61 protein. Cornish et al. (39) used a similar method to incorporate a different photoaffinity label. Still to be described is a similar method involving a labeled photoactivated amino acid replacement—the ultimate magic bullet.

(iii) Other considerations. One major disadvantage of using any cross-linking technique to detect protein-protein interactions is that it detects nearest neighbors which may not be in direct contact. The cross-linking reagent reaches out to any protein in close vicinity; thus, it may appear to detect protein interactions that are more like ships just passing in the night. This is more and more of a problem as the size of the cross-linking reagent increases. Any interaction detected by cross-linking should therefore be independently assessed for protein-protein interactions. However, cross-linking has three important advantages over other methods. First, it can "cement" weak interactions that would otherwise not be visible by other methods (see, for example, reference 48). Second, it can be used to detect transient contacts with different proteins at various stages in a dynamic process such as glycosylation, by freezing the process at different stages. Third, cross-linking can be done in vivo with membrane-permeable cross-linking reagents (48). It may also be possible to detect cross-linking in vivo after microinjection of a protein that is modified with a photoactivatable cross-linking group. To our knowledge, this has not yet been reported.

LIBRARY-BASED METHODS

A variety of methods have been developed to screen large libraries for genes or fragments of genes whose products may interact with a protein of interest. As these methods are by their nature highly qualitative, the interactions identified must be subsequently confirmed by biochemical approaches. However, the enormous advantage of these strategies is that the genes for these newly identified proteins or peptides are immediately available. This is in sharp contrast to the biochemical methods described in the section on physical methods to select and detect proteins that bind another protein, above, which generally result in the appearance of bands on a polyacrylamide gel. These library methods also differ from classical genetic techniques described in the section on genetic methods, below, which often require a specific phenotype before they can be carried out. Library screens are generally performed in bacteria or yeasts, organisms with rapid doubling times. Thus, these procedures can be completed rapidly.

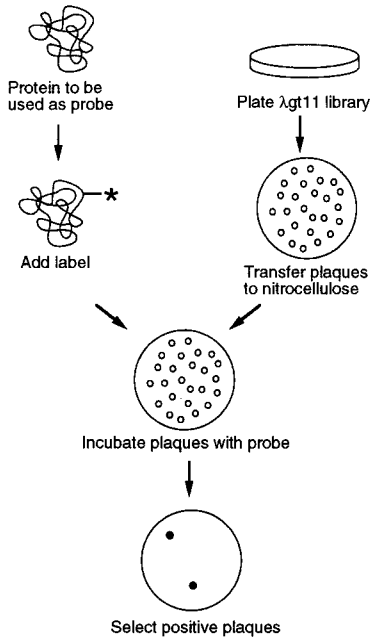

FIG. 4. Use of a labeled protein to probe an expression library.

Protein Probing

A labeled protein can be used as a probe to screen an expression library in order to identify genes encoding proteins that interact with this probe. Interactions occur on nitrocellulose filters between an immobilized protein (generally expressed in *E. coli* from a λgt11 cDNA library) and the labeled probe protein (Fig. 4). The method is highly general and therefore widely applicable, in that proteins as diverse as transcription factors and growth factor receptors have been used as probe. A variety of approaches can be used to label the protein ligand, or this ligand can be unlabeled and subsequently detected by specific antibody.

The method is based on the approach of Young and Davis (244), who showed that an antibody can be used to screen expression libraries to identify a gene encoding a protein antigen. The λgt11 libraries typically use an isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible promoter to express proteins fused to β-galactosidase. Proteins from the bacteriophage plaques are transferred to nitrocellulose filters, incubated with antibody, and washed to remove nonspecifically bound antibody. Protein ligands were first used as probes in this type of experiment by Sikela and Hahn (200), who identified a brain calmodulin-binding protein with ^{125}I-labeled calmodulin as the probe. The λgt11-expressed fusion protein bound calmodulin with a K_d between 3 and 10 nM, and binding was dependent on the presence of Ca^{2+}. The signal-to-noise ratio in these experiments was higher than that found with various antibody probes.

MacGregor et al. (135) used the leucine zipper and DNA-binding domain of Jun as a probe and identified the rat cAMP response element-binding protein type 1. In this case, the Jun

domain was biotinylated and detected with a streptavidin-alkaline phosphatase conjugate. Buffer conditions could be adjusted to distinguish a Jun-Jun homodimer from the more stable Fos-Jun heterodimer. Blackwood and Eisenman (17) used a similar approach with the basic-region helix-loop-helix leucine zipper domain (bHLH-zip) of the c-Myc protein. A 92-residue carboxy terminus of Myc, containing this domain, was expressed as a GST fusion protein, purified by glutathione-agarose affinity chromatography and ^{125}I labeled. This probe identified a new bHLH-zip protein termed Max, and gel shift experiments indicated that the Myc-Max complex exhibited site-specific DNA binding under conditions where neither Myc nor Max alone could bind. These results were critical in establishing a long-sought role for the Myc protein. Extending this result, Ayer et al. (6) used Max as a labeled probe to identify another member of this class, termed Mad.

A major advantage of the protein-probing approach is that the protein probe can be manipulated in vitro to provide, for example, a specific posttranslational modification or a metal cofactor. This modification or cofactor may be essential for the ability of the probe to bind to other proteins. This feature of the approach was exploited in the Ca^{2+}-dependent binding of calmodulin (200). Skolnik et al. (201) extended this use to phosphorylated probes in order to find proteins that bind to the carboxy-terminal phosphorylated tail of the epidermal growth factor (EGF) receptor. This tail is part of the intracellular domain of the receptor, which possesses a protein tyrosine kinase activity stimulated by binding of EGF. Skolnik et al. purified this domain from cells infected with a recombinant baculovirus, tyrosine phosphorylated it in vitro, and cleaved it to separate the phosphorylated carboxy-terminal tail from the protein kinase domain. Probing an expression library identified proteins containing the SH2 domain, which recognizes phosphotyrosyl-containing peptides. This cloning approach might be applied to the identification of proteins interacting with other activated phosphorylated receptors, including tyrosine- and serine-specific phosphatases as well as kinases. In addition, it should be possible to modify probe proteins by means other than phosphorylation to identify new proteins that recognize such modifications.

Probing expression libraries with labeled protein has numerous advantages. Since any protein or protein domain can be specifically labeled for use as a probe, the sophisticated arsenal of GST fusion vectors, other expression and tagging systems, and in vitro translation systems can be exploited; this makes preparation of the probe relatively straightforward. If specific antibody to the target protein is available, the probe protein need not be labeled; the antibody can be used in a second step to detect plaques that have bound the target protein. More than 10^6 plaques can be screened in an experiment, plating 5×10^4 plaques per 150-mm dish. The method not only results in the immediate availability of the cloned gene for the interacting protein but also can provide data regarding a specific domain involved in the interaction, because the λgt11 insert is often only a partial cDNA. Conditions of the wash cycles can be adjusted to vary the affinity required to yield a signal. As with many library-based methods, probing expression libraries compares equally all binary combinations of the probe protein and a library-encoded protein. Thus, less abundant proteins, proteins with weak binding constants, and proteins that temporally or spatially rarely interact with the probe protein in vivo can all be detected as long as their transcripts are present in the mRNA pool used to generate the library.

This method has certain intrinsic limitations. Proteins encoded by the library must be capable of folding correctly in *E. coli*, generally as fusion proteins, and of maintaining their structure on a nitrocellulose filter. However, proteins often can be renatured by subjecting the filters to a denaturation-renaturation cycle with 6 M guanidine hydrochloride as described by Vinson et al. (222). Binding conditions are arbitrarily imposed by the investigator, rather than reflecting the native environment of the cell. Since all combinations of protein-protein interactions are assayed, including those that might never occur in vivo, the possibility of identifying artifactual partners exists. In particular, the relative abundance of each potential partner expressed in a colony or plaque of the library is similar, instead of varying and potentially being compartmentalized as in the cell. Any posttranslational modifications necessary for efficient binding will generally not occur in bacteria (although some such modifications can be performed in vitro). Screening rather than direct selection is the means of detection, which inherently limits the number of plaques that can be assayed. The use of screening also restricts the further genetic manipulations that can be applied to the cDNA inserts. For example, in the analysis of point mutations, it is not possible to select directly for rare mutations that affect the interaction. Different protein probes are likely to behave variably in this approach, such that binding and washing conditions may have to be adjusted in each case to maximize the signal-to-noise ratio.

Phage Display

Basic approach. Smith (203) first demonstrated that an *E. coli* filamentous phage can express a fusion protein bearing a foreign peptide on its surface. These foreign amino acids were accessible to antibody, such that the "fusion phage" could be enriched over ordinary phage by immunoaffinity purification. Smith suggested that libraries of fusion phage might be constructed and screened to identify proteins that bind to a specific antibody. In the past few years, there have been numerous developments in this technology to make it applicable to a variety of protein-protein and protein-peptide interactions.

Filamentous phages such as M13, fd, and f1 have approximately five copies of the gene III coat protein on their surface; thus, a foreign DNA sequence inserted into this gene results in multiple copies of the fusion protein displayed by the phage. This is called polyvalent display. Similarly, the major coat protein encoded by gene VIII can also display a foreign insert (104). The gene VIII protein allows up to 2,700 copies of the insert per phage. Generally, polyvalent display is limited to small peptides (see the next section) because larger inserts interfere with the function of the coat proteins and the phage become poorly infective.

Random sequences can be inserted into gene III or gene VIII to generate a library of fusion phage (Fig. 5). Such a library can then be screened to identify specific phage that display any sequence for which there is a binding partner, such as an antibody. This screening is performed by a series of affinity purifications known as panning. The phage are bound to the antibody, which is immobilized on a plastic dish. Phage that do not bind are washed away, and bound phage are eluted and used to infect *E. coli*. Each cycle results in a 1,000-fold or greater enrichment of specific phage, such that after a few rounds, DNA sequencing of the tight-binding phage reveals only a small number of sequences. In addition to the advantage of high selectivity, a second advantage of this technology is that large phage libraries can be constructed (up to 10^9 to 10^{10} complexity) and the affinity purification step can be carried out at very high concentrations of phage ($>10^{13}$ phage per ml) (50). Third, the direct coupling of the fusion protein to its gene in a single phage allows the immediate availability of sequence

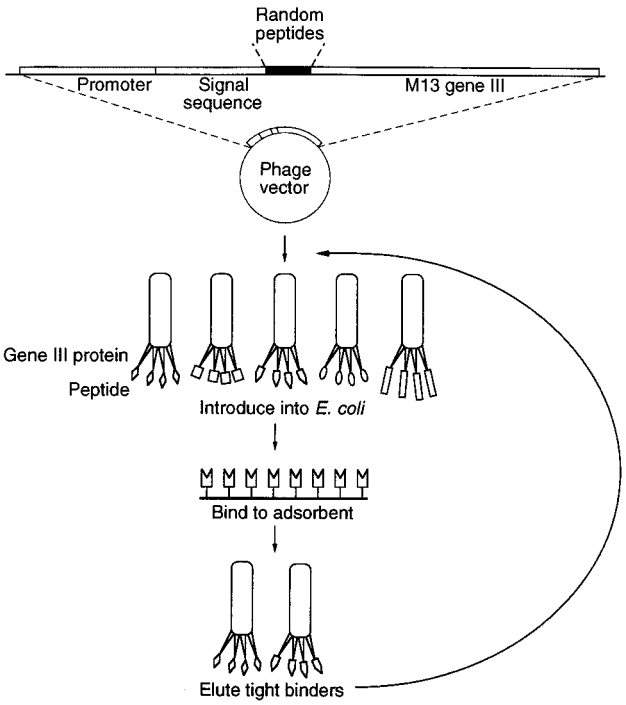

FIG. 5. A peptide library in a filamentous phage vector. The figure illustrates the process of panning, by which peptides that bind to an adsorbent are identified.

data to generate one or more consensus sequences of bound peptides or the sequences of variant proteins with a specific phenotype. Fourth, the phage can be used directly to assess the binding specificity of the encoded fusion proteins by varying the stringency of the wash procedures used in the panning cycles.

Random-sequence peptide libraries have been generated by cloning synthetic oligonucleotides into gene III (Fig. 5). Scott and Smith (198) generated a hexapeptide library and screened it to identify epitopes for two monoclonal antibodies specific for a hexapeptide from the protein myohemerythrin. Cwirla et al. (44) constructed a similar hexapeptide library to find peptides that can bind to a monoclonal antibody specific for a tetrapeptide from β-endorphin. Such epitope libraries allow rapid characterization of an unknown epitope recognized by either a monoclonal antibody or polyclonal serum. For example, monoclonal antibody pAB240, which recognizes the mutant conformation of the tumor suppressor p53 protein, was shown to bind to a 5-amino-acid motif in p53 (210). The binding partner for the phage-encoded peptides need not be an antibody. For example, Devlin et al. (50) constructed a 15-residue peptide library and used it to identify nine different peptides that bind to streptavidin.

A major advance in phage display came with the development of a monovalent system in which the coat protein fusion is expressed from a phagemid and a helper phage supplies a large excess of the wild-type coat protein (11, 131). Therefore, the phage are functional because the recombinant protein makes up only a small amount of the total coat protein. The vast majority (>99%) of the population of phage particles display either one or no copies of the fusion protein on their surface. Such phage can accommodate 50 kDa of foreign protein without any significant effect on phage infectivity. In addition, monovalent phage display avoids potential avidity effects observed with polyvalent display, in which the phage can attach to the adsorbent at multiple points.

Phage display has also been used to identify proteins with increased binding affinity. In some cases, the use of monovalent display was necessary to avoid potential avidity effects, attributed to multipoint attachment of the polyvalent phage to the absorbent (231). Lowman et al. (131) expressed nearly one million mutants of human growth hormone (191 residues) as fusion phage and identified variants that bound tightly to the growth hormone receptor. The mutations were directed to 12 sites known to be important for binding to the receptor. Some variants had binding affinities up to eightfold greater than that of the wild-type hormone. Roberts et al. (186) used polyvalent display of bovine pancreatic trypsin inhibitor and directed mutagenesis to five residues of the protein. They selected for high-affinity inhibitors of human neutrophil elastase and identified one variant with an affinity 3.6×10^6 higher than that of wild-type bovine pancreatic trypsin inhibitor.

A similar strategy can be used with nontargeted mutagenesis. For example, Pannekoek et al. (167) expressed human plasminogen activator inhibitor 1, a 42-kDa protein, as a gene III protein fusion under conditions for monovalent display. The phage-displayed inhibitor could specifically form complexes with serine protease tissue-type plasminogen activator. PCR mutagenesis was used to generate a library of mutant plasminogen activator inhibitor 1 proteins, which can be screened to analyze structure-function relationships.

Phage display presents several advantages for the study of protein-protein interactions. The very large sizes of either random libraries or pools of individual variants of a single sequence that can be generated mean that complex mixtures can be screened. While not strictly a genetic approach, in that there is no direct selection for an interacting partner, phage display has many of the properties of genetic selection through its use of panning cycles. It is a rapid procedure and should be widely applicable. Although screening a random library of cDNA by a panning procedure to identify proteins that interact with a protein of interest has not yet been demonstrated, this strategy should prove workable.

Disadvantages of phage display include the size limitation of protein sequence for polyvalent display; the requirement for proteins to be secreted from *E. coli*; and the use of a bacterial host which may preclude the correct folding or modification of some proteins. All phage-encoded proteins are fusion proteins, which may limit the activity or accessibility for binding of some proteins. Since binding is detected in vitro, the same considerations of an in vitro approach that are relevant for protein probing of expression libraries are relevant here.

Related methods. (i) Antibody phage. While we do not specifically address the vast topic of antigen-antibody interactions in this review, it is worth noting that phage display can be applied to these interactions. The principle of displaying antibody-combining domains on the surface of phage was first demonstrated by McCafferty et al. (141). The heavy- and light-chain variable domains of an anti-lysozyme antibody were linked on the same polypeptide and expressed as a gene III protein fusion. Over 1,000-fold enrichment of the antibody could be obtained by a single passage over a lysozyme-Sepharose column. This method was then extended by this and other groups to allow the display of libraries of combining domains, such that new antibodies or mutant versions of existing antibodies could be generated.

Kang et al. (110) used a vector to express a combinatorial library of functional Fab molecules (~50-kDa heterodimer) on the surface of a phage. The Fd chain, consisting of the variable region and constant domain 1 of the immunoglobulin heavy

chain, was synthesized as a gene VIII protein fusion, while the κ light chain contained no phage sequence. The two chains could assemble in the bacterial periplasm and become incorporated into the phage on coinfection with helper. Phage contained 1 to 24 antigen-binding sites per particle. The vector system described allows recombination of the two chains to generate large combinatorial libraries. A similar strategy to express Fabs by using the gene III protein has also been described (10). Additionally, a combinatorial library of linked heavy- and light-chain variable genes fused to the gene III protein has been shown to be capable of detecting a high-affinity binder (37). Kang et al. (110) suggested that such systems can be used for mutation and selection cycles to generate high-affinity antibodies. Moreover, they envisioned that the systems can be extended to analyze any protein recognition system, such as ligand-receptor interactions.

Phage display of Fab fragments was extended by Burton et al. (26), who generated a library of such fragments from the RNA of a human immunodeficiency virus-positive individual. After four rounds of panning with immobilized surface glycoprotein gp120 of the virus as the adsorbent, specific viral antibodies were obtained. A similar method was used to obtain human antibody Fabs that recognize the hepatitis B surface antigen (246).

(ii) Peptides on plasmids. In a method highly analogous to phage display, random peptides are fused to the C terminus of the *E. coli* Lac repressor and expressed from a plasmid that also contains Lac repressor-binding sites (43). Thus, the peptide fusions bind to the same plasmid that encodes them. The bacterial cells are lysed, and the peptide libraries are screened for peptides that bind to an immobilized receptor by using similar panning cycles to those for phage libraries. In this case, peptides become enriched because bound peptides carry their encoding plasmids with them, via the repressor-operator interaction, and these plasmids can be transformed back into *E. coli*. In the initial example, peptides that bind to a monoclonal antibody specific for dynorphin B were selected, and these peptides contained a hexapeptide sequence similar to a segment of dynorphin B (43). This method is distinguished from the phage display methods in that the peptides are exposed at the C terminus of the fusion protein and the fusions are cytoplasmic rather than exported to the periplasm.

Two-Hybrid System

The two-hybrid system (35, 65, 66) is a genetic method that uses transcriptional activity as a measure of protein-protein interaction. It relies on the modular nature of many site-specific transcriptional activators, which consist of a DNA-binding domain and a transcriptional activation domain (23, 97, 112). The DNA-binding domain serves to target the activator to the specific genes that will be expressed, and the activation domain contacts other proteins of the transcriptional machinery to enable transcription to occur. The two-hybrid system is based on the observation that the two domains of the activator need not be covalently linked and can be brought together by the interaction of any two proteins. The application of this system requires that two hybrids be constructed: a DNA-binding domain fused to some protein, X, and a transcription activation domain fused to some protein, Y. These two hybrids are expressed in a cell containing one or more reporter genes. If the X and Y proteins interact, they create a functional activator by bringing the activation domain into close proximity with the DNA-binding domain; this can be detected by expression of the reporter genes (Fig. 6). While the assay has been generally performed in yeast cells, it works similarly in mammalian cells

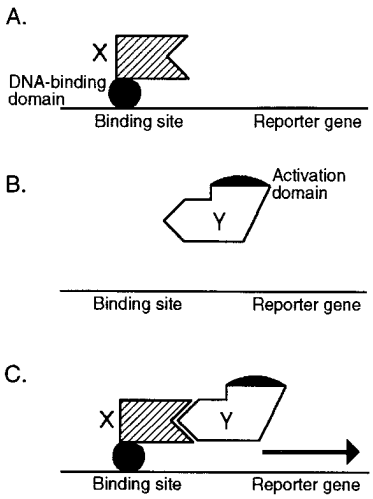

FIG. 6. The two-hybrid system. (A) The DNA-binding domain hybrid does not activate transcription if protein X does not contain an activation domain. (B) The activation domain hybrid does not activate transcription because it does not localize to the DNA-binding site. (C) Interaction between X and Y brings the activation domain into close proximity to the DNA-binding site and results in transcription.

(see, e.g., reference 46) and should be applicable to any other eukaryotic cells.

This method has been used with a wide variety of proteins, including some that normally reside in the nucleus, cytoplasm, or mitochondria, are peripherally associated with membranes, or are extracellular (see reference 66 for a review). It can be used to detect interactions between candidate proteins whose genes are available by constructing the appropriate hybrids and testing for reporter gene activity (220, 249). If an interaction is detected, deletions can be made in the DNA encoding one of the interacting proteins to identify a minimal domain for interaction (35). In addition, point mutations can be assayed to identify specific amino acid residues critical for the interaction (127). Most significantly, the two-hybrid system can be used to screen libraries of activation domain hybrids to identify proteins that bind to a protein of interest. These screens result in the immediate availability of the cloned gene for any new protein identified. In addition, since multiple clones that encode overlapping regions of protein are often identified, the minimal domain for interaction may be readily apparent from the initial screen (105, 223).

A variety of versions of the two-hybrid system exist, commonly involving DNA-binding domains that derive from the yeast Gal4 protein (35, 55) or the *E. coli* LexA protein (223, 247). Transcriptional activation domains are commonly derived from the Gal4 protein (35, 55) or the herpes simplex virus VP16 protein (45). Reporter genes include the *E. coli lacZ* gene (65) and selectable yeast genes such as *HIS3* (55) and *LEU2* (247). An increasing number of activation domain libraries are becoming available, such that screens are now feasible for proteins from many different organisms or specific mammalian tissues.

One field in which the two-hybrid system has been applied with considerable success has been the study of oncogenes and tumor suppressors and the related area of cell cycle control. For example, reconstruction experiments with previously cloned proteins indicated that interactions occur between Ras and the protein kinase Raf (220, 249), human Sos1 guanine nucleotide exchanger and the growth factor receptor-associ-

ated protein Grb2 (30), and Raf and the transcription factor inhibitor IκB (129). Two-hybrid searches with oncoproteins or tumor suppressors as targets have identified a leucine zipper protein that binds to Jun (34); protein phosphatase PP1α2, which binds to Rb (55); a bHLH-zip protein Mxi1, which binds to the Myc-associated protein Max (247); and the Rb-related protein p130, which binds to cyclins and was identified through its interaction with the cyclin-dependent kinase Cdk2 (83). A notable convergence of different approaches came about with the identification of another protein that binds to Cdk2, a 21-kDa protein termed Cip1, which inhibits the kinase activity (85). This protein turned out to be identical to a protein encoded by the major p53-inducible transcript (58), suggesting that the tumor suppressor role of p53 may be mediated by its activation of the gene for this 21-kDa protein.

The two-hybrid system has several features that make it useful for analysis of protein-protein interactions. It is highly sensitive, detecting interactions that are not detected by other methods (see, e.g., references 127 and 220). On the basis of binding of different proteins to the retinoblastoma protein, Durfee et al. (56) estimate that the minimal binding constant required to detect an interaction in their version of the two-hybrid system is on the order of 1 μM. This value suggests that the system should be applicable to a wide range of protein interactions. However, it is clear that the minimal affinity interaction detectable will depend on such variables as the level of expression of the hybrid proteins; the number, sequence, and arrangement of the DNA-binding sites in the reporter gene(s); and the amount of reporter protein required for a detectable phenotype. Given these variables, it is likely that some versions of the system may detect weak interactions with binding constants considerably greater than 1 μM. Another advantage is that the interactions are detected within the native environment of the cell and hence that no biochemical purification is required. The use of genetic-based organisms like yeast cells as the hosts for studying interactions allows both a direct selection for interacting proteins and the screening of a large number of variants to detect those that might interact either more or less strongly. With a reporter gene such as the yeast *HIS3* gene, the competitive inhibitor 3-aminotriazole can be used to directly select for constructs which yield increased affinity.

The two-hybrid system is limited to proteins that can be localized to the nucleus, which may prevent its use with certain extracellular proteins. Proteins must be able to fold and exist stably in yeast cells and to retain activity as fusion proteins. The use of protein fusions also means that the site of interaction may be occluded by one of the transcription factor domains. Interactions dependent on a posttranslational modification that does not occur in yeast cells will not be detected. Many proteins, including those not normally involved in transcription, will activate transcription when fused to a DNA-binding domain (134), and this activation prevents a library screen from being performed. However, it is often possible to delete a small region of a protein that activates transcription and hence to remove the activation function while retaining other properties of the protein.

Other Library-Based Methods

A number of other library strategies have been developed recently that, in principle, should result in the identification of proteins that interact with a protein of interest. However, because the first description of methods generally involves known combinations of proteins, the general applicability of a new method cannot be easily judged.

In one approach, the ability of the *E. coli* bacteriophage λ repressor to dimerize was used as a reporter for the interaction of leucine zipper domains (98). The N-terminal domain of repressor binds to DNA but dimerizes inefficiently; a separate C-terminal domain that mediates dimerization is required for efficient binding of the protein to its operator. The N-terminal DNA-binding domain was fused to the leucine zipper of the yeast Gcn4 protein, which allowed dimerization and repression of transcription in *E. coli*. This repression enabled the host cell to survive superinfection by λ phage. This phenomenon enabled Hu et al. (98) to introduce single-amino-acid mutations into the leucine zipper domain and to use a genetic assay in *E. coli* to determine whether dimerization of the zipper domain occurred. They suggested that this assay could be used to select clones from a library for proteins that bind to a target protein, which is expressed in *E. coli* as a repressor hybrid. Any phage that express a protein that binds to the target protein should compete for dimerization of the repressor and its ability to bind λ operators. These phage would be detected because they result in plaques. As described, this approach would be limited to target proteins that homodimerize. In addition, this method when applied to library screening is a competition assay; it would require that the library-encoded protein bind to the target protein in preference to the target protein interacting with itself.

Another *E. coli*-based assay involves tagging the target protein with biotin by fusing it to the biotin carboxylase carrier protein (74). This tag allows the protein to be bound by avidin, streptavidin, or anti-biotin antibody-coated filters. Potential interacting proteins are fused to the LacZ protein and expressed from a λ vector such that β-galactosidase activity is intact. These phage are infected into cells containing the biotin-tagged target protein, and interaction can occur in vivo between a library-encoded protein and the target protein. This interaction is then detected when the phage plaques are transferred to avidin filters and assayed for β-galactosidase activity. The method was shown to work by using biotinylated c-Jun protein and a c-Fos–LacZ fusion. Although the protein-protein interaction occurs within the living bacterial cells, the detection of this interaction occurs in vitro on filters that must be washed after transfer of the proteins. Thus, in principle, this method may have many of the same limitations that protein probing of expression libraries has.

GENETIC METHODS

For organisms for which powerful genetic analysis methods exist, sophisticated strategies can be designed to uncover genes that show interactions with other genes. In many cases, these newly uncovered genes encode proteins that physically interact with proteins encoded by the known genes. In other cases, genetic methods can be used to confirm interactions among previously identified proteins. These strategies are generally based on classical genetic approaches. For example, identification of extragenic suppressors often reveals mutations in genes whose products physically interact with the protein containing the original defect. Synthetic lethal screens yield mutations that, in combination with another nonlethal mutation, result in the inability of the organism to grow; this phenotype is commonly due to alterations in interacting proteins. Overproduction of certain proteins can lead to the suppression of mutations in interacting proteins. In other cases, overproduction disrupts a cellular process by altering the balance of the different components of a complex structure, or the overproduced protein is nonfunctional and acts in a dominant-negative manner.

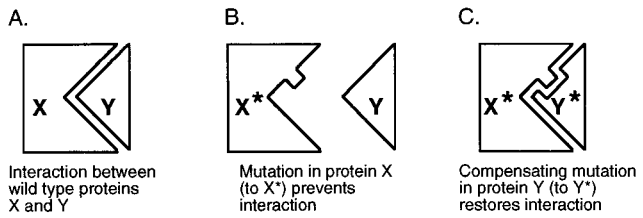

FIG. 7. Extragenic suppression due to restoration of a protein-protein interaction.

The value of some of these genetic approaches has been significantly increased by applying them to organisms not amenable to classical genetic techniques, using modern molecular tools. For example, the ability to generate mice either carrying novel genetic information or deleted for one or more of their endogenous genes allows this organism to be analyzed by some of the logic formerly reserved for much simpler creatures. However, it must be kept in mind with any genetic approach that identification of mutants with the correct phenotypes does not guarantee that the biochemical mechanisms invoked to explain these phenotypes are correct.

Extragenic Suppressors

Suppressor mutations are mutations that partially or fully revert the phenotype caused by an original mutation (see reference 86 for review). Extragenic suppressors occur in genes other than the gene carrying the primary mutation. This is illustrated in Fig. 7, in which a mutation of protein Y to Y* compensates for the defect X* to restore activity to the XY dimer. However, analysis of these suppressors is often difficult, because they lack any phenotype in the absence of the primary mutation. To circumvent this problem, Jarvik and Botstein (107) sought suppressors of temperature-sensitive mutants of phage P22 that resulted in a cold-sensitive phenotype. This cold-sensitive phenotype did not necessarily depend upon the presence of the original mutation causing temperature sensitivity, and thus mutations in new genes could be uncovered. It was proposed (107) that one mechanism of this suppression is that the original mutation and the suppressor lie in genes whose products physically interact and that the original mutation destroyed this interaction. The suppressor then produces a compensating alteration that restores the interaction.

This type of suppressor analysis has been exploited in studying fundamental processes in yeast cells, particularly cell cycle control, cytoskeleton structure, and RNA splicing. Moir et al. (152) isolated cold-sensitive cell division cycle (*cdc*) mutants of *Saccharomyces cerevisiae* and used them to identify temperature-sensitive revertants. Some of these revertants carried new mutations that alone resulted in a *cdc* phenotype at the restrictive temperature, suggesting that the mutated gene products might interact with the cold-sensitive protein. These results support the idea that only a few genes might be capable of mutation to generate an altered product that can suppress the original mutation. Thus, this approach can be applied to a process such as cell cycle control and reveal most or all of the interacting gene products.

In a similar strategy, suppressors of a temperature-sensitive mutation in the *S. cerevisiae* actin gene that acquired a cold-sensitive phenotype identified five new genes (160). Mutations in these genes, even in a background with the wild-type actin gene, led to phenotypes similar to those of actin mutants. These results suggested that these genes could encode proteins that are part of the actin cytoskeleton. In a related approach, dominant suppressors of an actin mutation also identified a gene whose product may interact with actin (3). In both these cases, the suppressor mutations showed allele specificity; some but not all actin alleles were suppressed by a given mutation. This allele specificity also supported the idea of a direct physical interaction, in that suppressor mutations that simply bypass the requirement for the protein containing the original mutation would not be expected to show such specificity.

The nematode *Caenorhabditis elegans* has also been used extensively for suppression analysis because large populations of individuals can be examined (96). If a temperature-sensitive mutant is available, it can be shifted to the restrictive temperature to apply a direct selection for suppressors. This approach has been used to study such processes as movement, egg laying, and sex determination. One example is the suppression of an *unc-22* mutation that resulted in muscle twitching (151). Some of these suppressors were mutations in the *unc-54* gene which encodes the major myosin gene. These results suggested that the *unc-22* and *unc-54* proteins physically interact, and this idea is supported by the finding that the *unc-22* protein, like myosin, is located in the A-bands of muscle (150).

Suppressor analysis can clearly uncover new mutations that affect a process under study, and analysis of the genes and proteins defined by these mutations sometimes indicates interacting proteins. While often used with temperature-sensitive and cold-sensitive mutations, many other types of spontaneous mutations can also be readily suppressed if an appropriate genetic selection is available. With the availability of numerous cloned genes, conditional alleles can now be generated by in vitro mutagenesis methods. An obvious limitation of this type of analysis is that it can generally be applied only to simple organisms such as phages, bacteria, yeasts, nematodes, and *Drosophila* species. It requires not only the gene of interest but also a useful mutant to initiate the analysis. For example, suppressors in an interacting protein may be difficult or impossible to obtain if the original mutation does not affect a domain of interaction. Furthermore, other mechanisms can yield suppressors. These include second intragenic mutations, gene duplication of the original mutant gene, suppression by epistasis, and informational suppression (see, for example, reference 96). Thus, identification of the suppressors of interest against a background of these other mutations can be a time-consuming process.

Synthetic Lethal Effects

Mutations in two genes can cause death (or another observable defect) while mutation in either alone does not. This phenomenon is called a synthetic effect and can result from physical interactions between two proteins required for the same essential function. This is illustrated in Fig. 8, in which the dimer XY is required for some function and loss of this function results in a detectable phenotype. Mutation in X or Y yields partial binding, but the double mutant X*Y* has no binding. Dobzhansky (52) first described synthetic lethal effects in *Drosophila* species. However, the search for synthetic lethal effects has been applied successfully most often in *S. cerevisiae*. One of the tools available for research in this organism is a colony-sectoring assay (93, 119), in which cells containing a plasmid are red and can therefore be easily distinguished from those that have lost the plasmid and are white. If maintenance of the plasmid is not essential for viability of the yeast, colonies appear with red-and-white sectoring. If the cells become dependent on a gene carried by the plasmid, the colonies appear uniformly red. For example, Bender and Pringle (15) used such an assay with a plasmid-borne copy of the *MSB1* gene, which

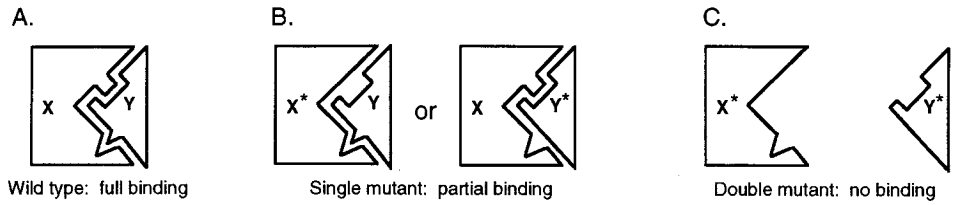

FIG. 8. Synthetic effect, in which either single mutant is functional but not the double mutant.

plays a role in bud formation. They mutagenized the plasmid-containing cells and screened for mutants in which *MSB1* had become essential for survival. This screen identified two new genes, *BEM1* and *BEM2*, in which mutations led to defects in cell polarity and bud emergence. In this approach, if the plasmid is maintained at high copy number in *S. cerevisiae*, it is also possible to identify mutations in new genes that are lethal but can be suppressed by multiple copies of the plasmid-borne gene.

A similar approach was taken by Costigan et al. (40) to identify mutants that require the Spa2 protein, which is also involved in polarized cell growth as well as in the morphogenetic changes that occur in yeast mating. The synthetic lethal screen identified the *SLK1* gene, which is necessary for morphogenesis in vegetatively growing yeast cells and in mating pheromone-treated cells. Costigan et al. pointed out that the synthetic lethal screen by the colony color assay is extremely sensitive and can identify mutants with low viability. Since both *spa2* and *slk1* mutants are individually healthy, the screen did not simply combine two mutations each causing unhealthiness to result in death, a common concern in using this method. Instead, it seems likely that the synthetic lethal effect often results from two different defects in the same cellular process.

Other synthetic lethal screens in yeast cells involve a poison assay in which the presence of a plasmid-borne gene on a particular medium is lethal; when yeast cells containing this plasmid are placed on such a medium, there is strong selection for cells that have lost the plasmid. However, mutants that cannot survive without the plasmid can be identified, because the plasmid also contains the gene of interest whose presence is required in these mutants. Such mutants do not grow on replica plates containing the poison. This approach was used to identify mutations in the 3-hydroxy-3-methylglutaryl coenzyme A reductase genes (12). Alternatively, the gene of interest can be expressed by using a regulated promoter, such that mutants that do not survive the repressed condition are identified. Inducible expression of the yeast *RAS2* gene led to the identification of mutations in the *CYR1* gene, which encodes adenylate cyclase (149). Finally, synthetic lethal effects can be uncovered by combining mutations identified in other genetic screens. For example, yeast cells containing a temperature-sensitive mutation in the *SEC4* gene, essential for secretion, are inviable at the permissive temperature when they also contain a temperature-sensitive mutation in certain other *SEC* genes (190). Yeast cells with mutations in both α-tubulin and β-tubulin are inviable (101).

While synthetic lethal screens often lead to the identification of interacting gene products, other explanations do not require this physical interaction (101). For example, the two proteins might both be components of the same structure, or one protein could regulate the activity of the other. Additionally, there are likely to be some cases in which the combination of two mutations, either of which causes poor growth on its own, leads to complete inviability.

Overproduction Phenotypes

Overproduction of wild-type proteins. The overproduction of some wild-type proteins can lead to phenotypes that provide insight into protein-protein interactions. In *S. cerevisiae*, a multicopy plasmid often suppresses mutations in genes other than the one carried on the plasmid (reviewed in reference 182). For example, a temperature-sensitive mutation in the *CDC28* gene, which encodes a protein kinase involved in controlling cell division, can be suppressed by multicopy plasmids carrying the *CLN1* or *CLN2* gene, which encode cyclins (82).

In other cases, overproduction of a protein can cause a phenotype that is altered by overproduction of an interacting protein. High-copy-number plasmids expressing either of the yeast histone pairs H2A and H2B or H3 and H4 caused an increased frequency of chromosome loss (142). However, overproduction of both pairs of histone proteins did not affect the fidelity of chromosome transmission, indicating that it is the imbalance of the two dimer sets with respect to one another that affects this fidelity (142). Overproduction of the yeast Gal4 protein, the transcriptional activator of the galactose-inducible genes, leads to galactose-independent transcription. However, proper regulation is restored if the Gal80 protein, a negative regulator that binds to the Gal4 protein, is also overproduced (159). While the phenotype due to an overproduced wild-type protein may reflect interactions with another protein (either mutant or wild type), there are several other mechanisms by which such phenotypes can occur. For example, an overproduced protein may bypass the transcriptional regulation due to another protein. In other cases, an overproduced protein may lead indirectly to the stabilization of a mutant protein.

Overproduction of mutant proteins. Overproduction of a nonfunctional version of a protein can result in a mutant phenotype due to disruption of the activity of the wild-type protein (Fig. 9) (reviewed in reference 90). The existence of such dominant-negative proteins can lead to a definition of the oligomerization domain of a protein. An early example of this came from studies of the *E. coli* Lac repressor, which has distinct domains for DNA binding and for oligomerization. A mixed oligomer of wild-type subunits and mutant subunits unable to bind DNA results in a nonfunctional repressor (143). This kind of mutant provides evidence for the multimeric nature of the repressor, and analysis of the sites of mutation defined the domains involved in DNA binding and in oligomerization.

A similar mechanism may operate in many human cancers. The wild-type p53 protein is a transcriptional regulator which is tetrameric, and its oligomerization domain is near the C terminus. Mutations in the central domain of p53 that occur in tumors produce dominant-negative mutant proteins that bind to and inactivate the function of the wild-type protein (67). The ability to manipulate cloned genes and reintroduce these mutant versions into cells now allows dominant-negative mutants to be created in many different organisms. For example,

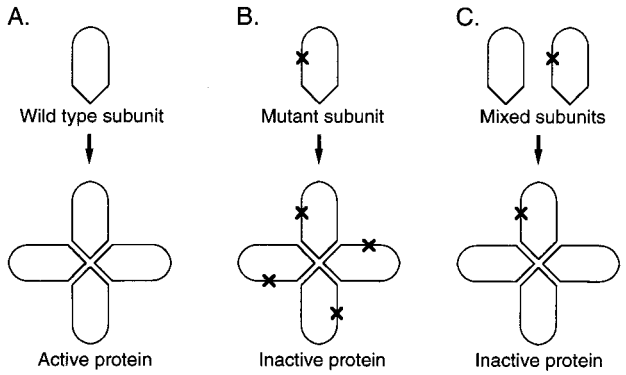

FIG. 9. Dominant-negative effect. Pure populations of wild-type (A) or mutant (B) subunits result in an active or inactive protein, respectively. A mixture of the two types (C) will also be inactive if the mutant subunit acts in a dominant fashion.

dominant-negative Myc proteins were overexpressed in fibroblasts and shown to inhibit transformation by the v-*abl* and *BCR-ABL* oncogenes (194). It was suggested that this effect was due to the mutant Myc proteins competing with the endogenous wild-type Myc protein for binding to the Max protein, thus forming nonfunctional heterodimers.

Unlinked Noncomplementation

Individuals heterozygous for two different recessive mutations sometimes display a mutant phenotype. This unlinked noncomplementation is often interpreted as being due to mutation in two genes that encode interacting products. In *Drosophila* spp., new recessive mutations were identified that failed to complement β_2-tubulin mutations and that mapped to other genes (176). At least one of these mutations mapped very close to an α-tubulin gene. A model for this noncomplementation is based on a minimal dosage requirement for the product of two interacting proteins. If the mutant proteins assemble randomly with the wild type, the double heterozygote would contain only one-fourth the normal level of complex, which would be insufficient for function. In addition, when homozygous, some of the second-site noncomplementing mutations lead to defects in tubulin function, and this property is consistent with the model.

POPULAR METHODS TO ESTIMATE AND DETERMINE BINDING CONSTANTS

Importance of Characterization of the Binding Interaction

The ultimate goal of studying protein-protein interactions is to understand the consequences of the interaction for cell function. This depends in turn on understanding the strength of the interaction in the cell. The determination that two proteins can interact with one another is only the first step in understanding if, and to what extent, the interaction takes place in vivo. Evaluation of the interaction requires the assessment of at least six parameters, which are discussed below.

Binding constant. For any simple interaction of one protein (P) with another (L, for ligand), the interaction is governed by the binding constant K_d, according to the simple equation $K_d = [P_f][L_f]/[PL]$. In this equation, $[P_f]$ and $[L_f]$ refer to the free (i.e., unbound) concentrations of P and L respectively. The interaction between protein and ligand is also expressed in two other ways. First, it is often expressed instead as an affinity constant, $K_a = [PL]/[P_f][L_f]$, i.e., $K_a = 1/K_d$. Second, it is often expressed as a ratio of two rate constants. The rate of formation of PL is $k_a\ [P_f][L_f]$, where k_a is the association rate constant, and the rate of breakdown of PL is $k_d\ [PL]$, where k_d is the dissociation rate constant. At equilibrium, the rate of formation of PL equals the rate of breakdown of PL, and $K_d = k_d/k_a$. Evaluation of the dissociation constant is the subject of this section.

Concentrations of species. To evaluate the extent to which two proteins can interact, the cellular (or compartmental) concentrations of P_t (the sum of bound and unbound concentrations) and L_t are required, in addition to the dissociation constant. These two parameters can drastically alter an evaluation of the population of molecules in a complex. For example, if $K_d = [P_t] = [L_t]$, 38% of the species are in the complex PL at any one time. If K_d is 10-fold higher (weaker binding), only 8.4% of the species are in the complex at one time, and if K_d is 10-fold lower (stronger binding), 73% of the species are in a complex. A similar effect holds for alterations in the concentrations of P and L in the cell. A simple way of calculating [PL] from the easily measured parameters $[P_t]$ and $[L_t]$ is as follows: $[PL] = \{([P_t] + [L_t] + K_d)/2\} - 1/2\ \{([P_t] + [L_t] + K_d)^2 - 4\ [L_t][P_t]\}^{1/2}$ (54).

Influence of competing proteins. Even if a protein has high affinity for a ligand protein, L, and the protein and ligand are present in sufficient quantities to interact functionally in the cell, they may not do so in vivo to the same extent as in vitro. Other ligands may effectively compete for the ligand protein if they are present at high enough concentration and interact with sufficient affinity. For example, if the concentration of P and L1 are both equal to the dissociation constant, 38% of the species are in a complex. If another ligand, L2 (or a set of potential ligands), is present at 1,000 times the concentration of L1 and has 10-fold-lower affinity for P, the interaction of P with L2 will titrate the vast majority of the protein P (99%, if L2 was the only interacting protein), leaving very little to interact with L1. This sort of consideration is addressed in part by protein affinity columns, coimmunoprecipitation experiments, and cross-linking, since all the proteins in the applied extract have equal opportunity to bind. It is not addressed in affinity blotting or library-based detection methods, in which gene products are tested individually.

Influence of cofactors. Two types of cofactors can influence protein-protein interactions. First, small effector molecules and ions such as ATP, GTP, and Ca^{2+} can influence many protein-protein interactions. Second, other macromolecules (DNA, RNA, and proteins) can affect protein-protein interactions by forming ternary (or larger) complexes. Such complexes can be very much more stable than the corresponding binary complexes.

Effect of cellular compartmentation. A protein that is interacting with a ligand or a set of ligands is also influenced by its location in the cell. For example, some transcription factors are regulated in part by their partitioning between the cytoplasm and nucleus; they can interact with the transcription machinery only when they are in the nucleus.

Solution conditions. Other factors that can affect the strength of protein-protein interactions include solution conditions (salt concentration, pH, etc.), as well as the effects of molecules such as polyethylene glycol, which causes macromolecular crowding and can significantly lower the observed binding constant of proteins (see, for example, reference 108).

Limits of Binding-Constant Considerations

The lower limit for the concentration of a protein in an organism of the size of the yeast *S. cerevisiae* is 0.1 nM (as-

TABLE 1. Dissociation constants for some well-defined protein-protein interactions

Complex[a]	K_d (M)	Method[b]	Reference(s)
PDEαβ:PDEγ	1.3×10^{-10}	Activity	103
	5×10^{-11}	fl. an.	25
	1×10^{-11}	fl. an.	234
	$<1 \times 10^{-11}$	Activity	233
TαGTPγS:PDEγ	$<1 \times 10^{-10}$	int. fl.	164
TαGDP:PDEγ	3×10^{-9}	int. fl.	164
CAP cAMP:RNA polh	3×10^{-5}	fl. an.	91
	1×10^{-6}	fl. an.	170
T7 gene 2.5 protein:T7 DNA polymerase	1.1×10^{-6}	fl. an.	115
λ repressor (dimer to tetramer)	2.3×10^{-6}	fl. an.	9
λ repressor (monomer:dimer)	2×10^{-8}	l.z. gf.	13, 193
Citrate synthase: malate dehydrogenase	1×10^{-6}	fl. an.	214
C4 binding protein: human protein S	6×10^{-10}	Solid phase	158
p85 (PI3K): tyrosine-phosphorylated peptide from PDGF	5.2×10^{-8}	SPR	166
CheY:CheA	3×10^{-8}	SPR	197
CheA:CheW	1.3×10^{-5}	eq. gf.	72
VAMP2:syntaxin A	4.7×10^{-6}	SPR	27
EGF:EGF receptor	4.1×10^{-7}	SPR	249
PKA-C:PKA-R	2.3×10^{-10}	SPR	88
PRI:angiogenin	7×10^{-16}	Fluorescence, exch	126
ras:raf	5×10^{-8}	GST ppt'n	227
NusB:S10	1×10^{-7}	Sucrose gradient sed'n	138
NusA: core RNA polymerase	1×10^{-7}	Sucrose gradient sed'n	80
		Fluorescence tag	76
Trypsin:pancreatic trypsin inhibitor	6×10^{-14}	Kinetics, comp'n	221

[a] Abbreviations: PDE, phosphodiesterase; TαGTPγS, α subunit of transducin complexed with GTPγS; TαGDP, α subunit of transducin complexed with GDP; CAP cAMP, catabolite gene activator protein complexed with cAMP; RNA polh, RNA polymerase holoenzyme; PDGF, platelet-derived growth factor; VAMP2, vesicle-associated membrane protein 2; PKA-C, catalytic subunit of protein kinase A; PKA-R, regulatory subunit of protein kinase A.

[b] Abbreviations: fl. an., fluorescence anisotropy; int. fl., intrinsic fluorescence; l.z. gf, large zone equilibrium gel filtration; eq. gf., equilibrium gel filtration; SPR, surface plasmon resonance; exch, exchange; ppt'n, precipitation; sed'n, sedimentation; comp'n, competition.

suming a radius of 1.5 μm and one molecule per cell), and for an animal cell with a radius of 10 μm, the lower limit is about 0.3 pM. Thus, for two such proteins to interact a significant percentage of the time, the dissociation constant must be at the same concentration (in which case they will interact 38% of the time). At the other extreme, some glycolytic proteins represent 1% or more of the soluble protein in the cell. Indeed, glyceraldehyde-3-phosphate dehydrogenase is reported to approach 20% of the soluble protein in *S. cerevisiae* under certain conditions. This upper limit corresponds to 1.7×10^7 protein molecules per cell and a cellular concentration of 1 mM, and it represents the upper limit for binding-constant considerations of two such proteins. In considering protein concentrations, it is worth noting that a typical yeast cell contains about 3×10^5 ribosomes (226), 100 to 500 molecules of tRNA splicing enzymes (169, 178), and 300,000 molecules of actin (157).

Methods for Determining Binding Constants

A number of methods have been described to measure binding constants. Some of the more commonly used ones are described below, together with a brief evaluation of the method. The values of dissociation constants for several protein-protein interactions are listed in Table 1.

Binding to immobilized proteins. Protein affinity chromatography can be used to estimate the binding constant. This method is well described in an excellent review (69). The form of the binding equation that is used in this sort of experiment expresses the fraction of L bound to protein P as follows: $[PL]/[L_t] = [P_f]/([P_f] + K_d)$. As long as the concentration of covalently bound protein $[P_t]$ is in great excess over that of the ligand, $[P_t] \approx [P_f]$ and the fraction of protein L that is bound is $[P_t]/(K_d + [P_t])$. Thus, if $[P_t] = 100\ K_d$, essentially all of L is bound (a little more than 99%), and if $[P_t] = 0.01\ K_d$, very little of L is bound (a little less than 1%).

Columns are prepared with different concentrations of covalently bound protein. Then a preparation of the interacting protein ligand is loaded on the column and washed with 10 column volumes of buffer, and bound protein is eluted with SDS. At a concentration of 20 K_d, the covalently bound protein retains 95% of the ligand in one column volume and therefore 0.95^{10} or 61% in 10 column volumes. Thus, the lowest concentration of bound protein that allows retention of most of the ligand is 20 K_d.

The percentage of bound ligand drops very quickly as the concentration of covalently bound P on the column is lowered, particularly as the concentration of P_t approaches K_d. At 5 K_d 16% of the ligand would be retained, at 2 K_d 1.7% of the protein would be retained, and at 1 K_d only 0.1% would be retained. It is for this reason that detection of interacting proteins by affinity chromatography depends critically on the concentration rather than the amount of bound protein (see the section on protein affinity chromatography, above).

An important parameter in this experiment is the amount of protein that is active on the column. Estimates range from 10% for gene 32 protein to about 50% for others (69). A second factor is the amount of pure protein available to be coupled. If protein is limiting, sufficiently high concentrations of bound protein on the gel are achieved only with appropriate microcolumns. Such columns, with as little as 20 μl of appropriate beads, are described in detail by Formosa et al. (69). With the recent widespread use of gene fusion technology, large quantities of protein are not a serious problem with most cloned structural genes. A third factor, which is evident from the discussion above, is the form of the protein that is used for the determination. Proteins that require modification to be active must be purified in that form for proper evaluation.

This method works well in estimating the binding constant. However, it is not clear that the values obtained represent a true equilibrium constant; if so, one would have to assume that the bound ligand is always in equilibrium with the solution ligand during flow of the column and that interactions of solid-phase bound protein with liquid-phase ligand are the same as interactions in the liquid state. Nonetheless, for interactions that have been measured by more than one method, the results agree well (see reference 69 and references therein).

Sedimentation through gradients. The method of sedimentation through gradients measures populations of complexes by monitoring the rate of sedimentation of a mixture of proteins through gradients of glycerol or sucrose. Fractions are assayed by appropriate methods (activity, immunoblotting, etc.) to determine the elution positions of each protein. Proteins will sediment as a complex at concentrations above the binding constant (provided that the complex is stable; see the discussion below) and at their native positions at concentrations below the binding constant. By varying the concentration of one or both of the proteins and taking into account the dilution of the species during sedimentation, one can reasonably accurately bracket the binding constant. For example, the binding constant of *E. coli* NusB protein and ribosomal protein S10 was estimated at 10^{-7} M based on the observation that S10 protein sedimented faster (with NusB protein) when both were at 6×10^{-7} M, slightly more slowly when both were at 3×10^{-7} M, and much more slowly (midway between its sedimentation position alone and its fully complexed sedimentation position) when both were at 1.5×10^{-7} M (138). There are two reasons that S10 sedimented at an intermediate position rather than at its own position during the run at 1.5×10^{-7} M of each protein. First, the proteins are usually about fivefold more dilute at the end of the sedimentation than when they are first loaded on the gradient; therefore, if S10 protein could bind at the beginning of the run (and sediment faster), it might not bind at the more dilute concentration at the end of the run. Thus, it would sediment at an intermediate position. Second, equilibrium binding is a dynamic process and molecules are constantly associating and dissociating. Therefore, an individual S10 molecule which dissociated from NusB at the trailing edge of the peak would be in a region with very much less NusB to bind. It would sediment at its native rate from that point on.

There are two problems associated with this technique. First, it is not an equilibrium determination, because of the changing conditions during the run. Therefore, failure to detect an interaction may be due to rapid equilibrium rather than a lack of interaction. As such, values obtained from this type of experiment represent an upper bound for the binding constant. Second, sedimentation through gradients does not resolve species that well. Sedimentation rates vary as $M^{2/3}$ for spherical molecules. Thus, dimerization of one spherical molecule with one that is 1/10 the mass will increase its sedimentation rate by only 6%, which is very difficult to detect; in contrast, the change in mobility of the smaller molecule will be fivefold under such conditions.

Although this method has limitations, it has been useful for estimating the upper limit of a binding interaction.

Gel filtration columns. Gel filtration is another simple way of estimating the binding constant. In gel filtration, the elution position of a protein or of a protein complex depends on its Stokes radius. This provides a very powerful and conceptually simple method for evaluating the strength of the interaction between two different proteins. Such sizing columns have been used in three distinct ways to measure or to estimate the binding constant.

(i) Nonequilibrium "small-zone" gel filtration columns. In the simplest approach, a solution containing a protein and a ligand protein is applied in a small volume to the column and the material is resolved in the usual way. This is called a "small-zone" column. The elution positions of the protein and ligand in the mixture are compared with those of the protein and ligand when each is chromatographed individually on the same column. If a complex has formed between the protein and ligand, the complex will elute earlier than either protein alone. From measurements of the concentrations of species required to form a complex, one can estimate the binding constant. This type of experiment has been used, for example, to measure the binding of *E. coli* NusA protein to core RNA polymerase and has yielded values very similar to those determined by fluorescence measurements (76). Similarly, Herberg and Taylor (89) quantitated the interaction of cAMP-dependent protein kinase with both the R1 subunit and PKI in the presence and absence of MgATP.

This direct-application method is not an equilibrium method. Since the concentrations of species change during gel filtration (by diffusion and by dilution), the results are subject to the same sources of error as those of sedimentation through sucrose gradients (see references 2 and 250 for a discussion). Thus, the binding constants calculated in this way can be vastly underestimated, particularly if the complex is in rapid equilibrium (see Fig. 3 of Gegner and Dahlquist [72]) for a vivid contrast between nonequilibrium and equilibrium gel filtration). However, several modeling systems have been described (see reference 211 and references therein).

(ii) Hummel-Dreyer method of equilibrium gel filtration. Gel filtration can also be used as an equilibrium method to establish the binding constant between a protein and its ligand protein. One such method is based on the classic paper by Hummel and Dreyer (102). In this gel filtration method, both the gel filtration buffer and the sample had ligand at the same concentration, but only the sample contained protein. Elution of a protein through such a column caused an increase in the concentration of ligand where the protein eluted, followed by a trough of ligand concentration representing ligand that had been removed in the binding. Evaluation of the binding constant of the protein-ligand complex was simply a matter of knowing the concentration of protein eluted, the free concentration of ligand (set by the column), and the concentration of ligand bound with protein (the concentration of ligand in samples containing protein).

This elegant method has been applied to the interaction of two proteins in only a few cases. As illustrated in Fig. 10, the gel filtration buffer contains protein ligand, and the applied sample contains gel filtration buffer (with the same concentration of protein ligand) as well as the other protein. Gegner and Dahlquist (72) used a column equilibrated with CheW to demonstrate and quantitate the interaction of CheA with CheW. They varied the CheW concentration in the initial sample (while maintaining a constant concentration of CheA in the sample and CheW in the buffer) and quantitated the peak area at the CheW position. The CheW concentration in the sample at which there was no resulting CheW peak or trough represented a sample at true equilibrium. From this, they could calculate a dissociation constant of the complex of 13 μM. A similar series of experiments was done by Yong et al. (243) to demonstrate an interaction between glycerol-3-phosphate dehydrogenase and lactate dehydrogenase over an extremely limited range of NADH concentrations. Such a complex was observed only when the NADH concentration was high enough for an interaction and low enough to be shared by the two enzymes, and it provided evidence for substrate channeling.

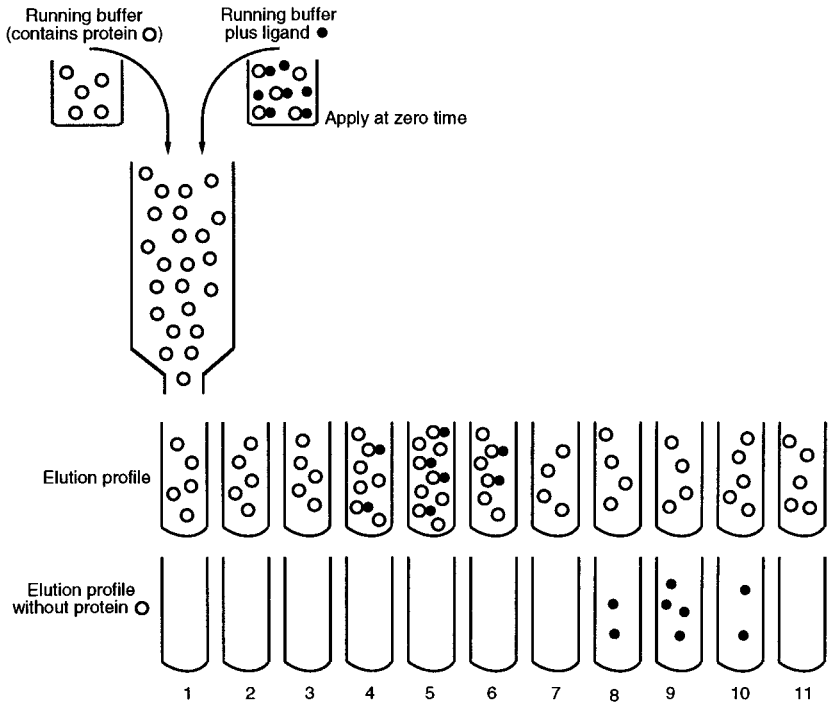

FIG. 10. Equilibrium gel filtration. A solution containing both protein ligand (solid circles) and interacting protein (open circles) is applied to a gel filtration column which is equilibrated with solution containing the interacting protein and developed with running buffer containing the interacting protein. The elution pattern is shown in the first row of test tubes at the bottom. The second row of test tubes indicates the elution pattern that would be observed in the absence of interacting protein.

This method is so simple and inexpensive that it is likely to become much more widely used than at present. Moreover, as an equilibrium experiment, it is without any flaw. The only requirements of this technique are that sufficient protein is available for the experiments and that the elution position of the complex differs from that of at least one of the interacting proteins. With the development of rapid techniques for large-scale protein purification through the use of fusion proteins, it should become relatively routine to obtain enough of any protein to use as a column eluant.

Another variation of Hummel-Dreyer columns is the partitioning method. In this technique, a protein and its ligand protein are mixed with a gel and allowed to equilibrate and the gel is centrifuged or filtered to separate the aqueous phase. From an analysis of the distribution of the protein and the ligand protein in the filtrate and in the gel when they are added separately or together, the K_d can be calculated. An example of this technique is the demonstration of a complex between transaminase and glutamate dehydrogenase which occurs with a dissociation constant of 16 to 67 μM, depending on the presence of various metabolites (63); this is another example of metabolite channeling. This method is also not in wide use, although it seems simple and accurate.

(iii) Large-zone equilibrium gel filtration. One final method of equilibrium gel filtration is the large-zone method (1, 2), in which a very large sample volume is applied to the column, followed by conventional buffer elution. Because a large volume is applied, the concentration of the eluted protein is fixed and constant during the experiment, except at the leading and trailing edges. The elution position of the leading or trailing edge (which measures the size of the molecule) is then monitored as a function of the sample concentration applied to the column. From such experiments, calculation of the dissociation constant is thermodynamically rigorous, as it is for the Hummel-Dreyer method. This large-zone method has been used to monitor self-association of proteins as well as interactions of dissimilar subunits (see, for example, references 75 and 122), but it has received only limited attention because of the large amounts of protein needed to do the experiments.

A variation of this method, first described by Sauer (193), monitors the change in elution position of radiolabeled protein mixed with different concentrations of unlabeled protein in different runs. The use of labeled protein allows simpler and more accurate determination of the elution position, thus allowing Sauer to determine a dimerization constant of 20 nM for repressor. Improvements in protein labeling have demonstrated that the lower limit of detection for this method is a K_d of the order of 10^{-12} M (13).

Sedimentation equilibrium. Although sedimentation equilibrium is a classical method of determining the molecular weight of a protein, it has not been widely used to study protein-protein interactions. However, recent progress makes this method much more accessible on a day-to-day basis (see reference 185 for a recent review). Sedimentation equilibrium can now be done in everyday preparative ultracentrifuges with swinging-bucket rotors, and samples can be readily collected because of the development of a highly reproducible BRANDEL microfraction collector (183). These developments allow the use of a variety of techniques to assay the protein content of each sample, including kinetic assays, radioactive tracers (183), and gel analysis of samples (47); the result is a huge increase in sensitivity over that obtained with the old model E centrifuge (184).

Fluorescence methods. Since fluorescence is a highly sensitive method for detecting proteins through their tryptophan residues, it is potentially a useful way of evaluating protein-protein interactions. Two such methods have been used and are described below.

(i) Fluorescence spectrum. Changes in the fluorescence emission spectrum on complex formation can occur either by a shift in the wavelength of maximum fluorescence emission or by a shift in fluorescence intensity caused by the mixing of two proteins. Therefore, the fluorescence intensity at a particular wavelength can be used to evaluate the dissociation constant. A good example of this technique is illustrated by the interaction of the γ subunit of cGMP phosphodiesterase (PDEγ) subunit with the transducin α subunit (Tα) in the presence of GTPγS or GDP (164).

An equimolar solution of TαGTPγS and PDEγ causes a blue shift in the fluorescence emission spectrum relative to the sum of the individual fluorescence spectra, resulting in a difference spectrum [F (complex) − F (sum)] with a positive component at low wavelengths (320 nm) and a negative component at higher wavelengths (357 nm). Titration of PDEγ into a solution of TαGTPγS therefore caused an enhanced increase in the fluorescence at 320 nm relative to that observed by titration of PDEγ into buffer alone (and a corresponding decrease at 357 nm) until the TαGTPγS was all complexed, after which further addition PDEγ caused no changes in fluorescence intensity relative to that observed in buffer alone. When corrected for PDEγ fluorescence, both curves yielded the same binding curve, and the K_d for the interaction was evaluated at <100 pM. The interaction of TαGDP with PDEγ results in a large increase (ca. 70%) in the intensity of the fluorescence emission spectrum relative to the sum of the individual spectra, and this was used to evaluate the K_d at 2.75 nM.

This technique has two limitations. First, the probability of detecting a change in the fluorescence spectrum decreases with the total number of tryptophan residues in the two proteins, since the fluorescence spectrum is the sum of the contributions from all the tryptophan residues. Since PDEγ has only one tryptophan residue and Tα has two, this condition was easily met in studying the Tα-PDEγ complex. Second, the sensitivity is limited by the intensity of the fluorescence change, which in turn depends on the inherent sensitivity of fluorescence (of the order of nanomolar) and the change that is observed (which is not easily predictable). Thus, the binding constant was too low to evaluate the TαGTPγS-PDEγ interaction (<100 pM) but was high enough to evaluate the interaction in the presence of GDP (2.75 nM).

Although these two limitations exclude the study of many interactions, a number of proteins have a small or limited number of tryptophan residues. For example, bovine Hsc70 has only two tryptophans, and its interaction with small peptides has been evaluated because of the resulting quenching of the fluorescence intensity (123). Similarly, the interaction of angiogenin (one tryptophan) with human placental RNase inhibitor (six tryptophan residues) causes a 50% increase in fluorescence (126), and the dissociation of mitochondrial creatine kinase (four tryptophans per monomer) from octamers to dimers results in a 25% decrease in fluorescence (81).

A second way in which fluorescence is used to measure the interaction of proteins is with a fluorescent tag. This allows for greater sensitivity of monitoring interactions, as long as the fluorescent adducts do not adversely affect the function of the modified protein or its interaction with other proteins. An example of this approach is the interaction of spinach calmodulin with smooth myosin light-chain kinase (146). Calmodulin from spinach has a single cysteine, which could be quantitatively labeled with 2-(4-maleimidoanilino)-naphthalene-6-sulfonic acid (MIANS). Calmodulin labeled with MIANS was as efficient as the wild type in activating calcineurin, in activating cGMP-dependent phosphodiesterase, and in binding terbium. The fluorescence of MIANS-labeled calmodulin increased 80% on binding calcineurin, more than fourfold when bound with myosin light-chain kinase, and twofold on binding caldesmon. In each case, the fluorescence change required the presence of calcium, and titrations were done to measure the K_d (<5, 9, and 250 nM, respectively).

(ii) Fluorescence polarization or anisotropy with tagged molecules. Because of the long lifetimes of excited fluorescent molecules (nanoseconds), fluorescence can also be used to monitor the rotational motion of molecules, which occurs on this timescale. This is accomplished experimentally by the use of plane-polarized light for excitation, followed by measurement of the emission at parallel and perpendicular planes. Since rotational correlation times depend on the size of the molecule (approximately 1 ns/2,400 Da for an idealized molecule), this method can be used to measure the affinity of two proteins for one another because of the increased rotational correlation time of the complex. Fluorescence anisotropy is done most often with a protein bearing a covalently added fluorescent group, which increases both the observed fluorescence lifetime of the excited state and the intensity of the fluorescent signal.

A good example of this technique is described by Weiel and Hershey (229), who studied the interaction of protein synthesis initiation factor 3 (IF3) with 30S ribosomal subunits by using fluorescein-labeled IF3. The labeled protein routinely had about one dye molecule per monomer, and most of the IF3 protein had one or two dye molecules attached. Fluorescein-labeled IF3 was biologically functional: it bound 30S ribosomal subunits, as measured by sucrose density gradients, at a saturable site(s) and had 80 to 100% of the activity of the native protein in stimulating binding of $tRNA^{Met}$ to 70S ribosomes in the presence of RNA. In the presence of 30S ribosomes, both the fluorescence emission spectrum and the fluorescence lifetime of the fluorescein-labeled IF3 were unchanged. Thus, the observed increase in fluorescence polarization which was associated with binding of 30S ribosomes was most consistent with the expected change in polarization as a result of binding a larger molecule. The Scatchard plot derived from the polarization data gave a stoichiometry of 1:1, and the dissociation constant from the polarization data was 3.2×10^{-8} M. Moreover, wild-type nonderivatized IF3 competed for the binding site with the same binding constant. Thus, the fluorescent probe had no effect on any measurable parameter and the measured K_d is likely to be accurate.

Similar experiments have been done with a variety of systems to evaluate the strength of protein-protein interactions. Fluorescein-labeled IF2 was slightly less active than nonderivatized protein, and the binding to 30S ribosomes was twofold weaker than that of the corresponding unlabeled protein (230). T7 gene 2.5 protein labeled with near-molar amounts of fluorescein isothiocyanate caused both a decrease in fluorescence and an increase in anisotropy when bound with T7 DNA polymerase. The fluorescein isothiocyanate-modified protein had no effect on activity, and the binding constant determined by anisotropy (1 μM) was nearly the same as that determined by anisotropy measurements of EDANS-labeled gene 2.5 protein (1.3 μM), for which the rotational correlation time indicated a 1:1 complex (115). The interaction of (florescein-labeled) citrate synthase and malate dehydrogenase was shown to be well within the physiological range (K_d = 1 μM) and varied as much as 25-fold in the presence of different metabolites (214). The tetramer-dimer equilibrium of λ repressor could be observed with dansylated λ repressor, because of its long fluorescence lifetime and high anisotropic value (indicating rigid orientation), but not with fluorescein, which was attached in the highly

mobile N-terminal arm of the repressor molecule (and therefore gave low values) (9).

A variation of this technique has been developed for the interaction of a DNA-binding protein with another protein, in which the DNA is fluorescently labeled (91). In this way, *E. coli* CAP could be shown to interact with RNA polymerase holoenzyme in the presence of cAMP and in the absence of a promoter site. The fluorescently labeled DNA oligonucleotide had a CAP-binding site but no RNA polymerase-binding site, and the resulting increase in polarization allowed the determination of a CAP-RNA polymerase binding constant (2.8×10^{-7} M). Since this interaction was not observed with a CAP mutant protein that was defective in transcription activation, it seems likely that the interaction is important physiologically. Other fluorescent polarization experiments suggest that the CAP-RNA polymerase interaction is much stronger in the presence of cAMP and requires σ factor (170).

Solution equilibrium measured with immobilized binding protein. A simple technique for measuring the dissociation constant of a solution of interacting proteins makes use of bound competitor protein to determine the amount of free protein in such a solution. This method was first described for antibody-antigen reactions (71) and later modified for general use to determine the interaction of C4b-binding protein (C4BP) with human protein S (HPS) (158). A solution containing C4BP and HPS was incubated until equilibrium was reached. The amount of free C4BP in the solution was then determined by incubating an aliquot on a plate containing immobilized HPS under conditions (short incubation time) in which a limited amount of the free C4BP binds the immobilized HPS. This resulted in little perturbation of the equilibrium during the assay for C4BP retained by the immobilized HPS, which was quantitated by an antibody-based method.

This method requires satisfaction of three criteria. First, the two proteins (HPS in solution and HPS immobilized on the plate) cannot bind each other. If they did, C4BP could be captured through HPS-HPS interactions. Second, HPS in solution and HPS immobilized on the plate must compete for the same binding site. This is obviously true in this case, but it is not necessarily true if, for example, anti-C4BP is used in the immobilized system to detect the amount of free C4BP. Third, the method requires that only free C4BP be measured during the incubation with immobilized HPS. This in turn requires that binding to the immobilized HPS remove only a small portion of the total C4BP ($<10\%$ was removed in this example) so that equilibrium of the solution is perturbed as little as possible. This condition also requires that the off rate of the complex is low compared with the time of incubation with the immobilized HPS; otherwise, HPS-C4BP complexes could dissociate during the incubation with immobilized HPS and the dissociated C4BP would be measured as free C4BP. Thus, this method, although simple, provides only an upper bound of the dissociation constant.

Surface plasmon resonance. The recent development of a machine to monitor protein-protein and ligand-receptor interactions by using changes in surface plasmon resonance measured in real time spells the beginning of a minor revolution in biology. This method measures complex formation by monitoring changes in the resonance angle of light impinging on a gold surface as a result of changes in the refractive index of the surface up to 300 nm away. A ligand of interest (peptide or protein in this case) is immobilized on a dextran polymer, and a solution of interacting protein is flowed through a cell, one wall of which is composed of this polymer. Protein that interacts with the immobilized ligand is retained on the polymer surface, which alters the resonance angle of impinging light as a result of the change in refractive index brought about by increased amounts of protein near the polymer. Since all proteins have the same refractive index and since there is a linear correlation between resonance angle shift and protein concentration near the surface, this allows one to measure changes in protein concentration at the surface due to protein-protein or protein-peptide binding. Furthermore, this can be done in real time, allowing direct measurement of both the on rate and the off rate of complex formation. A good layman's review of surface plasmon resonance is found in articles by Malmqvist (136) and Jonsson et al. (109), and a clear derivation of the appropriate equations is found in the article by Karlsson et al. (111).

In practice, determination of a binding constant requires measurement of two parameters. First, the increase in RU (resonance units) is measured as a function of time by passing a solution of interacting protein past the immobilized ligand until (usually) the RU values stabilize. Second, the decrease in RU is measured as a function of time with buffer lacking interacting protein. This produces a sensorgram for each concentration of protein, a continuous recording of RU versus time. This procedure is then repeated at a number of protein concentrations, after regeneration of the dextran surface. From these two sets of data, two lines are constructed whose slopes correspond to k_a (the on rate) and k_d (the off rate); from these data, K_d is calculated as k_d/k_a. An alternative determination of K_d can be made by using the steady-state RU values at different protein concentrations.

This system has several advantages. First, it requires very little material. Typically only 1 to 10 μg of protein has to be immobilized on a sensor chip, which can be reused up to 50 times after removal of adhering protein. Similarly, solutions of interacting protein are in the range of 0.01 to 1 ml, depending on the chosen flow rate (109). Second, the method is very fast. A typical run for a given protein takes about 10 min. Third, no modifications of the proteins are required, such as labeling or fluorescent tags. Fourth, interactions can be observed even in complex mixtures. Fifth, both the on rate and the off rate are readily obtained. Sixth, the system is useful over a wide range of protein concentrations. The practical lower limit of the original Biacore system is a change in resonance angle of 10^{-3} degrees (10 RU), corresponding to surface concentrations of 10 pg/mm^2; moreover, the system is linear up to RU values of 30,000 (109). Seventh, the system is quite sensitive; the practical limit for association rates is 10^6/M/s, and off rates as low as 1.1×10^{-5}/s have been measured by recording for 6 h with buffer (197).

This technique has been used successfully to monitor protein-peptide interactions. A good example is the determination of the binding interaction of different SH2 domains with two tyrosine-phosphorylated substrate peptides derived from platelet-derived growth factor (166). The corresponding peptides were attached to the dextran polymer chip via avidin on the chip and biotin on the peptides. Subsequent real-time analysis demonstrated that interaction of these peptides with the p85 subunit of phosphatidylinositol-3-kinase (PI3K) was characterized by a very high association rate (2×10^6/M/s) and dissociation rate (0.1/s) for the 12-mer peptide Y740P and that most of this binding was contributed by the C-terminal subunit of p85. In this particular case, the dissociation rate of bound p85 had to be determined in the presence of a sink of excess competing peptide in the buffer; otherwise, rebinding of dissociated p85 was a significant problem because of the very high on rate. A similar study of p85 SH2 domain interactions with different tyrosine-phosphorylated peptides (from IRS-1) led to the same conclusions of a high on rate and off rate, which was

also measured in the presence of a sink of peptide (64). In this case, the on rate was too high to measure directly (as high as 4.4×10^8/M/s for the C-terminal SH2 domain of p85) and was instead inferred from steady-state binding and off rate measurements and confirmed by competition experiments with free phosphorylated peptide (64). On rates in excess of 10^6/M/s can be limited by mass transport rates (fluid flow through the cell) rather than binding-reaction rates, although this can be partially compensated by either higher flow rates or a smaller amount of peptide on the chip (111). Competition experiments were also used to show that the affinity of p85 for phosphorylated peptides was 300- to 800-fold greater than for the corresponding nonphosphorylated peptide and was as much as 100-fold weaker with a glycine or arginine at the +1 position relative to the tyrosine compared with bulky hydrophobic groups or glutamate (64).

One final study demonstrated that a specific threonine residue in the SH2 domain of Src, when changed to a tryptophan, increased the affinity of the domain for phosphorylated peptides which were substrates for GRB2 and that the corresponding tryptophan of GRB2, when altered to threonine, weakened the affinity of GRB2 for this peptide (137). In each of these three examples, the primary determinant of specificity was the on rate rather than the off rate.

Surface plasmon resonance has also been used with great success to monitor protein-protein interactions. One such example is the demonstration of a quarternary complex of CheY with CheA, CheW, and Tar (197). CheY was bound to the dextran surface through a unique (and engineered) cysteine residue, which did not affect chemotaxis activity and which was remote from the interaction domain (197). CheA binds this immobilized CheY protein with a low association rate (368/M/s) and a very low off rate (1.14×10^{-5}/s). Moreover, CheA, CheW, and Tar probably form a quaternary complex with CheY; addition of all three proteins greatly increases the amount of protein bound to CheY relative to that obtained with CheA alone, although neither Tar nor CheW binds CheY individually or when present together.

Other examples of protein-protein interactions studied by surface plasmon resonance include the interaction of monoclonal antibodies with human immunodeficiency virus type 1 core protein p24 (111), EGF with the EGF receptor (249), the regulatory and catalytic domains of cAMP-dependent protein kinase (88), and VAMP2 and sytaxin 1A (27).

Two minor problems are associated with surface plasmon resonance measurements. First, immobilization of the ligand protein must be of such a nature that it does not impede or artificially enhance interactions. This is the same problem that is associated with protein affinity columns. Attachment of CheY was accomplished by using a single site remote from the interaction domain (197); this presents the interacting face to the solvent. Phosphorylated peptides were attached by biotinylation of the peptide at a single site (but variable position) with a long spacer followed by noncovalent interaction with an avidin-coupled sensor chip (166), and attachment of monoclonal antibodies to the chip was accomplished through noncovalent binding to covalently coupled rabbit anti-mouse IGGFc (111). Primary amines are often linked directly to the dextran polymer, leading to more homogeneous presentation of surfaces to the solvent but causing possible inhomogeneities in interaction (88). Second, the sensor chip has to be regenerated under conditions which do not denature the immobilized ligand protein. Protein adhering to the immobilized C subunit of protein kinase A was removed with cAMP (88), proteins binding to immobilized phosphorylated peptides were removed with a pulse of dilute SDS (166), and CheY was regenerated with a pulse of guanidine hydrochloride (197). In some cases, the ligand is deliberately removed before the next experiment; thus, monoclonal antibodies sticking to IGGFc were removed with dilute HCl before readdition of the monoclonal antibodies to act as a ligand for p24 binding (111).

Limits to Detection

Determination of the binding constant of tightly interacting species by standard methods described above depends on being able to determine and quantitate the fraction of protein ligand bound at a given protein concentration that spans the dissociation constant. For a standard 50,000-kDa protein, the practical limit of silver staining is of the order of 0.2 ng or 20 μl of a 10-ng/ml solution, which would be useful for a dissociation constant of 1 nM or greater. For in vitro translated protein, the practical limit is 1,000 Ci/mmol times the number of amino acid residues, or 1,000 dpm of ^{35}S-labeled protein per fmol (singly labeled); this corresponds to 10^{-12} M or, with 10 residues incorporated, 10^{-13} M; therefore, allowing for concentrations below K_d, the lower limit of detection is of the order of 10^{-12} M.

Some protein-protein interactions are too tight ($K_d < 10^{-12}$ M) to measure by the methods described above. For example, human placental RNase inhibitor (PRI) interacts very tightly with both angiogenin ($K_d = 7 \times 10^{-16}$ M) (126, 126a) and human placental RNase ($K_d = 9 \times 10^{-16}$ M) (199). For the interaction of PRI with angiogenin, the association rate constant, k_a, was measured by monitoring the change in intrinsic fluorescence by stopped-flow fluorescence techniques, and the dissociation rate constant, k_d, was measured by measuring the release of PRI in the presence of scavenger RNase, to which it binds and inhibits the activity.

A dissociation constant of the magnitude of 7×10^{-16} M for the PRI-angiogenin interactions means that the dissociation rate is measured in weeks! In this case, the $t_{1/2}$ for dissociation of the complex was 60 days (corresponding to $k_d = 1.3 \times 10^{-7}$/s). Furthermore, the overall on rate of 1.8×10^8/M/s liters · mol/s is near the diffusion limit for molecules of the size of proteins. It is hard to imagine what selective pressure would require or maintain such a tight interaction. This is particularly true since human placental RNase and angiogenin both bind PRI equally tightly and are substantially different at the amino acid level.

It is possible that a number of macroscopic protein-protein interactions operate at this level. Any protein composed of three or more subunits can have significant interactions among individual pairs of the component protein. If, for example, a subunit has a K_d of 10^{-7} M with each of two other subunits, the effective K_d of the dissociation of that subunit from the complex is 10^{-14} M (see reference 116 for a discussion of this point). Thus, complicated structures like the ribosome might effectively lock the proteins together in undissociable units. It is also possible that other, simpler interactions are this tight; the dissociation rate of the subunits of a number of proteins that purify as a complex tends never to be investigated.

EXAMPLES OF WELL-CHARACTERIZED DOMAINS

Given that a straightforward set of experiments is all that is required nowadays to identify two proteins that interact and to delineate the domains responsible for the binding, toward what ends does this analysis continue? To address this question, it is instructive to consider the case of some domains involved in protein-protein interaction that have been extensively characterized. Using a combination of numerous techniques, includ-

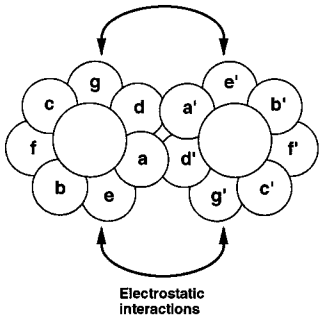

FIG. 11. Helical wheel representation of a leucine zipper. Adapted from reference 221a with permission of the publisher.

ing detailed structural approaches, investigators who have focused on the analysis of leucine zippers, SH2 domains, and SH3 domains have made tremendous advances in the last few years. These studies have considerably extended our understanding of transcriptional regulation and signal transduction. In the next sections, we provide a brief view of how these three domains function.

Leucine Zipper

The leucine zipper is a protein-protein interaction motif in which there is a cyclical occurrence of leucine residues every seventh residue over short stretches of a protein in an α-helix. These leucine residues project into an adjacent leucine zipper repeat by interdigitating into the adjacent helix, forming a stable coiled-coil. This motif was first described by Landschulz et al. (124) in connection with a new structure within DNA-binding proteins that might be responsible for interactions with a similar motif to promote specific DNA binding by basic amino acid residues adjacent to the leucine zipper motif (hence the name bZIP). The leucine zipper model was originally proposed on the basis of the leucine distribution and amino acid sequence of regions of C/EBP, Myc, Fos, Jun, and Gcn4. It is now known to be common to over 30 proteins (59). Subsequent experiments have confirmed the existence of this structure and have extended these observations.

Structure. The X-ray structure of the Gcn4 leucine zipper region (consisting of 33 amino acids) demonstrates that the leucine zipper consists of two parallel coiled coils of α-helices wrapped around each other and forming one-quarter of a turn of a left-handed supercoil (59, 161; also see reference 4). The dimer forms a smoothly bent cylinder about 45 Å (4.5 nm) long and 30 Å (3 nm) wide. On a helical-wheel representation of the α-helix (Fig. 11), the leucines occupy position d (and d′ of the adjacent helix) and share the interior with the residues at position a (a′), as well as parts of residues e and g (and e′ and g′). The packing corresponds to the "knobs into holes" model proposed by Crick (42), in which each interior amino acid residue is packed into a gap formed by four nearest neighbors from the opposite helix. More than 95% of the surface area that is buried upon dimerization is from the side chains of these residues.

Stability. The leucine zipper coiled coil is stabilized because of three factors: the hydrophobic groups that are buried (leucines at position d and hydrophobic or neutral residues at position a); constancy of size of the internally packing residues at each position; and several distinct ion pairs. Three such ion pairs appear to form, and each is between the e of one heptad and the g of the other. The leucine residues are critical for function in Gcn4. Although each individual leucine can itself be replaced by several different hydrophobic residues, randomized substitution of the leucines with other hydrophobic residues invariably causes the protein to lose function when more than one leucine is substituted; furthermore, isoleucine is by far the most easily tolerated substitution (98).

The binding constant of leucine zipper moieties that interact is estimated to be in the nanomolar range (163) and has been measured at 5×10^{-8} M for the Jun-Jun dimer at 4°C (196). Even a peptide corresponding to the Fos leucine zipper, which does not dimerize in vitro, has been shown to dimerize in the micromolar range (163).

The leucine zipper moieties that naturally interact do not necessarily have the maximal stability. For example, the Gcn4 dimer has a buried asparagine residue which is present within the hydrophobic core (59, 161). This Asn residue packs loosely in the crystal structure, and this position is particularly tolerant of other amino acids (98). Moreover, the asparagine residue (and resultant internal hydrogen bond) drastically destabilizes the Gcn4 zipper; its replacement with valine stabilizes the coiled coil about 1,000-fold (28). It has been speculated that the internal asparagine of Gcn4 (and, by extension, other buried polar groups in the a position in other leucine zippers) is present, so that the proteins do not bind too tightly and therefore can be subject to regulation, or that it keeps the coiled coils in register (4).

Specificity. The specificity of leucine zippers is the key to their regulatory properties. The oncoproteins Fos and Jun, for example, associate with each other to form a heterodimer in preference to the Jun-Jun homodimer. This preference has important consequences in that Fos-Jun heterodimers and Jun-Jun homodimers bend DNA in opposite orientations (114), which may explain the fact that Jun interaction with the glucocorticoid response element of the prolactin gene results in activation of the gene, whereas Fos-Jun interaction results in repression (51).

Specificity of Fos-Jun and Jun-Jun dimerization is achieved primarily by the electrostatic interactions of residues at the e and g positions at the periphery of the hydrophobic core (162). Fos has Glu residues at the g position, and Fos-Fos dimers are much more stable (as measured by T_m) at pH values at which these Glu residues are neutralized. Conversely, Jun is slightly more basic at the e and g positions, and Jun-Jun dimers are more stable at higher pH. Fos-Jun dimers, which are the preferential form, are uniformly stable over a wide range of pH values, because they are more neutral overall. A series of hybrid peptides in an otherwise Gcn4 peptide illustrate the point (162). Specificity (or antispecificity) is achieved by the 8 amino acids at the e and g positions of the peptide and not at other positions.

Regulation. Leucine zipper proteins are likely to be functionally regulated. Thus, the carboxyl-terminal zipper of the human and *Drosophila* heat shock factors may suppress formation of amino-terminal zippers in a way that is sensitive to heat shock (175). Similarly, the calphotin protein binds calcium at one end and has a distinctive leucine zipper at the other end (8). It may therefore be used to transmit signals by altering binding properties.

SH2 Domain

The SH2 domain was first recognized as a noncatalytic domain of Src that was homologous to the Fps protein (189) and is now recognized as a common motif involved in protein-protein interactions (117, 168). More than 20 SH2-containing proteins have been identified. They share a motif of about 100 amino acids that is involved in the recognition of proteins and

peptides containing phosphorylated tyrosines. This recognition is implicated in the mechanism of signal transduction, because the phosphorylated tyrosines that are recognized include those of growth factor receptors such as the platelet-derived growth factor receptor, the EGF receptor, and the fibroblast growth factor receptor. On binding their respective growth factors, the growth factor receptors have their tyrosine kinase activity activated, which allows them to autophosphorylate. The autophosphorylated receptor then binds various proteins containing SH2 domains, which are then phosphorylated to modulate their activity. Thus, the binding of growth factor on the outside of the cell results in phosphorylation on the inside of specific substrate proteins. The particular proteins that are phosphorylated depend on the binding specificity of the SH2 domains for the phosphorylated receptor. Binding of different peptides to different SH2 domains has yielded the following results.

Binding of SH2 proteins requires a large domain of the SH2 protein. The conserved domain of SH2 domains, which is common to more than two dozen proteins, has been crystallized for Src (224, 225) and solved by nuclear magnetic resonance spectroscopy techniques for c-Abl (165) and p85a of PI3K (20). In each case, this domain folds into a structure in which a set of internal antiparallel sheets is surrounded by two more or less symmetrical α-helices. The conserved amino acids tend to be part of the recognition for phosphotyrosine (e.g., Arg-175 of Src) or part of the hydrophobic pocket. Variable regions are responsible for sequence recognition (205) and may be parts of variable loops of unknown function (188).

Binding of SH2 proteins requires phosphorylated tyrosine in vitro. Thus, the binding constant of a peptide to an SH2 protein of p85 is between 50- and 800-fold weaker without the phosphate than with the phosphate (64). This preference is attributable to specific side chain contacts of the SH2 domain with the phosphoryl group of phosphotyrosine. The phosphoryl oxygens are hydrogen bonded with two guanidinium hydrogens, one from one arginine and one from another arginine, one hydroxyl hydrogen from threonine and one from serine, and a backbone amide hydrogen. One of the arginines appears to be acting both as a hydrogen bond donor and as an ion pair with the phosphate group. Thus, it cannot be substituted with lysine without loss of binding (140). These contacts are the same whether a weak-affinity (224) or a strong-affinity (225) phosphotyrosine-containing peptide is used.

SH2 domains make contacts with only a small region surrounding the phosphorylated tyrosine. Small peptides faithfully reproduce binding to SH2 domains and display binding constants of the order of nanomolar (64, 218). This is consistent with the crystallographic data of the SH2 domain of v-Src bound to a high-affinity 11-amino-acid peptide; the data clearly show significant peptide-protein interactions at 6 of the 11 positions of the phosphopeptide, from −2 to +3, relative to the tyrosine residue (225). These are the residues that have associated high electron density, indicating a fixed position in the crystal (except for the side chain portion of Gln-1). In addition to the phosphotyrosine-binding interactions described above, there are several ring interactions that define the rest of the phosphotyrosine pocket. There is also a very well-defined interaction of isoleucine at +3 with a deep pocket in the SH2 domain that results in protection of 95% of the surface of the amino acid side chain. The two glutamate residues at +1 and +2 are on the surface of the protein and largely exposed to solvent. Glu+1 appears to interact through its carboxyl group with a lysine amino group, and Glu+2 appears to be stabilized by a nearby arginine guanidinium and its associated H_2O molecules. The amino acids at positions −1 and −2 appear to cap the phosphotyrosine binding through the polypeptide backbone at position −1 and the proline ring at −2.

Other SH2 domain proteins bind different peptides through interactions at the same +1 to +3 positions relative to the phosphotyrosine. This has been elegantly investigated by Songyang et al. (205) through a study of selectivity of binding of random peptides to different SH2 domains. Although the results obtained in this experiment represent bulk selectivity for certain amino acids at certain positions relative to phosphotyrosine, rather than selectivity of individual peptides of known sequence, the results are clear. Each of the three positions following the phosphotyrosine plays an important role in determining the selectivity of binding in certain SH2 proteins, but the amino acids that are crucial and the extent to which they are crucial differs markedly. Thus, most of the discrimination of the C-terminal SH2 domain of p85 is due to its preference for methionine at +3, whereas most of the discrimination of Nck is at positions 1 and 2, where it prefers glutamate and aspartate, respectively (205).

SH3 Domain

The SH3 domain is a second noncatalytic domain of Src which is involved in protein-protein interactions and which is part of a motif shared by other proteins, including tyrosine kinases, phospholipase C-γ (PLC-γ) PI3K, GTPase-activating protein, the cell proliferation proteins Crk and Grb2/Sem5, and the cytoskeletal proteins spectrin, myosin 1, and an actin-binding protein (see references 117, 120, 154, and 168 for a recent list). More than 27 proteins have been shown to have an SH3 domain, which varies between about 55 and 75 amino acids, and its structure has been determined from four different specific domains: spectrin (154), Src (245), PI3K (120), and PLC (118). Each such structure is composed of antiparallel sheets oriented more or less at right angles to one another (or, for PLC, two partial greek key motifs of a barrel oriented such that the strands on opposite sides cross almost perpendicularly), and the amino acids in the conserved strands and a conserved C-terminal 3_{10} helix correspond to many of those that are conserved among SH3 proteins. In each case, a hydrophobic pocket is formed on the surface of the molecule; those of PI3K and Src are remarkably similar (120), and the location of the pocket is conserved between PLC and spectrin (118). This hydrophobic pocket has been implicated in peptide binding for Src (245), since binding of such a peptide perturbs the signal from these amino acids. There are notable differences among the protein structures; PLC, for example, is very similar in secondary structure to spectrin but not to Src, leading to different architectures (118). This property presumably leads to different binding specificities.

The substrates to which SH3-containing proteins bind include an uncharacterized protein similar to GTPase-activating protein-rho, detected with Abl (36); mSos1 and hSos1 (proteins similar to *Drosophila* Sos, which is required for Ras signaling), detected with Grb2 (187); formin and the rat m4 muscarinic receptor, detected with Abl (181); PI3K, detected with v-Src (130); and $p56^{lck}$ and $p59^{fyn}$ (172, 173).

Like the SH2 domain, the SH3 domain binds simple peptides with a high degree of sequence specificity and a high affinity. As judged on a qualitative basis, a 10-amino-acid proline-rich sequence is responsible for strong binding of the Abl SH3 domain to two proteins, called 3BP-1 and 3BP-2 (36, 181). This binding is specific in two ways. First, some but not all single-amino-acid alterations destroy detectable binding. Thus, prolines at positions 2, 7, and 10 are important but those at 5 and to some extent 9 are not. Nonproline residues do not

appear to be as important, except perhaps at position 1 (181). Second, peptide binding is SH3 domain specific. Thus, 3BP1 binds the SH3 domains of Abl and Src but not those of Neural Src or Crk (36), and 3BP2 binds most strongly to Abl SH3, less so to Src SH3 and Grb2, and poorly to Nck (181).

Similarly, binding of mSos1 to Grb2 appears to be through a proline-rich motif at the C terminus of the protein (187); any of several proline-rich 11-amino-acid peptides corresponding to sequences in this region all compete, and competition appears to require a C-terminal arginine. This arginine may add selectivity to the binding of mSos1 to Grb2. A peptide containing the relevant arginine-containing motif binds to Grb2 through its SH3 domain with a K_d of 25 nM (128).

CONCLUDING REMARKS

Alberts and Miake-Lye (5), summarizing a meeting entitled Proteins as Machines, described Tom Pollard's flow diagram for the detailed analysis of a cell biology process. First, a complete inventory of all the molecules making up the machine must be made. Second, a determination must be made of how and in what order the molecules interact with each other. Third, both detailed rate constants for each transition and structures of each component at atomic resolution must be obtained. While no process is yet completely understood at the three levels described by Pollard, enormous progress has been made in deciphering protein machines. In this review, we have tried to convey some of the classical and more recent approaches used to develop the inventory of proteins and the nature of their interactions.

Two factors are having a large impact on how cellular processes are viewed. First, the vast amount of DNA sequence information being obtained means that the identity of almost all proteins, at the level of primary sequence, may soon be known. Complete sequences for organisms such as *E. coli*, yeast cells, and the nematodes and nearly complete compilations of the cDNA sequences for human tissues should be available in the next few years. Second, the range of new procedures now available means that hundreds to thousands of new protein-protein interactions may be identified in the same period. Ten to twenty years ago, only a few complexes of proteins were well characterized as to their subunit composition and specific interactions; currently, a large number of such complexes are known. Relatively soon, there may be an enormous number. The continuing challenge will be for biochemists and cell, molecular, and structural biologists to use this information to understand how the cell works.

ACKNOWLEDGMENTS

We thank R. Angerer, M. Dumont, E. Grayhack, T. Miller, F. Sherman, and L. Tabak for many helpful comments on the manuscript and J. Nimmo for superb secretarial assistance.

Research from our laboratories has been supported by grants MCB-9118571 from the National Science Foundation and VM-99 from the American Cancer Society (to E.M.P.) and grants GM 49065 and CA 28146 from the National Institutes of Health (to S.F.).

REFERENCES

1. **Ackers, G. K.** 1970. Analytical gel chromatography of proteins. Adv. Protein Chem. **24:**343–441.
2. **Ackers, G. K.** 1975. Molecular sieve methods of analysis, p. 1–94. *In* H. Neurath and R. L. Hill (ed.), The proteins. Academic Press, Inc., New York.
3. **Adams, A. E. M., and D. Botstein.** 1989. Dominant suppressors of yeast actin mutations that are reciprocally suppressed. Genetics **121:**675–683.
4. **Alber, T.** 1992. Structure of the leucine zipper. Curr. Opin. Genet. Dev. **2:**205–210.
5. **Alberts, B., and R. Miake-Lye.** 1992. Unscrambling the puzzle of biological machines: the importance of the details. Cell **68:**415–420.
6. **Ayer, D. E., L. Kretzner, and R. N. Eisenman.** 1993. Mad: a heterodimeric partner for Max that antagonizes Myc transcriptional activity. Cell **72:**211–222.
7. **Baird, B. A., and G. G. Hammes.** 1976. Chemical cross-linking studies of chloroplast coupling factor 1. J. Biol. Chem. **251:**6953–6962.
8. **Ballinger, D. G., N. Xue, and K. D. Harshman.** 1993. A *Drosophila* photoreceptor cell-specific protein, calphotin, binds calcium and contains a leucine zipper. Proc. Natl. Acad. Sci. USA **90:**1536–1540.
9. **Banik, U., N. C. Mandal, B. Bhattacharyya, and S. Roy.** 1993. A fluorescence anisotropy study of tetramer-dimer equilibrium of λ repressor and its implication for function. J. Biol. Chem. **268:**3938–3943.
10. **Barbas, C. F., A. S. Kang, R. A. Lerner, and S. J. Benkovic.** 1991. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. Proc. Natl. Acad. Sci. USA **88:**7978–7982.
11. **Bass, S., R. Greene, and J. A. Wells.** 1990. Hormone phage: an enrichment method for variant proteins with altered binding properties. Proteins **8:**309–314.
12. **Basson, M. E., R. L. Moore, J. O'Rear, and J. Rine.** 1987. Identifying mutations in duplicated functions in *Saccharomyces cerevisiae* by use of the *LYS2* gene. Genetics **117:**645–655.
13. **Beckett, D., K. S. Koblan, and G. K. Ackers.** 1991. Quantitative study of protein association at picomolar concentrations: the λ phage cl repressor. Anal. Biochem. **196:**69–75.
14. **Beeckmans, S., and L. Kanarek.** 1981. Demonstration of physical interactions between consecutive enzymes of the citric acid cycle and of the aspartate-malate shuttle. Eur. J. Biochem. **117:**527–535.
15. **Bender, T., and J. R. Pringle.** 1991. Use of a screen for synthetic lethal and multicopy suppressee mutants to identify two new genes involved in morphogenesis in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **11:**1295–1305.
16. **Birge, R. B., J. E. Fajardo, C. Reichman, S. E. Shoelson, Z. Songyang, L. C. Cantley, and H. Hanafusa.** 1993. Identification and characterization of high-affinity interaction between v-Crk and tyrosine-phosphorylated paxillin in CT10-transformed fibroblasts. Mol. Cell. Biol. **13:**4648–4656.
17. **Blackwood, E. M., and R. N. Eisenman.** 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. Science **251:**1211–1217.
18. **Blanar, M. A., and W. J. Rutter.** 1992. Interaction cloning: Identification of a helix-loop-helix zipper protein that interacts with c-Fos. Science **256:**1014–1018.
19. **Bleil, J. D., and P. M. Wassarman.** 1990. Identification of a ZP3-binding protein on acrosome-intact mouse sperm by photoaffinity crosslinking. Proc. Natl. Acad. Sci. USA **87:**5563–5567.
20. **Booker, G. W., A. L. Breeze, A. K. Downing, G. Panayotou, I. Gout, M. D. Waterfield, and I. D. Campbell.** 1992. Structure of an SH2 domain of the p85α subunit of phosphatidylinositol-3-OH kinase. Nature (London) **358:**684–687.
21. **Bragg, P. D., and C. Hou.** 1980. A cross-linking study of the Ca^{2+}, Mg^{2+}-activated adenosine triphosphatase of *Escherichia coli*. Eur. J. Biochem. **106:**495–503.
22. **Bragg, P. D., and C. Hou.** 1986. Chemical crosslinking of α subunits in the F_1 adenosine triphosphatase of *Escherichia coli*. Arch. Biochem. Biophys. **244:**361–372.
23. **Brent, R., and M. Ptashne.** 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell **43:**729–736.
24. **Bretscher, A., and K. Weber.** 1980. Villin is a major protein of the microvillus cytoskeleton which binds both G and F actin in a calcium-dependent manner. Cell **20:**839–847.
25. **Brown, R. L.** 1992. Functional regions of the inhibitory subunit of retinal rod cGMP phosphodiesterase identified by site-specific mutagenesis and fluorescence spectroscopy. Biochemistry **31:**5918–5925.
26. **Burton, D. R., C. F. Barbas, M. A. A. Persson, S. Koenig, R. M. Chanock, and R. A. Lerner.** 1991. A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. Proc. Natl. Acad. Sci. USA **88:**10134–10137.
27. **Calakos, N., M. K. Bennett, K. E. Peterson, and R. H. Scheller.** 1994. Protein-protein interactions contributing to the specificity of intracellular vesicular trafficking. Science **263:**1146–1149.
28. **Carr, C., J. Petersen, P. S. Kim, and T. Alber.** 1992. Unpublished data.
29. **Carr, D. W., and J. D. Scott.** 1992. Blotting and band-shifting: techniques for studying protein-protein interactions. Trends Biochem. Sci. **17:**246–249.
30. **Chardin, P., J. H. Camonis, N. W. Gale, L. V. Aelst, J. Schlessinger, M. H. Wigler, and D. Bar-Sagi.** 1993. Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2. Science **260:**1338–1343.
31. **Chellappan, S., V. B. Kraus, B. Kroger, K. Munger, P. M. Howley, W. C. Phelps, and J. R. Nevins.** 1992. Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. Proc. Natl. Acad. Sci. USA **89:**4549–4553.
32. **Chellappan, S. P., S. Hiebert, M. Mudryj, J. M. Horowitz, and J. R. Nevins.**

1991. The E2F transcription factor is a cellular target for the RB protein. Cell **65:**1053–1061.
33. **Chen, Y., Y. W. Ebright, and R. H. Ebright.** 1994. Identification of the target of a transcription activator protein by protein-protein photocrosslinking. Science **265:**90–92.
34. **Chevray, P. M., and D. Nathans.** 1992. Protein interaction cloning in yeast: identification of mammalian proteins that react with the leucine zipper of Jun. Proc. Natl. Acad. Sci. USA **89:**5789–5793.
35. **Chien, C.-T., P. L. Bartel, R. Sternglanz, and S. Fields.** 1991. 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 **88:**9578–9582.
36. **Cicchetti, P., B. J. Mayer, G. Thiel, and D. Baltimore.** 1992. Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. Science **257:**803–806.
37. **Clackson, T., H. R. Hoogenboom, A. D. Griffiths, and G. Winter.** 1991. Making antibody fragments using phage display libraries. Nature (London) **352:**624–628.
38. **Cohen, B. D., D. R. Lowy, and J. T. Schiller.** 1993. The conserved C-terminal domain of the bovine papillomavirus E5 oncoprotein can associate with an α-adaptin-like molecule: a possible link between growth factor receptors and viral transformation. Mol. Cell. Biol. **13:**6462–6468.
39. **Cornish, V. W., D. R. Benson, C. A. Altenbach, K. Hideg, W. L. Hubbell, and P. G. Schultz.** 1994. Site-specific incorporation of biophysical probes into proteins. Proc. Natl. Acad. Sci. USA **91:**2910–2914.
40. **Costigan, C., S. Gehrung, and M. Snyder.** 1992. A synthetic lethal screen identifies SLK1, a novel protein kinase homolog implicated in yeast cell morphogenesis and cell growth. Mol. Cell. Biol. **12:**1162–1178.
41. **Cover, J. A., J. M. Lambert, C. M. Norman, and R. R. Traut.** 1981. Identification of proteins at the subunit interface of the *Escherichia coli* ribosome by cross-linking with dimethyl 3,3′-dithiobis(propionimidate). Biochemistry **20:**2843–2852.
42. **Crick, F. H. C.** 1953. The packing of α-helices: simple coiled-coils. Acta Crystallogr. **6:**689–697.
43. **Cull, M. G., J. F. Miller, and P. J. Schatz.** 1992. Screening for receptor ligands using large libraries of peptides linked to the C terminus of the *lac* repressor. Proc. Natl. Acad. Sci. USA **89:**1865–1869.
44. **Cwirla, S. E., E. A. Peters, R. W. Barrett, and W. J. Dower.** 1990. Peptides on phage: a vast library of peptides for identifying ligands. Proc. Natl. Acad. Sci. USA **87:**6378–6382.
45. **Dalton, S., and R. Treisman.** 1992. Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. Cell **68:**597–612.
46. **Dang, C. V., J. Barrett, M. Villa-Garcia, L. M. S. Resar, G. J. Kato, and E. R. Fearon.** 1991. Intracellular leucine zipper interactions suggest c-Myc hetero-oligomerization. Mol. Cell. Biol. **11:**954–962.
47. **Darawshe, S., G. Rivas, and A. P. Minton.** 1993. Sedimentation equilibrium-quantitative polyacrylamide gel electrophoresis (SE-QPAGE): a new technique for the detection of associations in multicomponent solutions. Anal. Biochem. **215:**236–242.
48. **de Gunzberg, J., R. Riehl, and R. A. Weinberg.** 1989. Identification of a protein associated with $p21^{ras}$ by chemical crosslinking. Proc. Natl. Acad. Sci. USA **86:**4007–4011.
49. **Denny, J. B., and B. Blobel.** 1984. ^{125}I-labeled crosslinking reagent that is hydrophilic, photoactivatable, and cleavable through an azo linkage. Proc. Natl. Acad. Sci. USA **81:**5286–5290.
50. **Devlin, J. J., L. C. Panganiban, and P. E. Devlin.** 1990. Random peptide libraries: a source of specific protein binding molecules. Science **249:**404–406.
51. **Diamond, M. I., J. N. Miner, S. K. Yoshinaga, and K. R. Yamamoto.** 1990. Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. Science **249:**1266–1272.
52. **Dobzhansky, T.** 1946. Genetics of natural populations. XIII. Recombination and variability in populations of *Drosophila pseudoobscura*. Genetics **31:**269–290.
53. **Drubin, D. G., K. G. Miller, and D. Botstein.** 1988. Yeast actin-binding proteins: evidence for a role in morphogenesis. J. Cell Biol. **107:**2551–2561.
54. **Dumont, M. E., A. F. Corin, and G. A. Campbell.** 1994. Noncovalent binding of heme induces a compact apocytochrome *c* structure. Biochemistry **33:**7368–7378.
55. **Durfee, T., K. Becherer, R.-L. Chen, S. H. Yeh, Y. Yang, A. E. Kilburn, W. H. Lee, and S. J. Elledge.** 1993. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. Genes Dev. **7:**555–569.
56. **Durfee, T., B. Gorovits, C. Hensey, and W.-H. Lee.** *In vivo* and *in vitro* determined binding affinities of complexes involving the retinoblastoma protein exhibit an exponential relationship. Submitted for publication.
57. **Eisenstein, E., and H. K. Schachman.** 1989. Determining the roles of subunits in protein function, p. 135–175. *In* T. E. Creighton (ed.), Protein function: a practical approach. IRL Press, Oxford.
58. **El-Deiry, W. S., T. Tokino, V. E. Velculescu, D. B. Levy, R. Parson, J. M. Trent, D. Lin, W. E. Mercer, K. W. Kinzler, and B. Vogelstein.** 1993. *WAF1*, a potential mediator of p53 tumor suppression. Cell **75:**817–825.
59. **Ellenberger, T. E., C. J. Brandl, K. Struhl, and S. C. Harrison.** 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystal structure of the protein-DNA complex. Cell **71:**1226–1237.
60. **Evans, P. R., G. W. Farrants, and P. J. Hudson.** 1981. Phosphofructokinase: structure and control. Philos. Trans. R. Soc. London **293:**53–62.
61. **Ewen, M. E., B. Faha, E. Harlow, and D. M. Livingston.** 1992. Interaction of p107 with cyclin A independent of complex formation with viral oncoproteins. Science **255:**85–87.
62. **Faha, B., M. E. Ewen, L.-H. Tsai, D. M. Livingston, and E. Harlow.** 1992. Interaction between human cyclin A and adenovirus E1A-associated p107 protein. Science **255:**87–90.
63. **Fahien, L. A., and S. E. Smith.** 1974. The enzyme-enzyme complex of transaminase and glutamate dehydrogenase. J. Biol. Chem. **249:**2696–2703.
64. **Felder, S., M. Zhou, P. Hu, J. Urena, A. Ullrich, M. Chaudhuri, S. E. Shoelson, and J. Schlessinger.** 1993. SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange. Mol. Cell. Biol. **13:**1449–1455.
65. **Fields, S., and O.-K. Song.** 1989. A novel genetic system to detect protein-protein interactions. Nature (London) **340:**245–246.
66. **Fields, S., and R. Sternglanz.** 1994. The two-hybrid system: An assay for protein-protein interactions. Trends Genet. **10:**286–292.
67. **Finlay, C. A., P. W. Hinds, and A. J. Levine.** 1989. The p53 proto-oncogene can act as a suppressor of transformation. Cell **57:**1083–1093.
68. **Flynn, D. C., T.-H. Leu, A. B. Reynolds, and J. T. Parson.** 1993. Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated $pp60^{src}$ substrate. Mol. Cell. Biol. **13:**7892–7900.
69. **Formosa, T., J. Barry, B. M. Alberts, and J. Greenblatt.** 1991. Using protein affinity chromatography to probe structure of protein machines. Methods Enzymol. **208:**24–45.
70. **Friedrich, P.** 1984. Supramolecular enzyme organization: quarternary structure and beyond. Pergamon Press, Oxford.
71. **Friguet, B., A. F. Chaffotte, L. Djavadi-Ohaniance, and M. Goldberg.** 1985. Measurements of the true affinity constant in solution of antigen-antibody complexes. B1 enzyme-linked immunosorbent assay. J. Immunol. Methods **77:**305–319.
72. **Gegner, J. A., and F. W. Dahlquist.** 1991. Signal transduction in bacteria: CheW forms a reversible complex with the protein kinase CheA. Proc. Natl. Acad. Sci. USA **88:**750–754.
73. **Geiser, J. R., H. A. Sundberg, B. H. Chang, E. G. D. Muller, and T. N. Davis.** 1993. The essential mitotic target of calmodulin is the 110-kilodalton component of the spindle pole body in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **13:**7913–7924.
74. **Germino, F. J., Z. X. Wang, and S. M. Weissman.** 1993. Screening for *in vivo* protein-protein interactions. Proc. Natl. Acad. Sci. USA **90:**933–937.
75. **Gilbert, G. A., and G. L. Kellett.** 1971. Interacting systems of the type A+B=C: ovalbumin and myoclobin. J. Biol. Chem. **246:**6079–6086.
76. **Gill, S. C., S. E. Weitzel, and P. H. Von Hippel.** 1991. *Escherichia coli* σ^{70} and NusA proteins. 1. Binding interactions with core RNA polymerase in solution and within the transcription complex. J. Mol. Biol. **220:**307–324.
77. **Glenney, J. R., and K. Weber.** 1993. Detection of calmodulin-binding polypeptides separated in SDS-polyacrylamide gels by a sensitive [^{125}I]calmodulin gel overlay assay. Methods Enzymol. **102:**204–210.
78. **Goodsell, D. S., and A. J. Olson.** 1993. Soluble proteins: size, shape and function. Trends Biochem. Sci. **18:**65–68.
79. **Gorlich, D., E. Hartmann, S. Prehn, and T. A. Rapoport.** 1992. A protein of the endoplasmic reticulum involved early in polypeptide translocation. Nature (London) **357:**47–52.
80. **Greenblatt, J., and J. Li.** 1981. The *nusA* gene protein of *Escherichia coli*. J. Mol. Biol. **147:**11–23.
81. **Gross, M., and T. Wallimann.** 1993. Kinetics of assembly and dissociation of the mitochondrial creatine kinase octamer: a fluorescence study. Biochemistry **32:**13933–13940.
82. **Hadwiger, J. A., C. Wittenberg, H. E. Richardson, M. D. Lopes, and S. I. Reed.** 1989. A family of cyclin homologs that control the G1 phase in yeast. Proc. Natl. Acad. Sci. USA **86:**6255–6259.
83. **Hannon, G. J., D. Demetrick, and D. Beach.** 1993. Isolation of the Rb-related p130 through its interaction with CDK2 and cyclins. Genes Dev. **7:**2378–2391.
84. **Harlow, E., P. Whyte, B. R. Franza, and C. Schley.** 1986. Association of adenovirus early-region 1A proteins with cellular polypeptides. Mol. Cell. Biol. **6:**1579–1589.
85. **Harper, J. W., G. R. Adami, N. Wei, K. Keyomarsi, and S. J. Elledge.** 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell **75:**805–816.
86. **Hartman, P. E., and J. R. Roth.** 1973. Mechanisms of suppression. Adv. Genet. **17:**1–105.
87. **Hartmann, E., M. Wiedmann, and T. A. Rapoport.** 1989. A membrane component of the endoplasmic reticulum that may be essential for protein translocation. EMBO J. **8:**2225–2229.
88. **Herberg, F. W., W. R. G. Dostmann, M. Zorn, S. J. Davis, and S. S. Taylor.** 1994. Crosstalk between domains in the regulatory subunit of cAMP-dependent protein kinase: Influence of amino terminus on cAMP binding and

holoenzyme formation. Biochemistry **33:**7485–7494.
89. **Herberg, F. W., and S. S. Taylor.** 1993. Physiological inhibitors of the catalytic subunit of cAMP-dependent protein kinase: effect of MgATP on protein-protein interactions. Biochemistry **32:**14015–14022.
90. **Herskowitz, I.** 1987. Functional inactivation of genes by dominant negative mutations. Nature (London) **329:**219–222.
91. **Heyduk, T., J. C. Lee, Y. W. Ebright, E. E. Blatter, Y. Zhou, and R. H. Ebright.** 1993. CAP interacts with RNA polymerase in solution in the absence of promoter DNA. Nature (London) **364:**548–549.
92. **Hiebert, S. W., S. P. Chellappan, J. M. Horowitz, and J. R. Nevins.** 1992. The interaction of RB with E2F coincides with an inhibition of the transcriptional activity of E2F. Genes Dev. **6:**177–185.
93. **Hieter, P., C. Mann, M. Snyder, and R. W. Davis.** 1985. Mitotic stability of yeast chromosomes: a colony color assay that measures nondisjunction and chromosome loss. Cell **40:**381–392.
94. **High, S., B. Bartoglio, D. Gorlich, S. S. L. Andersen, A. J. Ashford, A. Giner, E. Hartmann, S. Prehn, T. A. Rapoport, B. Dobberstein, and J. Brunner.** 1993. Site-specific photocross-linking reveals that Sec61p and TRAM contact different regions of a membrane-inserted signal sequence. J. Biol. Chem. **368:**28745–28751.
95. **Hill, R. L., and K. Brew.** 1975. Lactose synthetase. Adv. Enzymol. **43:**411–490.
96. **Hodgkin, J., K. Kondo, and R. H. Waterston.** 1987. Suppression in the nematode *Caenorhabditis elegans*. Trends Genet. **3:**325–329.
97. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell **46:**885–894.
98. **Hu, J. C., E. K. O'Shea, P. S. Kim, and R. T. Sauer.** 1990. Sequence requirements for coiled-coils: analysis with λ repressor-GCN4 leucine zipper fusions. Science **250:**1400–1403.
99. **Hu, P., B. Margolis, E. Y. Skolnik, R. Lammers, A. Ullrich, and J. Schlessinger.** 1992. Interaction of phosphatidylinositol 3-kinase-associated p85 with epidermal growth factor and platelet-derived growth factor receptors. Mol. Cell. Biol. **12:**981–990.
100. **Huang, S., W.-H. Lee, and E. Y.-H. P. Lee.** 1991. A cellular protein that competes with SV40 T antigen for binding to the retinoblastoma gene product. Nature (London) **350:**160–162.
101. **Huffaker, T. C., M. A. Hoyt, and D. Botstein.** 1987. Genetic analysis of the yeast cytoskeleton. Annu. Rev. Genet. **21:**259–284.
102. **Hummel, J. P., and W. J. Dreyer.** 1962. Measurement of protein-binding phenomena by gel filtration. Biochim. Biophys. Acta **63:**530–532.
103. **Hurley, J. B., and L. Stryer.** 1992. Purification and characterization of the γ regulatory subunit of the cyclic GMP phosphodiesterase from retinal rod outer segments. J. Biol. Chem. **257:**11094–11099.
104. **Il'lchev, A. A., O. O. Minenkova, S. I. Tatkov, N. N. Karpyshev, A. M. Eroshkin, V. A. Petrenko, and L. S. Sandakhchiev.** 1989. M13 filamentous bacteriophage in protein engineering. Dokl. Akad. Nauk SSSR **307:**431–433.
105. **Iwabuchi, K., B. Li, P. L. Bartel, and S. Fields.** 1993. Use of the two-hybrid system to identify the domain of p53 involved in oligomerization. Oncogene **8:**1693–1696.
106. **James, P., M. Inui, M. Tada, M. Chiesi, and E. Carafoli.** 1989. Nature and site of phospholamban regulation of the Ca^{2+} pump of sarcoplasmic reticulum. Nature (London) **342:**90–92.
107. **Jarvik, J., and D. Botstein.** 1975. Conditional-lethal mutations that suppress genetic defects in morphogenesis by altering structural proteins. Proc. Natl. Acad. Sci. USA **72:**2738–2742.
108. **Jarvis, T. C., D. M. Ring, S. S. Daube, and P. H. Von Hippel.** 1990. "Macromolecular crowding": thermodynamic consequences for protein-protein interactions within the T4 DNA replication complex. J. Biol. Chem. **265:**15160–15167.
109. **Jonsson, U., L. Fagerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Lofas, B. Persson, H. Roos, I. Ronnberg, S. Sjolander, E. Stenberg, R. Stahlberg, C. Urbaniczky, H. Ostlin, and M. Malmqvist.** 1991. Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. BioTechniques. **11:**620–627.
110. **Kang, A. S., C. F. Barbas, K. D. Janda, S. J. Benkovic, and R. A. Lerner.** 1991. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. Proc. Natl. Acad. Sci. USA **88:**4363–4366.
111. **Karlsson, R., A. Michaelsson, and L. Mattsson.** 1991. Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system. J. Immunol. Methods **145:**229–240.
112. **Keegan, L., G. Gill, and M. Ptashne.** 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science **231:**699–704.
113. **Kellogg, D. R., C. M. Field, and B. M. Alberts.** 1989. Identification of microtubule-associated proteins in the centrosome, spindle, and kinetochore of the early *Drosophila* embryo. J. Cell Biol. **109:**2977–2991.
114. **Kerppola, T. K., and T. Curran.** 1991. Fos-Jun heterodimers and Jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. Cell **66:**317–326.
115. **Kim, Y. T., S. Tabor, J. E. Churchich, and C. C. Richardson.** 1992. Interactions of gene 2.5 protein and DNA polymerase of bacteriophage T7. J. Biol. Chem. **267:**15032–15040.
116. **Klotz, I. M., D. W. Darnall, and N. R. Langerman.** 1975. Quarternary structure of proteins, p. 293–411. *In* H. Neurath and R. L. Hill (ed.), The proteins. Academic Press, Inc., New York.
117. **Koch, C. A., D. Anderson, M. F. Moran, C. Willis, and T. Pawson.** 1991. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science **252:**668–674.
118. **Kohda, D., H. Hatanaka, M. Odaka, V. Mandiyan, A. Ullrich, J. Schlessinger, and F. I. Nagaki.** 1993. Solution structure of the SH3 domain of phospholipase C-g. Cell **72:**953–960.
119. **Koshland, D., J. C. Kent, and L. H. Hartwell.** 1985. Genetic analysis of the mitotic transmission of minochromosomes. Cell **40:**393–403.
120. **Koyama, S., H. Yu, D. C. Dalgarno, T. B. Shin, L. D. Zydowsky, and S. L. Schreiber.** 1993. Structure of the PI3K SH3 domain and analysis of the SH3 family. Cell **72:**945–952.
121. **Krieg, U. C., A. E. Johnson, and P. Walter.** 1989. Protein translocation across the endoplasmic reticulum membrane: identification by photocross-linking of a 39-kD integral membrane glycoprotein as part of a putative translocation tunnel. J. Cell Biol. **109:**2033–2043.
122. **Kukuruzinska, M. A., B. W. Turner, G. K. Ackers, and S. Roseman.** 1984. Subunit association of enzyme I of the *Salmonella typhimurium* phosphoenolpyruvate:glycose phosphotransferase system. J. Biol. Chem. **259:**11679–11681.
123. **Kwon, O.-S., and J. E. Churchich.** 1994. Interaction of 70-kDa heat shock cognate protein with peptides and *myo*-inositol monophosphatase. J. Biol. Chem. **269:**266–271.
124. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240:**1759–1764.
125. **Lazarides, E., and U. Lindberg.** 1974. Actin is the naturally occurring inhibitor of deoxyribonuclease I. Proc. Natl. Acad. Sci. USA **71:**4742–4746.
126. **Lee, F. S., D. S. Auld, and B. L. Vallee.** 1989. Tryptophan fluorescence as a probe of placental ribonuclease inhibitor binding to angiogenin. Biochemistry **28:**219–224.
126a. **Lee, F. S., R. Shapiro, and B. L. Vallee.** 1989. Tight-binding inhibition of angiogenin and ribonuclease A by placental ribonuclease inhibitor. Biochemistry **28:**225–230.
127. **Li, B., and S. Fields.** 1993. Identification of mutations in p53 that affect its binding to SV40 large T antigen by using the yeast two-hybrid system. FASEB J. **7:**957–963.
128. **Li, N., A. Batzer, R. Daly, V. Yajnik, E. Skolnik, P. Chardin, D. Bar-Sagi, B. Margolis, and J. Schlessinger.** 1993. Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinases to Ras signalling. Nature (London) **363:**85–88.
129. **Li, S., and J. M. Sedivy.** 1993. Raf-1 protein kinase activates the NF-κB transcription factor by dissociating the cytoplasmic NF-kB-IkB complex. Proc. Natl. Acad. Sci. USA **90:**9247–9251.
130. **Liu, X., L. E. M. Marengere, C. A. Koch, and T. Pawson.** 1993. The v-Src SH3 domain binds phosphatidylinositol 3′-kinase. Mol. Cell. Biol. **13:**5225–5232.
131. **Lowman, H. B., S. H. Bass, N. Simpson, and J. A. Wells.** 1991. Selecting high-affinity binding proteins by monovalent phage display. Biochemistry **30:**10832–10838.
132. **Ludlow, J. W., J. A. DeCaprio, C.-M. Huang, W.-H. Lee, E. Paucha, and D. M. Livingston.** 1989. SV40 large T antigen binds preferentially to an underphosphorylated member of the retinoblastoma susceptibility gene product family. Cell **56:**57–65.
133. **Ludlow, J. W., C. L. Glendening, D. M. Livingston, and J. A. DeCaprio.** 1993. Specific enzymatic dephosphorylation of the retinoblastoma protein. Mol. Cell. Biol. **13:**367–372.
134. **Ma, J., and M. Ptashne.** 1987. A new class of yeast transcriptional activators. Cell **51:**113–119.
135. **MacGregor, P. F., C. Abate, and T. Curran.** 1990. Direct cloning of leucine zipper proteins: jun binds cooperatively to the CRE with CRE-BP1. Oncogene **5:**451–458.
136. **Malmqvist, M.** 1993. Biospecific interaction analysis using biosensor technology. Nature (London) **361:**186–187.
137. **Marengere, L. E. M., Z. Songyang, G. D. Gish, M. D. Schaller, J. T. Parsons, M. J. Stern, L. C. Cantley, and T. Pawson.** 1994. SH2 domain specificity and activity modified by a single residue. Nature (London) **369:**502–505.
138. **Mason, S. W., J. Li, and J. Greenblatt.** 1992. Direct interaction between two *Escherichia coli* transcription antitermination factors, NusB and ribosomal protein S10. J. Mol. Biol. **223:**55–66.
139. **Mayer, B. J., P. K. Jackson, and D. Baltimore.** 1991. The noncatalytic src homology region 2 segment of *abl* tyrosine kinase binds to tyrosine-phosphorylated cellular proteins with high affinity. Proc. Natl. Acad. Sci. USA **88:**627–631.
140. **Mayer, B. J., P. K. Jackson, R. A. Van Etten, and D. Baltimore.** 1992. Point mutations in the *abl* SH2 domain coordinately impair phosphotyrosine binding in vitro and transforming activity in vivo. Mol. Cell. Biol. **12:**609–618.

141. **McCafferty, J., A. D. Griffiths, G. Winter, and D. J. Chiswell.** 1990. Phage antibodies: filamentous phage displaying antibody variable domains. Nature (London) **348:**552–554.
142. **Meeks-Wagner, D., and L. H. Hartwell.** 1986. Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell **44:**43–52.
143. **Miller, J. H.** 1978. The *lacI* gene: its role in *lac* operon control and its use as a genetic system, p. 31–88. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
144. **Miller, K. G., and B. M. Alberts.** 1989. F-actin affinity chromatography: technique for isolating previously unidentified actin-binding proteins. Proc. Natl. Acad. Sci. USA **86:**4808–4812.
145. **Miller, K. G., C. M. Field, B. M. Alberts, and D. R. Kellogg.** 1991. Use of actin filament and microtubule affinity chromatography to identify proteins that bind to the cytoskeleton. Methods Enzymol. **196:**303–319.
146. **Mills, J. S., M. P. Walsh, K. Nemcek, and J. D. Johnson.** 1988. Biologically active fluorescent derivatives of spinach calmodulin that report calmodulin target protein binding. Biochemistry **27:**991–996.
147. **Mita, S., A. Tominaga, Y. Hitoshi, K. Sakamoto, T. Honjo, M. Akagi, Y. Kikuchi, N. Yamaguchi, and K. Takatsu.** 1989. Characterization of high-affinity receptors for interleukin 5 on interleukin 5-dependent cell lines. Proc. Natl. Acad. Sci. USA **86:**2311–2315.
148. **Mitchell, D. A., T. K. Marshall, and R. J. Deschenes.** 1993. Vectors for the inducible overexpression of glutathione S-transferase fusion proteins in yeast. Yeast **9:**715–723.
149. **Mitsuzawa, H., I. Uno, T. Oshima, and T. Ishikawa.** 1989. Isolation and characterization of temperature-sensitive mutations in the *RAS2* and *CYR!* genes of *Saccharomyces cerevisiae*. Genetics **123:**739–748.
150. **Moerman, D. G., G. M. Benian, R. J. Barstead, L. A. Schriefer, and R. H. Waterston.** 1988. Identification and intracellular localization of the *unc-22* gene product of *Caenorhabditis elegans*. Genes Dev. **2:**93–105.
151. **Moerman, D. G., S. Plurad, and R. H. Waterston.** 1982. Mutations in the *unc-54* myosin heavy chain gene of *Caenorhabditis elegans* that alter contractility but not muscle structure. Cell **29:**773–781.
152. **Moir, D., S. E. Stewart, B. C. Osmond, and D. Botstein.** 1982. Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. Genetics **100:**547–563.
153. **Monod, J., J. Wyman, and J.-P. Changeux.** 1965. On the nature of allosteric transitions: a plausible model. J. Mol. Biol. **12:**88–118.
154. **Musacchio, A., M. Noble, R. Pauptit, R. Wierenga, and M. Saraste.** 1992. Crystal structure of a Src-homology 3 (SH3) domain. Nature (London) **359:**851–855.
155. **Musch, A., M. Wiedmann, and T. A. Rapoport.** 1992. Yeast Sec proteins interact with polypeptides traversing the endoplasmic reticulum membrane. Cell **69:**343–352.
156. **Nakai, H., and C. C. Richardson.** 1990. The gene 1.2 protein of bacteriophage T7 interacts with the *Escherichia coli* dGTP triphosphohydrolase to form a GTP-binding protein. J. Biol. Chem. **265:**4411–4419.
157. **Nefsky, B., and A. Bretscher.** 1992. Yeast actin is relatively well behaved. Eur. J. Biochem. **206:**949–955.
158. **Nelson, R. M., and G. L. Long.** 1991. Solution-phase equilibrium binding interaction of human protein S with C4b-binding protein. Biochemistry **30:**2384–2390.
159. **Nogi, Y., K. Matsumoto, A. Toh-e, and Y. Oshima.** 1977. Regulation of expression of the galactose gene cluster in *Saccharomyces cerevisiae*. II. The isolation and dosage effect of the regulatory gene Gal80. Mol. Gen. Genet. **195:**29–34.
160. **Novick, P., B. C. Osmond, and D. Botstein.** 1989. Suppressors of yeast actin mutations. Genetics **121:**659–674.
161. **O'Shea, E. K., J. D. Klemm, P. S. Kim, and T. Alber.** 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science **254:**539–544.
162. **O'Shea, E. K., R. Rutkowski, and P. S. Kim.** 1992. Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell **68:**699–708.
163. **O'Shea, E. K., R. Rutkowski, W. F. Stafford III, and P. S. Kim.** 1989. Preferential heterodimer formation by isolated leucine zippers from fos and jun. Science **245:**646–648.
164. **Otto-Bruc, A., B. Antonny, T. M. Vuong, P. Chardin, and M. Chabre.** 1993. Interaction between retinal cyclic GMP phosphodiesterase inhibitor and transducin: kinetics and affinity studies. Biochemistry **32:**8636–8645.
165. **Overduin, M., C. B. Rios, B. J. Mayer, D. Baltimore, and D. Cowburn.** 1992. Three dimensional solution structure of the scr homology 2 domain of c-abl. Cell **70:**697–704.
166. **Panayotou, G., G. Gish, P. End, O. Truong, I. Gout, R. Dhand, M. J. Fry, I. Hiles, T. Pawson, and M. D. Waterfield.** 1993. Interactions between SH2 domains and tyrosine-phosphorylated platelet-derived growth factor β-receptor sequences: analysis of kinetic parameters by a novel biosensor-based approach. Mol. Cell. Biol. **13:**3567–3576.
167. **Pannekoek, H., M. VanMeijer, R. R. Schleef, D. J. Loskutoff, and C. F. Barbas.** 1993. Functional display of human plasminogen-activator inhibitor 1 (PAI-1) on phages: novel perspectives for structure-function analysis by error-prone DNA synthesis. Gene **128:**135–140.
168. **Pawson, T., and G. D. Gish.** 1992. SH2 and SH3 domains: from structure to function. Cell **71:**359–362.
169. **Phizicky, E. M., R. C. Schwartz, and J. Abelson.** 1986. *Saccharomyces cerevisiae* tRNA ligase. J. Biol. Chem. **261:**2978–2986.
170. **Pinkney, M., and J. G. Hoggett.** 1988. Binding of the cyclic AMP receptor protein of *Escherichia coli* to RNA polymerase. Biochem. J. **250:**897–902.
171. **Porpaczy, Z., B. Sumegi, and I. Alkonyi.** 1983. Association between the a-ketoglutarate dehydrogenase complex and succinate thiokinase. Biochim. Biophys. Acta **749:**172–179.
172. **Prasad, K. V. S., O. Janssen, R. Kapeller, M. Raab, L. C. Cantley, and C. E. Rudd.** 1993. Src-homology 3 domain of protein kinase $p59^{fyn}$ mediates binding to phosphatidylinositol 3-kinase in T cells. Proc. Natl. Acad. Sci. USA **90:**7366–7370.
173. **Prasad, K. V. S., R. Kapeller, O. Janssen, H. Repke, J. S. Duke-Cohan, L. C. Cantley, and C. E. Rudd.** 1993. Phosphatidylinositol (PI) 3-kinase and PI 4-kinase binding to the CD4-p56lck SH3 domain binds to PI 3-kinase but not PI 4-kinase. Mol. Cell Biol. **13:**7708–7717.
174. **Prelich, G., C.-K. Tan, M. Kostura, M. B. Mathews, A. G. So, K. M. Downey, and B. Stillman.** 1989. Functional identity of proliferating cell nuclear antigen and a DNA polymerase-δ auxiliary protein. Nature (London) **326:**517–520.
175. **Rabidran, S. K., R. I. Haroun, J. Clos, J. Wisniewski, and C. Wu.** 1993. Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. Science **259:**230–234.
176. **Raff, E. C., and M. T. Fuller.** 1984. Genetic analysis of microtubule function in *Drosophila*, p. 293–304. *In* G. G. Borisy, D. W. Cleveland, and D. B. Murphy (ed.), Molecular biology of the cytoskeleton. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
177. **Ratner, D.** 1974. The interaction of bacterial and phage proteins with immobilized *Escherichia coli* RNA polymerase. J. Mol. Biol. **88:**373–383.
178. **Rauhut, R., P. R. Green, and J. Abelson.** 1990. Yeast tRNA-splicing endonuclease is a heterotrimeric enzyme. J. Biol. Chem. **265:**18180–18184.
179. **Ray, S. K., M. Arroyo, S. Bigchi, and P. Raychaudhuri.** 1992. Identification of a 60-kilodalton Rb-binding protein, RMP60, that allows the Rb-E2F complex to bind DNA. Mol. Cell. Biol. **12:**4327–4333.
180. **Redl, B., J. Walleczek, M. Stoffler-Meilicke, and G. Stoffler.** 1989. Immunoblotting analysis of protein-protein crosslinks within the 50S ribosomal subunit of *Escherichia coli*: a study using dimethylsuberimidate as crosslinking reagent. Eur. J. Biochem. **181:**351–356.
181. **Ren, R., B. J. Mayer, P. Cicchetti, and D. Baltimore.** 1993. Identification of a ten-amino acid proline-rich SH3 binding site. Science **259:**1157–1161.
182. **Rine, J.** 1991. Gene overexpression in studies of *Saccharomyces cerevisiae*. Methods Enzymol. **194:**239–251.
183. **Rivas, G., K. C. Ingham, and A. P. Minton.** 1992. Calcium-linked self-association of human complement Cls. Biochemistry **31:**11707–11712.
184. **Rivas, G., K. C. Ingham, and A. P. Minton.** 1994. Ca^{2+}-linked association of human complement Cls and Clr. Biochemistry **33:**2341–2348.
185. **Rivas, G., and A. P. Minton.** 1993. New developments in the study of biomolecular associations via sedimentation equilibrium. Trends Biochem. Sci. **18:**284–287.
186. **Roberts, B. L., W. Markland, A. C. Ley, R. B. Kent, D. W. White, S. K. Gutterman, and R. C. Ladner.** 1992. Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. Proc. Natl. Acad. Sci. USA **89:**2429–2433.
187. **Rozakis-Adcock, M., R. Fernley, J. Wade, T. Pawson, and D. Bowtell.** 1993. The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. Nature (London) **363:**83–85.
188. **Russell, R. B., J. Breed, and G. J. Barton.** 1992. Conservation analysis and structure prediction of the SH2 family of phosphotyrosine binding domains. FEBS Lett. **304:**15–20.
189. **Sadowski, I., J. C. Stone, and T. Pawson.** 1986. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus $P130^{gag\text{-}fps}$. Mol. Cell. Biol. **6:**4396–4408.
190. **Salminen, A., and P. J. Novick.** 1987. A *ras*-like protein is required for a post-Golgi event in yeast secretion. Cell **49:**527–538.
191. **Sanders, S. L., K. M. Whitfield, J. P. Vogel, M. D. Rose, and R. W. Schekman.** 1992. Sec61p and BiP directly facilitate polypeptide translocation into the ER. Cell **69:**353–365.
192. **Sarkar, F. H., and S. L. Gupta.** 1984. Receptors for human γ interferon: binding and crosslinking of ^{125}I-labeled recombinant human γ interferon to receptors on WISH cells. Proc. Natl. Acad. Sci. USA **81:**5160–5164.
193. **Sauer, R. T.** 1979. Molecular characterization of the repressor and its gene cI. Ph.D. thesis. Harvard University, Cambridge, Mass.
194. **Sawyers, C. L., W. Callahan, and O. N. Witte.** 1992. Dominant negative MYC blocks transformation by ABL oncogenes. Cell **70:**901–910.
195. **Scherer, P. E., U. C. Manning-Krieg, P. Jeno, G. Schatz, and M. Horst.** 1992. Identification of a 45-kDa protein at the protein import site of the yeast mitochondrial inner membrane. Proc. Natl. Acad. Sci. USA **89:**11930–11934.
196. **Schmidt-Dorr, T., P. Oertel-Buchheit, C. Pernelle, L. Bracco, M. Schnarr, and M. Granger-Schnarr.** 1991. Construction, purification and character-

ization of a hybrid protein comprising the DNA binding domain of the LexA repressor and the jun leucine zipper: a circular dichroism and mutagenesis study. Biochemistry **30:**9657–9664.

197. **Schuster, S. C., R. V. Swanson, L. A. Alex, R. B. Bourret, and M. I. Simon.** 1993. Assembly and function of a quarternary signal transduction complex monitored by surface plasmon resonance. Nature (London) **365:**343–346.
198. **Scott, J. K., and G. P. Smith.** 1990. Searching for peptide ligands with an epitope library. Science **249:**386–390.
199. **Shapiro, R., and B. L. Vallee.** 1991. Interaction of human placental ribonuclease with placental ribonuclease inhibitor. Biochemistry **30:**2246–2255.
200. **Sikela, J. M., and W. E. Hahn.** 1987. Screening an expression library with a ligand probe: isolation and sequence of a cDNA corresponding to a brain calmodulin-binding protein. Proc. Natl. Acad. Sci. USA **84:**3038–3042.
201. **Skolnik, E. Y., B. Margolis, M. Mohammadi, E. Lowenstein, R. Fischer, A. Drepps, A. Ullrich, and J. Schlessinger.** 1991. Cloning of PI3 kinase-associated p85 utilizing a novel method for expression/cloning of target proteins for receptor tyrosine kinases. Cell **65:**83–90.
202. **Smith, D. B., and K. S. Johnson.** 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. Gene **67:**31–40.
203. **Smith, G. P.** 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science **228:**1315–1317.
204. **Sollner, T., J. Rassow, M. Wiedmann, J. Schlossmann, P. Keil, W. Neupert, and N. Pfanner.** 1992. Mapping of the protein import machinery in the mitochondrial outer membrane by crosslinking of translocation intermediates. Nature (London) **355:**84–87.
205. **Songyang, Z., S. E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W. G. Haser, F. King, T. Robertys, S. Ratnofsky, R. J. Lechleider, B. G. Neel, R. B. Birge, J. E. Fajardo, M. M. Chou, H. Hanafusa, B. Schaffhausen, and L. C. Cantley.** 1993. SH2 domains recognize specific phosphopeptide sequences. Cell **72:**767–778.
206. **Sopta, M., R. W. Carthew, and J. Greenblatt.** 1985. Isolation of three proteins that bind to mammalian RNA polymerase II. J. Biol. Chem. **260:**10353–10360.
207. **Soutar, A. K., and D. P. Wade.** 1990. Ligand blotting, p. 55–76. *In* T. Creighton (ed.), Protein function: a practical approach. IRL Press, Oxford.
208. **Srere, P. A.** 1987. Complexes of sequential metabolic enzymes. Annu. Rev. Biochem. **56:**89–124.
209. **Steitz, T. A., W. F. Anderson, R. J. Fletterick, and C. M. Anderson.** 1977. High resolution crystal structures of yeast hexokinase complexes with substrates, activators, and inhibitors. J. Biol. Chem. **252:**4494–4500.
210. **Stephen, C. W., and D. P. Lane.** 1992. Mutant conformation of p53: precise epitope mapping using a filamentous phage epitope library. J. Mol. Biol. **225:**577–583.
211. **Stevens, F. J.** 1986. Analysis of protein-protein interaction by simulation of small-zone size-exclusive chromatography: application to an antibody-antigen association. Biochemistry **25:**981–993.
212. **Stoffler, G., B. Redl, J. Walleczek, and M. Stoffler-Meilicke.** 1988. Identification of protein-protein cross-links within the *Escherichia coli* ribosome by immunobloting techniques. Methods Enzymol. **164:**64–76.
213. **Susskind, M. M., and P. Youderian.** 1983. Bacteriophage P22 antirepressor and its control, p. 347–363. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
214. **Tompa, P., J. Batke, J. Ovadi, G. R. Welch, and P. A. Srere.** 1987. Quantitation of the interaction between citrate synthase and malate dehydrogenase. J. Biol. Chem. **282:**6089–6092.
215. **Traut, R. R., C. Casiano, and N. Zecherle.** 1989. Crosslinking of protein subunits and ligands by the introduction of disulphide bonds, p. 101–133. *In* T. E. Creighton (ed.), Protein function: a practical approach. IRL Press, Oxford.
216. **Truant, R., H. Xiao, C. J. Ingles, and J. Greenblatt.** 1993. Direct interaction between the transcriptional activation domain of human p53 and the TATA box-binding protein. J. Biol. Chem. **268:**2284–2287.
217. **Uchiumi, T., A. J. Wahba, and R. R. Traut.** 1987. Topography and stoichiometry of acidic proteins in large ribosomal subunits from *Artemia salina* as determined by crosslinking. Proc. Natl. Acad. Sci. USA **84:**5580–5584.
218. **Urena, J., et al.** 1993. Unpublished data.
219. **Vallee, R. B., and C. A. Collins.** 1986. Purification of microtubules and microtubule-associated proteins from sea urchin eggs and cultured mammalian cells using Taxol, and use of exogenous taxol-stabilized brain microtubules for purifying microtubule-associated proteins. Methods Enzymol. **134:**116–121.
220. **Van Aelst, L., M. Barr, S. Marcus, A. Polverino, and M. Wigler.** 1993. Complex formation between RAS and RAF and other protein kinases. Proc. Natl. Acad. Sci. USA **90:**6213–6217.
221. **Vincent, J.-P., and M. Lazdunski.** 1972. Trypsin-pancreatic trypsin inhibitor association. Dynamics of the interaction and role of disulfide bridges. Biochemistry **11:**2967–2977.

221a. **Vinson, C. R., T. Hai, and S. M. Boyd.** 1993. Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: prediction and rational design. Genes Dev. **7:**1047–1058.

222. **Vinson, C. R., K. L. LaMarco, P. F. Johnson, W. H. Landschulz, and S. L. McKnight.** 1988. In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev. **2:**807–806.
223. **Vojtek, A. B., S. M. Hollenberg, and J. A. Cooper.** 1993. Mammalian Ras interacts directly with the serine/threonine kinase Raf. Cell **74:**205–214.
224. **Waksman, G., D. Kominos, S. C. Robertson, N. Pant, D. Baltimore, R. B. Birge, D. Cowburn, H. Hanafusa, B. J. Mayer, M. Overduin, M. D. Resh, C. B. Rios, L. Silverman, and J. Kuriyan.** 1992. Crystal structure of the phosphotyrosine recognition domain SH2 of V-src complexed with tyrosine-phosphorylated peptides. Nature (London) **358:**646–653.
225. **Waksman, G., S. E. Shoelson, N. Pant, D. Cowburn, and J. Kuriyan.** 1993. Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. Cell **72:**779–790.
226. **Waldron, C., and F. Lacroute.** 1975. Effect of growth rate on the amounts of ribosomal and transfer ribonucleic acids in yeast. J. Bacteriol. **122:**855–865.
227. **Warne, P. H., P. R. Viciana, and J. Downward.** 1993. Direct interaction of Ras and the amino-terminal region of Raf-1 *in vitro*. Nature (London) **364:**352–355.
228. **Weber, J., R. S.-F. Lee, S. Wilke-Mounts, E. Grell, and A. E. Senior.** 1993. Combined application of site-directed mutagenesis, 2-azido-ATP labeling, and *lin*-benzo-ATP binding to study the noncatalytic sites of *Escherichia coli* F_1-ATPase. J. Biol. Chem. **268:**6241–6247.
229. **Weiel, J., and J. W. B. Hershey.** 1981. Fluorescence polarization studies of the interaction of *Escherichia coli* protein synthesis initiation factor 3 with 30S ribosomal subunits. Biochemistry **20:**5859–5865.
230. **Weiel, J., and J. W. B. Hershey.** 1982. The binding of fluorescein-labeled protein synthesis initiation factor 2 to *Escherichia coli* 30 S ribosomal subunits determined by fluorescence polarization. J. Biol. Chem. **257:**1215–1220.
231. **Wells, J. A., and H. B. Lowman.** 1992. Rapid evolution of peptide and protein binding properties *in vitro*. Curr. Opin. Struct. Biol. **2:**355–362.
232. **Weng, Z., J. A. Taylor, C. E. Turner, J. S. Brugge, and C. Seidel-Dugan.** 1993. Detection of Src homology 3-binding proteins, including paxillin, in normal and v-Src-transformed Balb/c 3T3 cells. J. Biol. Chem. **268:**14956–14963.
233. **Wensel, T. G., and L. Stryer.** 1986. Reciprocal control of retinal rod cyclic GMP phosphodiesterase by its γ subunit and transducin. Proteins Struct. Funct. Genet. **1:**90–99.
234. **Wensel, T. G., and S. Stryer.** 1990. Activation mechanism of retinal rod cyclic GMP phosphodiesterase probed by fluorescein-labeled inhibitory subunit. Biochemistry **29:**2155–2161.
235. **White, R. J., and S. P. Jackson.** 1992. The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. Trends Genet. **8:**284–288.
236. **Whyte, P., K. J. Buchkovich, J. M. Horowitz, S. H. Friend, M. Raybuck, R. A. Weinberg, and E. Harlow.** 1988. Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. Nature (London) **334:**124–129.
237. **Whyte, P., N. M. Williamson, and E. Harlow.** 1989. Cellular targets for transformation by the adenovirus E1A proteins. Cell **56:**67–75.
238. **Wiedmann, M., D. Goerlich, E. Hartmann, T. V. Kurzchalia, and T. A. Rapoport.** 1989. Photocrosslinking demonstrates proximity of a 34 kDa membrane protein to different portions of preprolactin during translocation through the endoplasmic reticulum. FEBS Lett. **257:**263–268.
239. **Wiedmann, M., T. V. Kurzchalia, E. Hartmann, and T. A. Rapoport.** 1987. A signal sequence receptor in the endoplasmic reticulum membrane. Nature (London) **328:**830–833.
240. **Wiman, K. G.** 1993. The retinoblastoma gene: role in cell cycle control and cell differentiation. FASEB J. **7:**841–845.
241. **Yanofsky, C., and M. Rochmeler.** 1958. The exclusion of free indole as an intermediate in the biosynthesis of tryptophan in *Neurospora crassa*. Biochim. Biophys. Acta **28:**640–645.
242. **Yee, S.-P., and P. E. Branton.** 1985. Detection of cellular proteins associated with human adenovirus type 5 early region 1A polypeptides. Virology **147:**142–153.
243. **Yong, H., G. A. Thomas, and W. L. Peticolas.** 1993. Metabolite-modulated complex formation between α-glycerophosphate dehydrogenase and lactate dehydrogenase. Biochemistry **32:**11124–11131.
244. **Young, R., and R. Davis.** 1983. Yeast RNA polymerase II genes: Isolation with antibody probes. Science **222:**778–782.
245. **Yu, H., M. K. Rosen, T. B. Shin, C. Seidel-Dugan, J. S. Brugge, and S. L. Schreiber.** 1992. Solution structure of the SH3 domain of Src and identification of its ligand-binding site. Science **258:**1655–1668.
246. **Zebedee, S. L., C. F. Barbas III, Y.-L. Hom, R. H. Caothien, R. Graff, J. DeGraw, J. Pyati, R. LaPolla, D. R. Burton, R. A. Lerner, and G. B. Thornton.** 1992. Human combinatorial antibody libraries to hepatitis B surface antigens. Proc. Natl. Acad. Sci. USA **89:**3175–3179.
247. **Zervos, A. S., J. Gyuris, and R. Brent.** 1993. Mxi1, a protein that specifically

interacts with Max to bind Myc-Max recognition sites. Cell **72:**223–232.
248. **Zhang, X.-F., J. Settleman, J. M. Kyriakis, E. Takeuchi-Suzuki, S. J. Elledge, M. S. Marshall, J. T. Bruder, U. R. Rapp, and J. Avruch.** 1993. Normal and oncogenic p21^{ras} proteins bind to the amino-terminal regulatory domain of c-Raf-1. Nature (London) **364:**308–313.
249. **Zhou, M., S. Felder, M. Rubinstein, D. R. Hurwitz, A. Ullrich, I. Lax, and J. Schlessinger.** 1993. Real-time measurements of kinetics of EGF binding to soluble EGF receptor monomers and dimers support the dimerization model for receptor activation. Biochemistry **32:**8193–8198.
250. **Zimmerman, J. K., and G. K. Ackers.** 1971. Molecular sieve studies of interacting protein systems: behavior of small zone profiles for reversibly self-associating solutes. J. Biol. Chem. **246:**7289–7292.